Universidade de Santiago de Compostela

Grao en Matemáticas

Facultade de Matemáticas

Álxebra Lineal e Multilineal

Óscar Rivero Salgado

Santiago de Compostela, Decembro 2024

Índice xeral

Introdución

Este libro está pensado para un segundo curso de Álxebra Lineal, no que se supón que o alumnado xa está familiarizado coas nocións de espazos vectoriais e aplicacións lineais, e tamén que ten certa familiaridade traballando con matrices.
O texto adáptase ao temario da materia de *Álxebra Lineal e Multilineal*, que se imparte na Facultade de Matemáticas da Universidade de Santiago de Compostela desde o curso 2009-2010. En certo sentido, pódese entender como unha continuación de *Espazos Vectoriais e Cálculo Matricial*, especialmente nos temas que continúan o estudo das aplicacións lineais, introducindo os conceptos de diagonalización, polinomio mínimo dun endomorfismo e forma de Jordan. Ao mesmo tempo, a materia funciona tamén como unha preparación para a materia de *Xeometría Lineal*, xa que o tema de formas bilineais e cuadráticas introduce conceptos e técnicas esenciais, como as propiedades do produto escalar ou da proxección ortogonal, o teorema de clasificación das isometrías, ou o método de congruencia pivote para a clasificación de formas cuadráticas. Por último, a materia serve tamén para ir introducindo estruturas alxébricas máis abstractas que se estudarán de xeito máis sistemático no curso de *Estruturas Alxébricas*: aparecen varios exemplos de grupos e álxebras, trabállanse os aneis de polinomios sobre un corpo e ao introducir os tensores discútense por primeira vez conceptos como o de propiedade universal.

Contidos. O libro contén 5 capítulos, cada un deles correspondentes a un dos temas da materia. Ademais, consta dun Capítulo 0 no que se fai un repaso de conceptos básicos que se dan por sabidos ao longo do texto, especialmente os relativos ao espazo dual e ao espazo cociente.

- **Capítulo 1. Polinomios.** Introdúcense as propiedades principais do anel de polinomios sobre un corpo e discútense algúns resultados de irredutibilidade.

- **Capítulo 2. Aplicacións multilineais e determinantes.** Introdúcense as aplicacións multilineais, con énfase no caso das alternadas; iso serve tamén para dar un primeiro acercamento aos tensores e motivar algúns dos resultados do capítulo 5. Unha sección do capítulo dedícase ao estudo do grupo simétrico, que é un tema que pode ter interese por si mesmo máis alá da álxebra lineal. Finalmente, defínese o determinante e trabállanse as súas propiedades.

- **Capítulo 3. Estrutura das aplicacións lineais.** Trabállase a diagonalización de endomorfismos e a forma de Jordan, discutindo tamén algunhas das súas aplicacións. Preséntase igualmente o concepto de polinomio mínimo dun endomorfismo e demóstrase o teorema de Cayley–Hamilton.

- **Capítulo 4. Formas bilineais e cuadráticas.** Introdúcense as nocións básicas de formas bilineais e cuadráticas (produtos escalares e hermíticos, proxección ortogonal ou clasificacións de formas cuadráticas). Estúdanse tamén os teoremas espectrais real e complexo, as isometrías e a descomposición en valores singulares.

- **Capítulo 5. Tensores e álxebra tensorial.** Faise unha introdución aos tensores, con énfase nas propiedades dos tensores simétricos e antisimétricos, e tamén no produto exterior. Inclúense unhas seccións de carácter máis técnico nas que se define o produto tensorial de espazos vectoriais en termos dunha propiedade universal. Finalmente, danse algunhas aplicacións á teoría da integración e á física.

Tódolos capítulos se estruturan do mesmo xeito. Primeiro inclúense as diferentes seccións teóricas, nas que se presentan os resultados principais, e logo na última sección discútense diferentes problemas, todos eles resoltos, co obxectivo de favorecer o estudo autónomo por parte do lector. Neste sentido, pensamos que a obra pode ser útil tanto para a aprendizaxe autodidacta da materia como para que sirva de manual de referencia para calquera curso universitario.

Libros e textos de referencias. Algunhas partes dos capítulos 2 e 3, así como a sección de isometrías no capítulo 4 están inspiradas no libro de Castellet e Llerena [CL00]. Partes do capítulo 4, así como algúns dos exercicios doutros temas, foron collidos do texto de Axler *Linear Algebra Done Right* [A23]. O libro de Hernández [H98], outro clásico, usouse para explicar a forma de Jordan real e a diagonalización simultánea de endomorfismos. Os textos de Keith Conrad, todos eles accesibles no seu sitio web persoal, foron tamén un excelente apoio para preparar o último capítulo, especialmente [C23a] e [C23b]. Algúns textos como [LM15] ou [MS21] foron tamén empregados para consultas puntuais. Outros, como [Ro08], ofrecen unha perspectiva máis avanzada, traballando no contexto de módulos sobre ideais principais e recuperando a teoría dos espazos vectoriais como un caso particular.
Hai temas específicos para os que neste libro unicamente se da unha primeira introdución. Un deles son os usos da álxebra na resolución de ecuacións diferenciais. Hai textos introdutorios que tratan o tema en detalle, como o de Logan [L06], mentres que algúns como o Kapitula [K15] mesmo teñen como eixe central as interaccións entre as dúas disciplinas. No último capítulo preséntanse as formas diferenciais como unha das principais aplicacións dos tensores; unha introdución espléndida e accesible á materia pode atoparse no coñecido libro de Do Carmo [DC94].
Por último, a estrutura dalgunhas seccións está inspirada nos apuntamentos dos cursos de Álxebra Lineal [Q13] e Álxebra Multilineal e Xeometría [P14] que cursei na Universitat Politècnica de Catalunya nos cursos 2012-2013 e 2013-2014 cos profesores Jordi Quer e Francesc Planas, respectivamente. En particular, as demostracións do segundo teorema de descomposición, do criterio de Sylvester e do teorema espectral real seguen a primeira das referencias, mentres que partes da sección de formas cuadráticas, así como as partes do tema de tensores relativas ao produto exterior e á propiedade universal, seguen a segunda. As materias ás que fago referencia están tamén inspiradas no libro de Ferran Puerta [P05], que impartiu a materia durante varios anos. Finalmente, algúns dos exercicios son parte das colección de exames da Universitat Politècnica de Catalunya.

A álxebra lineal en galego. Ata onde sabemos, existen poucas referencias de textos de álxebra en galego. Un deles é *Álxebra Linear*, do profesor Ramón González Rodríguez

[G22], que se centra en aspectos máis introdutorios da materia e que sería un bo acompañamento para un primeiro curso. Durante a escrita deste manual, tivemos que facer algunhas escollas relativas á terminoloxía, dadas as poucas referencias existentes en lingua galega. En primeiro lugar, os adxectivos *lineal* e *linear* son ambos válidos. Nesta obra optouse de xeito bastante arbitrario polo primeiro, aínda que o nome oficial da materia no Grao en Matemáticas sexa *Álxebra Linear e Multilinear*. Un termo que foi fonte de conflito é *cuadrático*. Nin cuadrático nin cadrático aparecen recollidos no dicionario da Real Academia Galega, pero este si recolle termos como cuadrante, cuadrangular ou cuadratura; parécenos coherente entender que cuadrático, ao igual que estes outros termos, é un cultismo no que se mantivo a orixe latina da palabra e non se produciu a evolución que se observa en cadrado.

Agradecementos. Este libro foise escribindo durante o primeiro cuadrimestre do curso 2023-2024, mentres impartía a materia *Álxebra Lineal e Multilineal*, correspondente ao segundo curso do Grao en Matemáticas. As clases de problemas impartinas conxuntamente con Andrés Pérez e Brais Ramos, a quen agradezo as longas e produtivas discusións sobre a orientación da materia e sobre os exercicios. Tamén lles agradezo a lectura atenta das exposicións teóricas e os seus comentarios sobre elas. O texto complementouse e mellorouse durante o curso 2024-2025, no que volvín a impartir a materia, nesta ocasión conxuntamente con Ana Peón, con quen tamén compartín discusións sobre algunhas demostracións e que contribuíu a pulir este manual.
O alumnado tamén foi esencial na mellora deste documento, avisándome de diferentes erros e incorreccións en versións preliminares do texto; fago un agradecemento moi especial a Manuel Vilar por terme enviado unha listaxe exhaustiva de correccións e erros nunha versión preliminar do texto. Por outra banda, agradézolle a Felipe Gago a axuda con algunhas cuestións de LaTeX e sobre a organización do texto, ademais de múltiples correccións e comentarios.
Quero expresar finalmente a miña gratitude cara aos revisores anónimos do texto, polas súas suxestións de mellora e por teren detectado varios erros, e cara ao Servizo de Publicacións da USC por xestionar o proceso de publicación da obra.

Conceptos previos

Estruturas alxébricas

Nesta sección lembramos as definicións de grupo, anel e corpo, que aparecerán frecuentemente ao longo do texto.

Definición 0.1. Un *grupo* é un conxunto onde hai definida unha operación asociativa, con elemento neutro, e tal que todo elemento ten inverso. O grupo dise que é *abeliano* ou *conmutativo* se a operación é conmutativa.
Un *subgrupo* H dun grupo G é un subconxunto que, coa operación de G, é un grupo (é dicir, a operación é pechada, contén o elemento neutro e os inversos de cada elemento).

Definición 0.2. Un *anel* é un conxunto A con dúas operacións, suma e produto, de maneira que coa suma é un grupo abeliano e o produto é asociativo e é distributivo respecto á suma. Se o produto é conmutativo, o anel chámase *conmutativo*; se o produto ten elemento neutro chámase *unitario*.

Definición 0.3. Un corpo é un anel conmutativo e unitario no que $1 \neq 0$ e no que todo elemento diferente de 0 ten inverso para o produto.

Imos comentar algúns exemplos.

- Os enteiros, $\mathbb{Z}$, son un anel, pero non un corpo.
- Os racionais, $\mathbb{Q}$, os reais, $\mathbb{R}$, e os complexos, $\mathbb{C}$, son corpos.
- Se A é un anel,

  $$A[X] = \{a_0 + a_1X + \ldots + a_nX^n \text{ con } n \geq 0,\ a_0, \ldots, a_n \in A\}$$

  é tamén un anel. Fálase do *anel de polinomios na variable A*. Outro anel importante é o *anel de series de potencias na variable A*,

  $$A[\![X]\!] = \{a_0 + a_1X + \ldots, \text{ con } a_i \in A\}.$$

- Se A é un anel, o conxunto de elementos que teñen inverso para o produto denótase como $A^\times$ e ten estrutura de grupo. Por exemplo, $A[X]^\times = A^\times$.
- Se n é un número enteiro positivo, $\mathbb{Z}/n\mathbb{Z}$ é un anel coa suma e co produto. É un corpo se, e soamente se, n é primo.
- Sexa K un corpo e $\mathcal{M}_n(K)$ o conxunto de matrices cadradas $n \times n$ con entradas en K. Nese caso, $\mathcal{M}_n(K)$ ten estrutura de anel e as matrices invertibles forman un grupo, que se chama *grupo lineal*. Nesta materia traballaremos con detalle algunhas das súas propiedades e centrarémonos no estudo de varios dos seus subgrupos.

A modo de notación, cando K é un corpo, escribiremos $K^n = K \times \cdots \times K$ para refirirnos ao produto cartesiano de n copias de K. Os seus elementos son as n-tuplas $x = (x_1, \ldots, x_n)$, onde $x_i \in K$. En K^n podemos definir de xeito natural dúas operacións.

- Suma. Dados $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ en K^n, a súa suma é
$$x + y = (x_1 + y_1, \ldots, x_n + y_n) \in K^n.$$
- Produto por un escalar. Dados $x = (x_1, \ldots, x_n) \in K^n$ e $\lambda \in K$, o produto λx é
$$\lambda x = (\lambda x_1, \ldots, \lambda x_n) \in K^n.$$

Espazos vectoriais e aplicacións lineais

Definición 0.4. Sexa K un corpo. Un *espazo vectorial* sobre K, ou K-espazo vectorial, é un conxunto E, cuxos elementos se chaman *vectores*, onde hai definidas:

- unha operación interna, a suma de vectores, que fai corresponder a dous vectores $u, v \in E$ o vector suma $u + v \in E$;
- unha operación externa de K en E, o produto dun escalar por un vector, que a un vector $v \in E$ e a un escalar $\lambda \in K$ lles fai corresponder o vector produto $\lambda v \in E$.

Ademais, esíxese que se cumpran as seguintes propiedades:

- O conxunto E coa operación suma é un grupo abeliano. É dicir, a suma é asociativa, conmutativa, ten elemento neutro (que se denota por 0) e todo vector $v \in E$ ten un oposto con respecto á suma (que se denota por $-v$).
- O produto por escalares cumpre
$$(\lambda + \mu)v = \lambda v + \mu v, \quad (\lambda\mu)v = \lambda(\mu v), \quad 1v = v, \quad \lambda(u + v) = \lambda u + \lambda v,$$
onde $\lambda, \mu, 1 \in K$ e $u, v \in E$.

Algúns exemplos cos que traballaremos frecuentemente son os seguintes:

- Se K é un corpo, o espazo K^n de n-tuplas $x = (x_1, \ldots, x_n)$ de elementos de K.
- O espazo vectorial das matrices $m \times n$ sobre K, que se denota como $\mathcal{M}_{m\times n}(K)$.
- Os polinomios na variable X con coeficientes nun corpo K, $K[X]$.

No caso no que a multiplicación externa non é por un corpo, senón por un anel, falamos de módulos.

Definición 0.5. Sexa A un anel conmutativo. Un A-*módulo* é un grupo abeliano M cunha operación externa do anel A sobre M, é dicir,
$$A \times M \longrightarrow M, \quad (a, x) \mapsto ax,$$
de maneira que:
$$a(x + y) = ax + ay, \quad (a + b)x = ax + bx, \quad a(bx) = (ab)x, \quad 1x = x,$$
onde $a, b, 1 \in A$ e $x, y \in M$.

Por exemplo, $\mathbb{Q}$, $\mathbb{R}$, $n\mathbb{Z}$ e $\mathbb{Z}/n\mathbb{Z}$ son $\mathbb{Z}$-módulos. Máis en xeral, todo grupo abeliano é un $\mathbb{Z}$-módulo, onde a multiplicación nx defínese, se n é positivo, como $x + x + \ldots + x$, a suma de $-n$ veces o elementos x; e se n é negativo, como $-x - x - \ldots - x$, a suma de n veces o elementos $-x$.
Neste curso imos traballar case sempre sobre espazos vectoriais e non sobre módulos, aínda que nalgúns momentos faremos comentarios puntuais sobre as diferenzas entre os dous casos.

Definición 0.6. Sexa E un K-espazo vectorial. Un *subespazo vectorial* de E é un subconxunto $F \subset E$ que, ao considerarmos nel a suma de vectores e o produto por escalares que herda como subconxunto de E, é el mesmo un K-espazo vectorial. É dicir, cúmprense as seguintes propiedades:

(a) se $u, v \in F$, entón $u + v \in F$;

(b) $0 \in F$;

(c) se $\lambda \in K$ e $v \in F$, entón $\lambda v \in F$.

Dados $v_1, \ldots, v_n$ vectores dun K-espazo vectorial E, escribimos $\langle v_1, \ldots, v_n\rangle$ para referirnos ao conxunto de vectores que son combinacións lineais de $v_1, \ldots, v_n$, é dicir,

$$\langle v_1, \ldots, v_n\rangle = \{\lambda_1 v_1 + \cdots + \lambda_n v_n \mid \lambda_i \in K\} \subset E.$$

Este subconxunto é un subespazo vectorial de E.

Definición 0.7. Sexan $v_1, \ldots, v_n$ vectores dun K-espazo vectorial E.

- Dise que os vectores son *independentes* ou *linealmente independentes* se toda combinación lineal súa que sexa igual a 0 ten todos os coeficientes iguais a 0, é dicir, se $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$, entón $\lambda_1 = \cdots = \lambda_n = 0$.
- Dise que os vectores son *xeradores* se $E = \langle v_1, \ldots, v_n\rangle$.
- Dise que os vectores son *base* se son ao mesmo tempo independentes e xeradores.

Dise que un K-espazo vectorial E é finitamente xerado se existe unha familia finita $v_1, \ldots, v_n$ de vectores que xeran o espazo. Cúmprese que todo espazo vectorial finitamente xerado ten algunha base. Máis en xeral, tense o seguinte resultado, que se coñece como *Teorema de Steinitz.*

Proposición 0.1. Sexan $u_1, \ldots, u_n$ vectores independentes, e sexan $v_1, \ldots, v_m$ vectores xeradores dun espazo vectorial E. Entón, cúmprese que $n \leq m$ e pódense substituír n vectores da familia $\{v_j\}$ polos vectores da familia $\{u_i\}$, de forma que a familia que resulte siga a ser xeradora. En particular, todas as bases dun espazo vectorial finitamente xerado teñen o mesmo número de vectores.

A *dimensión* dun espazo finitamente xerado é o número de vectores dunha base. Os espazos que non son finitamente xerados dise que teñen dimensión infinita.

Definición 0.8. Sexan E e F espazos vectoriais sobre un corpo K. Unha aplicación lineal entre E e F é unha aplicación $f\colon E \to F$ tal que

$$f(u + v) = f(u) + f(v) \quad \text{e} \quad f(\lambda v) = \lambda f(v) \quad \text{para todo } u, v \in E,\ \lambda \in K.$$

O conxunto das aplicacións lineais de E a F ten estrutura de espazo vectorial e denótase por $\mathcal{L}(E, F)$.

Se $f\colon E \to E'$ é unha aplicación lineal e $F \subset E$ e $F' \subset E'$, son subespazos, entón a imaxe $f(F)$ e a imaxe recíproca (ou antiimaxe) $f^{-1}(F')$ son subespazos de E' e de E, respectivamente.

Definición 0.9. Sexa $f\colon E \to F$ unha aplicación lineal. O seu núcleo, que se denota como $\ker f$, é o subespazo de E formado polos vectores con imaxe 0:

$$\ker f = \{v \in E \mid f(v) = 0\} = f^{-1}(\{0\}).$$

A súa imaxe, que se denota $\operatorname{im} f$, é o subespazo de F formado polos vectores que son imaxe dalgún vector de U:

$$\operatorname{im} f = \{f(v) \mid v \in E\} = \{w \in F \mid \text{ existe } v \in E \text{ con } f(v) = w\} = f(E).$$

Coas notacións da definición anterior tense que f é inxectiva se, e soamente se, $\ker f = \{0\}$, e f é sobrexectiva se, e soamente se, $\operatorname{im} f = F$.
Sexa $f\colon E \to F$ unha aplicación lineal entre dous K-espazos vectoriais de dimensión finita. Sexan $\mathcal{B}_u = \{u_1, \dots, u_n\}$ e $\mathcal{B}_v = \{v_1, \dots, v_m\}$ bases dos espazos E e F, respectivamente.

Definición 0.10. A matriz de f nas bases $\mathcal{B}_u$ e $\mathcal{B}_v$ é a matriz $(a_{ij}) \in \mathcal{M}_{m\times n}(K)$ con entradas determinadas polas identidades

$$f(u_j) = \sum_{i=1}^{m} a_{ij} v_i, \quad j = 1, \dots, n.$$

É dicir, é a matriz que ten por columnas as coordenadas dos vectores $f(u_j)$, imaxe dos vectores da base $\mathcal{B}_u$ do espazo de saída, na base $\mathcal{B}_v$ do espazo de chegada F. Denotaremos esta matriz como $(f)_{\mathcal{B}_u,\mathcal{B}_v}$.

Un caso particular que se traballará con frecuencia ocorre cando $E = F$ e f é a aplicación identidade; nese caso, falamos de *matrices de cambios de base* e pomos $(\mathbb{I})_{\mathcal{B}_u,\mathcal{B}_v}$. Finalmente, se f é un endomorfismo e se considera a mesma base $\mathcal{B}_u$ nos espazos de saída e de chegada, pomos simplemente $(f)_{\mathcal{B}_u}$.

O espazo cociente

Sexa E un K-espazo vectorial e sexa $F \subset E$ un subespazo. En E, definimos unha relación de equivalencia da seguinte maneira: $u \sim_F v$ se, e soamente se, $u - v \in F$. É inmediato comprobar que se trata, efectivamente, dunha relación de equivalencia.

Definición 0.11. Escribimos E/F para referirnos ao conxunto de clases de equivalencia de vectores de E coa relación de equivalencia anterior. Pomos

$$[v] = \{u \in E \mid u \sim v\} = \{v + w \mid w \in F\} = v + F.$$

Como en toda relación de equivalencia hai unha aplicación canónica $\pi\colon E \twoheadrightarrow E/F$ que envía cada vector á súa clase: $\pi(v) = [v]$. No conxunto cociente E/F é posible definir unha suma de clases e un produto de clases por escalares a partir da suma e o produto por escalares de vectores representantes das clases:

$$[u] + [v] := [u + v], \quad \lambda[v] := [\lambda v].$$

Proposición 0.2. A suma e o produto por escalares de clases a partir de representantes están ben definidos e dan ao conxunto cociente a estrutura de K-espazo vectorial. Ademais, a aplicación canónica $\pi\colon E \to E/F$ é sobrexectiva con núcleo $\ker \pi = F$.

Da propoposición anterior dedúcese que $\dim(E/F) = \dim E - \dim F$.

Proposición 0.3 (Primeiro teorema de isomorfismo)**.** Para toda aplicación lineal $f\colon E \to F$, a aplicación $E/\ker f \to \operatorname{im} f \subset F$ que envía cada clase $[v]$ ao vector $f(v)$ está ben definida e é un isomorfismo de espazos vectoriais. Isto exprésase a través do seguinte diagrama conmutativo

$$\begin{tikzcd}
E \arrow[r, "f"] \arrow[d, two heads, "\pi"] & F \\
E/_{\ker f} \arrow[r, tail, two heads, dashed] & \mathrm{im} f, \arrow[u, hook]
\end{tikzcd}$$

no que π é a proxección ao espazo cociente.

O espazo dual

Ao longo desta sección, sexa K un corpo e E un espazo vectorial de dimensión finita. Sexa $\{e_1, \ldots, e_n\}$ unha base de E.

Definición 0.12. O *espazo dual* de E, que se denota como E^* ou $\mathcal{L}(E, K)$ é o conxunto de aplicacións lineais de E a K:

$$\mathcal{L}(E, K) = \Big\{ f\colon E \longrightarrow K \mid f \text{ lineal} \Big\}.$$

O espazo dual ten estrutura de espazo vectorial sobre K e a súa dimensión é n. Segundo a referencia que se consulte, é tamén frecuente empregar as notacións $\hat{E}$ ou $E^\vee$ para o espazo dual.

Definición 0.13. A base dual de $\{e_1, \ldots, e_n\}$ está formada polos elementos $\{e_1^*, \ldots, e_n^*\}$ de E^* tales que $e_i^*(e_j) = \delta_{i,j}$, onde $\delta_{i,j}$ é a delta de Kronecker.

Tense tamén que toda base do dual E^* nun espazo de dimensión finita E é a base dual dalgunha base de E.
Dada unha aplicación lineal $f\colon E \to F$, existe unha aplicación dual $f^*\colon F^* \to E^*$ definida por $f^*(\varphi) = \varphi \circ f$.

Proposición 0.4. Sexan E e F dous K-espazos vectorais con bases $\mathcal{B}_u$ e $\mathcal{B}_v$ respectivamente. Se A é a matriz da aplicación lineal $f\colon E \to F$ con respecto a esas bases, a matriz de $f^*\colon F^* \to E^*$ con respecto ás bases duais é A^t.

Problemas

Espazo cociente.

Problema 0.1. Demostrar que $\mathcal{B}_1 = \{[(2,1,1)], [(1,2,1)]\}$ e $\mathcal{B}_2 = \{[(3,1,1)], [(1,3,1)]\}$ son dúas bases do espazo $\mathbb{R}^3/\langle(1,1,1)\rangle$. Se $[u]$ ten coordenadas (a,b) na base $\mathcal{B}_1$, que coordenadas terá na base $\mathcal{B}_2$?

Solución. Para demostrar que $\mathcal{B}_1$ é base, é suficiente con comprobar que os vectores $\{(2,1,1),(1,2,1),(1,1,1)\}$ son unha base de $\mathbb{R}^3$. Alternativamente, iso é equivalente a ver que

$$\begin{aligned} 2a+b+c&=0\\ a+2b+c&=0\\ a+b+c&=0 \end{aligned}$$

é un sistema compatible determinado que ten por única solución o $(0,0,0)$. Iso é inmediato aplicando os métodos de redución gaussiana. O mesmo procedemento funciona para $\mathcal{B}_2$.
Para a segunda pregunta, o feito de que $[u]$ teña coordenadas (a,b) en $\mathcal{B}_1$ quere dicir que

$$u = a(2,1,1)+b(1,2,1)+c(1,1,1) = (2a+b+c, a+2b+c, a+b+c),$$

para algún $c \in \mathbb{R}$. Cómpre buscar agora $(a',b',c') \in \mathbb{R}^3$ de maneira que

$$(2a+b+c, a+2b+c, a+b+c) = a'(3,1,1)+b'(1,3,1)+c'(1,1,1).$$

Un cálculo rutineiro amosa que $a' = a/2$, $b' = b/2$ e $c' = (a+b)/2 + c$. Polo tanto, as coordenadas na base $\mathcal{B}_2$ son $(a/2, b/2)$.
Alternativamente, tamén se pode proceder empregando matrices de cambio de base.

Problema 0.2. Sexa F un subespazo vectorial de $\mathcal{M}_2(\mathbb{R})$. Demostrar que as seguintes condicións son equivalentes:

(i) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in F$.

(ii) Para toda matriz $B \in \mathcal{M}_2(\mathbb{R})$ tense a igualdade $[B] = [B^t]$ no espazo cociente $\mathcal{M}_2(\mathbb{R})/F$.

(iii) Para toda matriz $B \in \mathcal{M}_2(\mathbb{R})$ tense a igualdade $[B] = [\frac{1}{2}(B+B^t)]$ no espazo cociente $\mathcal{M}_2(\mathbb{R})/F$.

Solución. Para pasar da primeira á segunda condición, sexa $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ unha matriz arbitraria. Temos que

$$B - B^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} a & c \\ b & d \end{pmatrix} = (b-c)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

polo que se pode concluír que $B - B^t \in F$ e, polo tanto, $[B] = [B^t]$ no espazo cociente $\mathcal{M}_2(\mathbb{R})/F$.
Para irmos da segunda á terceira, basta con observar que

$$\left[\frac{1}{2}(B+B^t)\right] = \frac{1}{2}[B] + \frac{1}{2}[B^t] = \frac{1}{2}[B] + \frac{1}{2}[B] = [B]$$

en $\mathcal{M}_2(\mathbb{R})/F$.
Finalmente, para demostrar (i) usando (iii), impomos a condición coa matriz $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Entón,

$$[B] = \frac{1}{2}[B+B^t] = \left[\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right].$$

Polo tanto, $[B] = [0]$, o que quere dicir que $B \in F$.

Problema 0.3. O obxectivo deste problema é demostrar o segundo e o terceiro teorema de isomorfismo. Sexa U un K-espazo vectorial.

(a) Sexan $V, W \subset U$ subespazos vectoriais. Consideremos a aplicación lineal

$$V \hookrightarrow V + W \twoheadrightarrow (V + W)/W$$

definida como a composición da inclusión coa proxección natural no espazo cociente, que envía cada vector $v \in V$ á clase $[v] = v + W \subset V + W$. Demostrar que a aplicación é sobrexectiva, calcular o seu núcleo e establecer un isomorfismo de $(V + W)/W$ cun espazo cociente do espazo V.

(b) Sexan $W \subset V \subset U$ inclusións de espazos vectoriais. Demostrar que a aplicación $U/W \to U/V$ que envía $[u]_W = u + W$ á clase $[u]_V = u + V$ está ben definida e é sobrexectiva. Calcular o seu núcleo e deducir un isomorfismo entre U/V e un espazo cociente do espazo U/W.

Solución. (a) Sexa $[x] \in (V + W)/W$, onde $x \in V + W$ é un representante calquera. Entón temos que $x = v + w$, con $v \in V$ e $w \in W$ e por definición de espazo cociente, $[x] = [v]$. Deste xeito, para ver que a aplicación é sobrexectiva é suficiente con notar que a imaxe de $v \in V$ é simplemente $[v] = [x]$.

Como a aplicación $V \hookrightarrow V + W$ é inxectiva, un elemento $v \in V$ estará no núcleo da composición se, e soamente se, $v \in V \cap W$.

Polo tanto, o primeiro teorema de isomorfismo asegura que dada unha aplicación $f \colon A \to B$ entre dous espazos vectoriais, esta induce un isomorfismo

$$\bar{f} \colon A/\ker(f) \longrightarrow \operatorname{im}(f), \quad a \mapsto f(a).$$

Neste caso, como a aplicación xa é sobrexectiva, temos que o isomorfismo inducido será

$$V/_{V \cap W} \longrightarrow V+W/_{W}, \quad v + V \cap W \mapsto v + W.$$

O seguinte diagrama resume a situación que se presenta no teorema:

$$\begin{tikzcd}
V \cap W \arrow[r, hook] \arrow[d, hook] & W \arrow[r, two heads] \arrow[d, hook] & W/_{V \cap W} \arrow[d, dashed, "\simeq"] \\
V \arrow[r, hook] \arrow[d, two heads] & V + W \arrow[r, two heads] \arrow[d, two heads] & V+W/_{V} \\
V/_{V \cap W} \arrow[r, dashed, "\simeq"] & V+W/_{W} &
\end{tikzcd}$$

(b) Para ver que a aplicación está ben definida, observamos que se $[u_1]_W = [u_2]_W$, entón $u_1 - u_2 \in W$, e como $W \subset V$, tamén teremos que $u_1 - u_2 \in V$ e polo tanto $[u_1]_V = [u_2]_V$.

Para ver que a aplicación é sobrexectiva, sexa $[u]_V$ un elemento de U/V, e sexa $u \in U$ un representante calquera. Entón, $[u]_W$, a clase en U/W representada polo elemento u, ten por imaxe $[u]_V$, polo que a aplicación é sobrexectiva.

Por último, observamos que a imaxe de $[u]_W$ é cero se, e soamente se, $u \in V$. En particular, isto quere dicir que o núcleo é V/W. Polo tanto, o primeiro teorema de isomorfismo dános unha aplicación

$$ {}^{U}\!/_{W}\Big/{}^{V}\!/_{W} \longrightarrow U/V, \quad [u]_W + V/W \mapsto [u]_V $$

que é un isomorfismo. O seguinte diagrama resume a situación que se presenta no teorema:

$$
\begin{tikzcd}
& W \arrow[dl, hook'] \arrow[dr, hook] & & \\
V \arrow[rr, hook] \arrow[d, two heads] & & U \arrow[r, two heads] \arrow[d, two heads] & U/V \\
V/W \arrow[rr, tail] & & U/W \arrow[r, two heads] & {}^{U}\!/_{W}\Big/{}^{V}\!/_{W} \arrow[u, dashed, "\simeq"']
\end{tikzcd}
$$

Espazo dual.

Problema 0.4. Sexa $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ unha aplicación lineal definida por $f(x, y, z) = (2x + y, y - z)$ e sexan

$$ \mathcal{B}_u = \{(1, 1, 0), (0, 1, 0), (0, 1, -1)\} \quad \text{e} \quad \mathcal{B}_v = \{(1, 0), (1, 1)\} $$

bases de $\mathbb{R}^3$ e $\mathbb{R}^2$ respectivamente.

(a) Sexa $\mathcal{B}_e^* = \{e_1^*, e_2^*, e_3^*\}$ a base dual da base canónica $\mathcal{B}_e$ de $\mathbb{R}^3$. Expresar a base dual $\mathcal{B}_u^* = \{u_1^*, u_2^*, u_3^*\}$ en termos de $\mathcal{B}_e^*$.

(b) Encontrar as compoñentes de $w = 2e_1^* + e_2^* - e_3^*$ na base $\mathcal{B}_u^*$.

(c) Determinar a matriz $(f^*)_{\mathcal{B}_v^*, \mathcal{B}_u^*}$ da aplicación dual de f nas bases $\mathcal{B}_v^*$ e $\mathcal{B}_u^*$, respectivamente.

Solución. (a) Hai dúas maneiras de resolver este apartado. A primeira consiste en considerar a matriz de cambio de base de $\mathcal{B}_u$ a $\mathcal{B}_e$,

$$ (\mathbb{I})_{\mathcal{B}_u, \mathcal{B}_e} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}, $$

e recordar o resultado que afirma que

$$ (\mathbb{I})_{\mathcal{B}_u^*, \mathcal{B}_e^*} = ((\mathbb{I})^{-1}_{\mathcal{B}_u, \mathcal{B}_e})^t = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}. $$

Polo tanto, $u_1^* = e_1^*$, $u_2^* = -e_1^* + e_2^* + e_3^*$ e $u_3^* = -e_3^*$.

Alternativamente, para cada un dos vectores da base dual podemos pór $u_1^* = a_{11}e_1^* + a_{12}e_2^* + a_{13}e_3^*$. Aplicando a definición de base dual, $u_1^*(u_j) = \delta_{1,j}$, polo que $a_{11} + a_{12} = 0$, $a_{12} = 0$ e $a_{12} - a_{13} = 0$. Resolvendo o sistema, temos que $(a_{11}, a_{12}, a_{13}) = (1, 0, 0)$. Facendo o mesmo con u_2^* e con u_3^* temos o resultado anterior.

(b) Novamente hai dúas maneiras de proceder. Usando a matriz $(\mathbb{I})_{\mathcal{B}_e^*,\mathcal{B}_u^*}$, o resultado é simplemente

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

Alternativamente, podemos escribir $w = au_1^* + bu_2^* + cu_3^*$ e usar as definicións. Por exemplo,

$$2 = w(e_1) = au_1^*(e_1) + bu_2^*(e_1) + cu_3^*(e_1) = a - b.$$

Do mesmo xeito, $1 = b$ e $-1 = b - c$, recuperándose a solución $(a, b, c) = (3, 1, 2)$.

(c) Sexa $(f^*)_{\mathcal{B}_e^*}$ a matriz de f^* na dual da base canónica; cun pequeno abuso de notación, poremos $\mathcal{B}_e$ para as bases canónicas de $\mathbb{R}^3$ e $\mathbb{R}^2$, xa que non hai perigo de confusión. Usando as matrices de cambio de base, temos que

$$(f^*)_{\mathcal{B}_v^*,\mathcal{B}_u^*} = (\mathbb{I})_{\mathcal{B}_e^*,\mathcal{B}_u^*} \cdot (f^*)_{\mathcal{B}_e^*} \cdot (\mathbb{I})_{\mathcal{B}_v^*,\mathcal{B}_e^*}.$$

A matriz $(\mathbb{I})_{\mathcal{B}_e^*,\mathcal{B}_u^*}$ foi calculada no primeiro apartado. Por outro lado,

$$(\mathbb{I})_{\mathcal{B}_v^*,\mathcal{B}_e^*} = (\mathbb{I})^t_{\mathcal{B}_e,\mathcal{B}_v} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Combinando estes resultados, temos que

$$(f^*)_{\mathcal{B}_v^*,\mathcal{B}_u^*} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

Alternativamente, nas bases canónicas sabemos que $f^*(e_1^*) = 2e_1^* + e_2^*$ e $f^*(e_2^*) = e_2^* - e_3^*$. En $\mathbb{R}^2$ temos que $v_1^* = e_1^* - e_2^*$ e $v_2^* = e_2^*$. Imos calcular entón a imaxe de v_1^*. Entón

$$\begin{aligned} f(v_1^*) &= f(e_1^*) - f(e_2^*) = 2e_1^* + e_3^* \\ &= 2u_1^* - u_3^*. \end{aligned}$$

Precisamente, $(2, 0, -1)$ é a primeira columna da matriz. Facendo o mesmo con v_2^* obteriamos a segunda.

Problema 0.5. Sexa $f\colon \mathbb{R}^2 \to \mathbb{R}^4$ a aplicación lineal definida por

$$f(x, y) = (3x - y, 5x - 2y, -4x + 2y, -7x + 3y).$$

Sexan $v_1, v_2 \in \mathbb{R}^2$ os vectores $v_1 = (1, 2)$ e $v_2 = (1, 1)$. Sexan $w_1, w_2, w_3, w_4 \in \mathbb{R}^4$ os vectores $w_1 = (1, 1, -1, -1)$, $w_2 = (1, 1, -1, -2)$, $w_3 = (0, -1, 1, 1)$ e $w_4 = (-1, 1, 0, 0)$.

(a) Demostrar que $\mathcal{B}_v = \{v_1, v_2\}$ e $\mathcal{B}_w = \{w_1, w_2, w_3, w_4\}$ son bases de $\mathbb{R}^2$ e $\mathbb{R}^4$, respectivamente. Calcular as súas bases duais.

(b) Determinar a matriz da aplicación dual f^* nas bases duais obtidas no apartado anterior.

Solución. (a) A comprobación de que son base é rutineira e omitímola. Para achar as bases duais, podemos proceder de varias maneiras. Por exemplo, sexa $\mathcal{B}_v^* = \{v_1^*, v_2^*\}$ a base dual de $\mathcal{B}_v$. Se $v_1^* = \alpha e_1^* + \beta e_2^*$, temos que

$$\begin{cases} \alpha + 2\beta = 1 \\ \alpha + \beta = 0. \end{cases}$$

Resolvendo o sistema, quédanos que $v_1^* = -e_1^* + e_2^*$. De xeito análogo, $v_2^* = 2e_1^* - e_2^*$. Procedendo do mesmo xeito con $\mathcal{B}_w$, e empregando cun pequeno abuso de notación novamente $e_1^*, \ldots, e_4^*$ para a base canónica, obtemos que $w_1^* = e_1^* + e_2^* + e_4^*$, $w_2^* = e_3^* - e_4^*$, $w_3^* = e_1^* + e_2^* + 2e_3^*$ e $w_4^* = e_2^* + e_3^*$.

(b) Como $f(1,2) = (1,1,0,-1)$, pomos

$$(1,1,0,-1) = \alpha(1,1,-1,-1) + \beta(1,1,-1,-2) + \gamma(0,-1,1,1) + \delta(-1,1,0,0).$$

Isto quere dicir que as coordenadas de $f(v_1)$ na base $\mathcal{B}_w$ son $(1,1,2,1)$. De xeito similar, $f(v_2) = f(1,1) = (2,3,-2,-4)$ ten coordenadas $(-1,2,-1,-1)$. Polo tanto, a matriz de f con respecto ás bases $\mathcal{B}_v$ e $\mathcal{B}_w$ é

$$\begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 2 & -1 \\ 1 & -1 \end{pmatrix},$$

polo que a matriz buscada é a transposta:

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ -1 & 2 & -1 & -1 \end{pmatrix}.$$

Problema 0.6. Sexa E un espazo vectorial, e fixemos $w \in E$ e $\phi \in E^*$. Demostrar que a aplicación $f : E \longrightarrow E$ dada por $f(v) = v + \phi(v)w$ é un endomorfismo, e calcular a aplicación dual f^*.

Solución. En primeiro lugar, vemos que f está ben definida, dado que tanto v como $\phi(v)w$ son elementos de E. Para ver que é lineal, observamos que

$$f(u+v) = u + v + \phi(u+v)w = u + \phi(u)w + v + \phi(v)w = f(u) + f(v)$$

e que

$$f(\lambda u) = \lambda u + \phi(\lambda u)w = \lambda u + \lambda\phi(u)w = \lambda f(u).$$

Para achar a aplicación dual f^*, observamos que por definición ten que cumprir que $f^*(\psi) = \psi \circ f$ para toda ψ. Iso é equivalente a dicir que $f^*(\psi)(v) = (\psi \circ f)(v)$ para todo v. Agora ben,

$$(\psi \circ f)(v) = \psi(v) + \phi(v)\psi(w).$$

Polo tanto, se pomos $f^*(\psi) = \psi + \psi(w) \cdot \phi$, resulta claro que

$$f^*(\psi)(v) = \psi(v) + \phi(v)\psi(w),$$

co cal cumpre a definición de aplicación dual.

Problema 0.7. Para cada $\alpha \in \mathbb{R}$, sexa $\phi_\alpha \colon \mathbb{R}_n[X] \longrightarrow \mathbb{R}$ a aplicación que a un polinomio $P(X)$ lle fai corresponder o seu valor $P(\alpha)$. Sexan $\alpha_0, \alpha_1, \ldots, \alpha_n$ unha familia de $n+1$ números reais diferentes.

(a) Demostrar que as $n+1$ aplicacións lineais $\phi_{\alpha_0}, \ldots, \phi_{\alpha_n}$ son unha base do espazo dual $\mathbb{R}_n[X]^*$. Sexa $\mathcal{B} = \{P_0(X), \ldots, P_n(X)\}$ a base de $\mathbb{R}_n[X]$ que a ten por dual.

(b) Encontrar as coordenadas dun polinomio $P(X) \in \mathbb{R}_n[X]$ na base $\mathcal{B}$.

(c) Demostrar que, dados $n+1$ números reais calquera $\beta_0, \ldots, \beta_n$, o polinomio

$$P(X) = \sum_{i=0}^{n} \beta_i P_i(X)$$

é o único polinomio de $\mathbb{R}_n[X]$ tal que $P(\alpha_i) = \beta_i$ para todo $i = 0, \ldots, n$.

(d) Pomos agora $n = 3$ e para $i = 0, 1, 2, 3$ pomos $\alpha_i = i$. Encontrar a base $\{P_i(X)\}$ de $\mathbb{R}_3[X]$ que ten como dual a base $\{\phi_0, \phi_1, \phi_2, \phi_3\}$.

(e) Sexan $a < b$ dous números reais. Demostrar que existe unha única $(n+1)$-tupla de números reais $(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}$ tal que

$$\int_a^b P(t)\, dt = \sum_{i=0}^{n} x_i P(\alpha_i)$$

para todo polinomio $P(X) \in \mathbb{R}_n[X]$.

(f) Calcular os coeficientes $x_i \in \mathbb{R}^4$ que corresponden á forma lineal $\int_0^1 P(t)\, dt$ de $\mathbb{R}_3[X]$ e aos números $\alpha_i = 0, 1, 2, 3$.

Solución. (a) Definimos os polinomios $q_j(X) = \prod_{i \neq j}(X - \alpha_i)$, que teñen por raíces todos os α_i a excepción do α_j. Se avaliamos unha combinación lineal $\sum_{i=0}^{n} \lambda_i \phi_{\alpha_i}$ en $q_j(X)$ quédanos $\lambda_j q_j(\alpha_j)$; como α_j non é raíz de $p_j(X)$ necesariamente $\lambda_j = 0$. Este razoamente vale para calquera j, polo que a combinación lineal é novamente a trivial. Isto demostra que os ϕ_{α_i} son base.

(b) Escribamos $P(X) = \sum_{i=0}^{n} \mu_i P_i(X)$. Se consideramos a avaliación de ϕ_{α_i} en $P(X)$ quédanos $P(\alpha_i) = \mu_i$, polo que

$$P(X) = \sum_{i=0}^{n} P(\alpha_i) P_i(X).$$

(c) Supoñamos que existe $q(X) \in \mathbb{R}_n[X]$, con $q(X) \neq p(X)$, e de maneira que $q(\alpha_i) = \beta_i$ para todo $i = 0, 1, \ldots, n$. Este polinomio $q(X)$ escríbese na base dada polos $P_i(X)$ como $q(X) = \sum_{i=0}^{n} q(\alpha_i) P_i(X)$. Como $q(\alpha_i) = \beta_i$, tense que $q(X) = \sum_{i=0}^{n} \beta_i P_i(X) = P(X)$.

(d) Imos amosar tamén dúas maneiras de resolver este apartado. Chamando $P_0(X) = a_{00} + a_{01}X + a_{02}X^2 + a_{03}X^3$, temos que

$$P_0(0) = 1,\ P_0(1) = 0,\ P_0(2) = 0,\ P_0(3) = 0.$$

Polo tanto, temos o sistema

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \end{pmatrix} \begin{pmatrix} a_{00} \\ a_{01} \\ a_{02} \\ a_{03} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

A solución correspóndese con $P_0(X) = \frac{-X^3+6X^2-11X+6}{6}$.

Alternativamente, podemos usar que $P_0(X)$ é un polinomio de grao 3 con raíces 1, 2 e 3, polo que $P_0(X) = \alpha_0(X-1)(X-2)(X-3)$. Impondo que $P_0(0) = 1$ temos que $1 = -6\alpha$, e obtemos a mesma solución.

De xeito análogo, temos que os outros polinomios da base son

$$P_1(X) = \frac{X^3 - 5X^2 + 6X}{2},\ P_2(X) = \frac{-X^3 + 4X^2 - 3X}{2},\ P_3(X) = \frac{X^3 - 3X^2 + 2X}{6}.$$

(e) Dado $P(X) \in \mathbb{R}_n[X]$, podemos escribir $P(X) = \sum_{i=0}^n P(\alpha_i)p_{\alpha_i}(X)$. Ao integrar con respecto a x tense que

$$\int_a^b P(t)\,dt = \int_a^b \sum_{i=0}^n P(\alpha_i)p_{\alpha_i}(t)\,dt = \sum_{i=0}^n P(\alpha_i) \cdot \int_a^b p_{\alpha_i}(t)\,dt.$$

Se definimos $x_i = \int_a^b P_{\alpha_i}(t)\,dt$, temos que

$$\int_a^b P(t)\,dt = \sum_{i=0}^n x_i \cdot P(\alpha_i).$$

(f) Polo establecido no apartado anterior, é suficiente integrar os polinomios que achamos no apartado (d). Por exemplo,

$$x_0 = \int_0^1 p_0(x)\,dx = \int_0^1 \Big(\frac{-x^3}{6} + x^2 - \frac{11x}{6} + 1\Big)\,dx = -\frac{1}{24} + \frac{1}{3} - \frac{11}{12} + 1 = \frac{3}{8}.$$

Os valores buscados son $(x_0, x_1, x_2, x_3) = \left(\frac{3}{8}, \frac{19}{24}, \frac{-5}{24}, \frac{1}{24}\right)$.

Capítulo 1

Polinomios

O obxectivo deste capítulo é introducir algúns conceptos básicos sobre polinomios que serán útiles posteriormente no estudo e clasificación de endomorfismos. En particular, estableceremos que moitas das propiedades sobre divisibilidade que se cumpren para o anel dos números enteiros tamén se poden estender aos polinomios.

1.1. Definicións básicas

De agora en adiante, sexa K un corpo arbitrario.

Definición 1.1. O *anel de polinomios nunha variable con coeficientes en* K é o conxunto

$$K[X] = \{f(X) = a_0 + a_1X + \ldots + a_nX^n \text{ con } a_i \in K\}$$

coas operacións de suma e multiplicación habituais.

Non faremos a comprobación de que efectivamente é un anel conmutativo, por tratarse dun cálculo rutineiro. Ademais, $K[X]$ pode verse como un K-espazo vectorial de dimensión infinita.

Definición 1.2. O *grao* dun polinomio $f \neq 0$ é o maior enteiro n tal que $a_n \neq 0$. Escribiremos $\mathrm{gr}(f)$ ou $\deg(f)$ (polo termo inglés *degree*). O polinomio $f = 0$ non ten grao definido; por convenio, poremos $\deg(0) = -\infty$.
Un polinomio de grao n con $a_n = 1$ chámase *mónico*.

Proposición 1.1. O grao ten as seguintes propiedades.

(a) $\deg(f + g) \leq \text{máx}\{\deg(f), \deg(g)\}$, con igualdade se $\deg(f) \neq \deg(g)$.

(b) $\deg(fg) = \deg(f) + \deg(g)$.

O resultado anterior tamén se aplica cando algún dos polinomios (ou ámbolos dous) é igual a 0.

1.2. Algoritmo da división

A seguinte proposición afirma que o algoritmo da división de números enteiros tamén funciona no anel de polinomios. A demostración basearase en reproducir o algoritmo.

Proposición 1.2. Dados polinomios $f(X), g(X) \in K[X]$, con $g \neq 0$, existen polinomios $q(X)$ e $r(X)$ únicos e tales que

$$f(X) = g(X)q(X) + r(X), \quad \deg(r) < \deg(g).$$

Demostración. Imos comezar vendo que se existen, os polinomios necesariamente son únicos. De ter dúas igualdades da forma

$$\begin{aligned} f(X) &= g(X)q_1(X) + r_1(X) \\ f(X) &= g(X)q_2(X) + r_2(X), \end{aligned}$$

podemos restar unha da outra e obter

$$g(X)(q_1(X) - q_2(X)) = r_2(X) - r_1(X).$$

Porén, se $q_1(X) - q_2(X) \neq 0$, temos que

$$\deg(g(X)) > \deg(r_2(X) - r_1(X)) = \deg(g(X)(q_1(X) - q_2(X))) \geq \deg(g(X)).$$

Polo tanto, $q_1(X) - q_2(X) = 0$ e de aquí deducimos que $r_1(X) = r_2(X)$.
Imos comprobar agora a existencia. Se $f = 0$, o resultado é trivial pondo $q(X) = r(X) = 0$. Senón, pomos

$$f(X) = a_0 + a_1X + \ldots + a_nX^n, \quad g(X) = b_0 + b_1X + \ldots + b_mX^m,$$

con $a_n, b_m \neq 0$. Se $n < m$, entón $f(X) = g(X) \cdot 0 + f(X)$ e o resultado é certo. Se $n \geq m$, podemos escribir

$$f(X) - g(X) \cdot \frac{a_n}{b_m} X^{n-m} = r_1(X),$$

onde o polinomio $r_1(X)$ é 0 ou ten grao $n_1 < n$. Se $r_1(X) = 0$ ou $n_1 < m$, pomos $q(X) = \frac{a_n}{b_m} X^{n-m}$ e $r(X) = r_1(X)$.
No caso $n_1 \geq m$, iteramos o proceso novamente e pomos $r_1(X) = c_0 + c_1X + \ldots + c_{n_1}X^{n_1}$, con $c_{n_1} \neq 0$; novamente,

$$r_1(X) - g(X)\frac{c_{n_1}}{b_m} X^{n_1-m} = r_2(X),$$

con $r_2(X) = 0$ ou de grao $n_2 < n_1$. En calquera caso,

$$f(X) = g(X)\Big(\frac{a_n}{b_m}X^{n-m} + \frac{c_{n_1}}{b_m}X^{n_1-m}\Big) + r_2(X).$$

Se $n_2 \geq m$, hai que repetir o proceso novamente, pero como o valor $n_i - m$ é estritamente menor despois de cada iteración, acabaremos despois dun número finito de pasos cun polinomio que será ou ben 0 ou ben de grao menor que m. Polo tanto, despois de k pasos teremos

$$f(X) = g(X)\Big(\frac{a_n}{b_m}X^{n-m} + \frac{c_{n_1}}{b_m}X^{n_1-m} + \ldots + \frac{c_{n_{k-1}}}{b_m}X^{n_{k-1}-m}\Big) + r_k(X).$$

$\square$

Cando $r(X) = 0$ diremos que $f(X)$ é un múltiplo de $g(X)$. Grazas á existencia de división euclidiana, podemos considerar os conceptos habituais de divisibilidade sobre os enteiros como mínimo común múltiplo, máximo común divisor, algoritmo de Euclides ou identidade de Bézout. Frecuentemente denotaremos o mínimo común múltiplo como lcm, polas súas siglas en inglés, e o máximo común divisor como gcd.
Coas notacións do resultado anterior, o algoritmo de Euclides di que $\gcd(f(X), g(X)) = \gcd(g(X), r(X))$; por outro lado, a identidade de Bézout asegura que existen polinomios $u(X)$ e $v(X)$ tales que

$$u(X) \cdot f(X) + v(X) \cdot g(X) = \gcd(f(X), g(X)).$$

Ademais, o algoritmo de Euclides proporciona unha maneira construtiva de achar os polinomios $u(X)$ e $v(X)$.

Exemplo. Consideremos o caso dos polinomios $f(X) = X^4 + X^3 + 1$ e $g(X) = X^2 + 1$. Entón, en dúas iteracións, o algoritmo de Euclides danos que

$$\begin{aligned} x^4 + x^3 + 1 &= (x^2 + 1)(x^2 + x - 1) + (-x + 2) \\ x^2 + 1 &= (-x + 2)(-x - 2) + 5 \end{aligned}.$$

Polo tanto, $\gcd(x^4 + x^3 + 1, x^2 + 1) = 1$, dado que calquera constante do corpo (en particular o 5) divide un polinomio arbitrario. Procedendo agora en sentido inverso,

$$\begin{aligned} 5 &= 1 \cdot (x^2 + 1) + (x + 2) \cdot (-x + 2) \\ &= 1 \cdot (x^2 + 1) + (x + 2) \cdot (x^4 + x^3 + 1 - (x^2 + 1) \cdot (x^2 + x - 1)) \\ &= (x + 2) \cdot (x^4 + x^3 + 1) + (-x^3 - 3x^2 - x + 3) \cdot (x^2 + 1). \end{aligned}$$

Alternativamente,

$$1 = \frac{x + 2}{5} \cdot (x^4 + x^3 + 1) + \frac{-x^3 - 3x^2 - x + 3}{5} \cdot (x^2 + 1).$$

Exemplo. Igual que sucede no caso da aritmética modular, podemos considerar sistemas de congruencias co teorema chinés do resto. En concreto, sexan $p_1(X), \ldots, p_k(X)$ polinomios en $K[X]$, e $q_1(X), \ldots, p_k(X)$ polinomios en $K[X]$ relativamente primos dous a dous. Pomos $Q(X) = q_1(X)q_2(X) \cdots q_k(X)$. Entón, o sistema de congruencias $f(X) \equiv p_i(X) \pmod{q_i(X)}$, con $i \in \{1, \ldots, k\}$ ten solución, e se $f_0(X)$ é unha solución particular, o conxunto de solucións é

$$f_0(X) + Q(X) \cdot K[X] = \{f_0(X) + Q(X)R(X) \mid R(X) \in K[X]\}.$$

Máis en concreto, sexa $t_i(X)$ o inverso de $\frac{Q(X)}{q_i(X)}$ módulo $q_i(X)$. Entón, podemos escoller

$$f_0(X) = p_1(X)t_1(X) \cdot \frac{Q(X)}{q_1(X)} + \ldots + p_k(X)t_k(X)\frac{Q(X)}{q_k(X)}.$$

Por exemplo, como sabemos que $X^2 + 1$ e $X + 1$ son relativamente coprimos en $\mathbb{Q}[X]$, podemos asegurar que existe un único polinomio módulo $(X + 3)(X^2 + 1)$ de xeito que

$$\begin{cases} f(X) \equiv X + 3 & \pmod{X^2 + 1} \\ f(X) \equiv 1 & \pmod{X + 1}. \end{cases}$$

En concreto, podemos dar expresións exactas para $f(X)$, observando que o inverso de $X+1$ módulo X^2+1 é $\frac{1-X}{2}$ e o inverso de $X^2+1 \equiv 2$ módulo $X+1$ é $\frac{1}{2}$. Polo tanto,

$$f_0(X) = \frac{(X+3)(1-X)(X+1)}{2} + \frac{X^2+1}{2} = \frac{-X^3-2X^2+X+4}{2}.$$

Observamos ademais que é posible coller calquera outro polinomio obtido ao sumarlle a $f(X)$ un múltiplo de $(X+1)(X^2+1)$.

1.3. Polinomios irredutibles

Definición 1.3. Un polinomio non constante $f(X)$ dise que é *irredutible* se os seus únicos divisores son da forma k e $k \cdot f(X)$, con $k \in K$.

En teoría de aneis, é habitual definir unha noción de primo, dicindo que $f(X)$ é primo se sempre que divide a un produto $g(X)h(X)$, entón divide a $g(X)$ ou a $h(X)$. Neste caso, os dous conceptos son equivalentes, pero imos traballar polo de agora coa noción de irredutible. Por outro lado, é conveniente excluír da definición o caso dos polinomios constantes.

Proposición 1.3. Todo polinomio $f(X) \neq 0$ de grao maior que 0 é produto de irredutibles

Demostración. Se $f(X)$ é irredutible, o resultado é certo. En caso contrario, sexa $g_1(X)$ un divisor de grao mínimo entre os de $f(X)$. Entón, $g_1(X)$ é irredutible, xa que todos os seus divisores tamén o son de $f(X)$. Podemos pór entón $f(X) = g_1(X) \cdot f_1(X)$. Se $f_1(X)$ é irredutible, acabamos; en caso contrario, consideramos novamente un dos seus divisores de grao mínimo, $g_2(X)$, e escribimos $f(X) = g_1(X) \cdot g_2(X) \cdot f_2(X)$. Como $\deg(f) > \deg(f_1) > \deg(f_2) > \ldots$, o proceso remata despois dun número finito de pasos. □

Por último, é importante observar que a descomposición en irredutibles, como sucede no caso dos enteiros, é única (salvo multiplicación por constantes). Para demostrar o resultado, imos usar que se $p(X)$ divide o produto $q(X)r(X)$ e $\gcd(p(X), q(X)) = 1$, entón $p(X)$ divide $r(X)$. Isto vese aplicando a identidade de Bézout, que nos permite considerar polinomios $u(X)$ e $v(X)$ tales que $1 = p(X)u(X) + q(X)v(X)$. Polo tanto,

$$r(X) = p(X)r(X)u(X) + q(X)r(X)v(X),$$

e o polinomio $p(X)$ divide os dous sumandos da dereita e por conseguinte tamén divide $r(X)$.

Proposición 1.4. Se $f(X) = p_1(X) \cdots p_n(X) = q_1(X) \cdots q_m(X)$ e todos os factores son polinomios irredutibles, entón $n = m$ e os polinomios $\{p_i(X)\}$ son os mesmos que os $\{q_j(X)\}$ salvo factores do corpo K.

Demostración. Faremos indución sobre n. Cando $n = 1$, é obviamente certo que $m = 1$ e $p_1(X) = q_1(X)$. Supoñamos logo que o resultado é certo cando $n \leq r-1$ e imos establecelo para $n = r$. Dada a igualdade $p_1(X) \cdots p_r(X) = q_1(X) \cdots q_m(X)$, temos que $p_r(X)$ divide o produto $q_1(X)(q_2(X) \cdots q_m(X))$. Se $p_r(X)$ non coincide con $q_1(X)$ (salvo multiplicación por factores de K), entón é relativamente primo con el e divide a $q_2(X) \cdots q_m(X)$. Repetindo o razoamento, chegamos a que un $q_j(X)$ é igual a $p_r(X)$

salvo un factor de K, isto é, $p_r(X) = k \cdot q_j(X)$, con $k \in K$ diferente de 0. Suprimindo este factor de ambos lados da igualdade, podemos aplicar a hipótese de inducción e concluír que cada un dos $p_i(X)$, con $1 \leq i \leq r-1$ coincide cos $q_j(X)$ restantes e en particular $r = m$. □

Observamos que a proba deste resultado usa de forma esencial a existencia de división euclidiana no anel de polinomios, dado que usamos a identidade de Bézout para demostrar que se $p(X)$ divide o produto $q(X)r(X)$ e $\gcd(p(X), q(X)) = 1$, entón $p(X)$ divide $r(X)$.

1.4. Raíces de polinomios

Definición 1.4. Un elemento $k \in K$ é un *cero* do polinomio $f(X)$ se $f(k) = 0$. Ás veces tamén se usa o termo *raíz*.

Se L é un corpo de maneira que $K \subset L$, un polinomio $p(X) \in K[X]$ pódese ver de xeito natural como un polinomio en $L[X]$; polo tanto, ten sentido falar das raíces de $p(X)$ en $L[X]$. Por exemplo, dicimos que o polinomio $X^2 + 1 \in \mathbb{R}[X]$ ten raíces $\pm i \in \mathbb{C}$, xa que estamos considerando implicitamente que $X^2 + 1 \in \mathbb{C}[X] \supset \mathbb{R}[X]$.
É habitual introducir a aplicación lineal *avaliación en* k, que denotaremos por av_k e que está dada por

$$\mathrm{av}_k \colon K[X] \longrightarrow K, \quad f(X) \mapsto f(k).$$

Entón, k é un cero de $f(X)$ se e soamente se $f(X)$ está no núcleo de av_k.

Proposición 1.5. Un elemento $k \in K$ é un cero de $f(X)$ se e soamente se $X - k$ divide $f(X)$.

Demostración. Se $f(X) = (X-k)q(X)$, avaliando en $X = k$ temos que $f(k) = 0$. Reciprocamente, supoñamos que $k \in K$ é un cero de $f(X)$, con $f \neq 0$ (o caso $f = 0$ é trivial). Podemos considerar a división euclidiana e escribir $f(X) = (X-k) \cdot q(X) + r$, onde $r \in K$. Avaliando en $X = k$, temos que $0 = f(k) = r$, co cal $X - k$ divide $f(X)$. □

Definición 1.5. Se $f(X) = (X-k)^m \cdot g(X)$, con $\gcd(X-k, g(X)) = 1$, diremos que k é unha raíz de $f(X)$ de *multiplicidade* m. Se $m > 1$, diremos que k é unha *raíz múltiple*.

Definición 1.6. Un corpo K é *alxebricamente pechado* se todo polinomio $f(X)$ non constante ten unha raíz.

Enunciamos agora o coñecido como teorema fundamental da álxebra. As probas máis sinxelas do mesmo usan a teoría da variable complexa (o teorema de Liouville) ou o concepto de *grao* topolóxico. Por esta razón, ímola omitir polo de agora.

Teorema 1.1. O corpo dos números complexos, $\mathbb{C}$, é alxebricamente pechado.

Corolario 1.1. Os factores irredutibles dun polinomio $f(X) \in \mathbb{R}[X]$ teñen grao como moito dous.

Demostración. Sobre os números complexos, o polinomio $f(X)$ ten tantas raíces como o seu grao. Imos demostrar que se $X - \alpha \in \mathbb{C}[X]$ é un factor irredutible, entón $X - \bar{\alpha} \in \mathbb{C}[X]$ tamén o será; para iso, é suficiente con observar que

$$0 = f(\alpha) = \overline{f(\alpha)} = f(\bar{\alpha}).$$

Polo tanto, $(X-\alpha)(X-\bar{\alpha})$ é un factor de $f(X)$, que se pode escribir como

$$(X-\alpha)(X-\bar{\alpha}) = X^2 - (\alpha+\bar{\alpha})X + \alpha\bar{\alpha} \in \mathbb{R}[X].$$

Dividindo por este factor e iterando o proceso, concluímos a demostración. $\square$

O caso dos polinomios sobre $\mathbb{Q}$ é polo xeral máis complicado. Se un polinomio ten coeficientes enteiros, diremos que é *primitivo* cando o máximo común divisor dos seus coeficientes é 1. Nese caso, cúmprese que o polinomio é irredutible sobre $\mathbb{Q}[X]$ se e soamente se o é tamén sobre $\mathbb{Z}[X]$ (é o coñecido como lema de Gauss). En $\mathbb{Z}[X]$ podemos aplicar os resultados habituais de divisibilidade; por exemplo, que un polinomio mónico ten todas as súas raíces entre os divisores do termo independente.

Proposición 1.6. Se $P(X)$ e $Q(X)$ son dous polinomios primitivos en $\mathbb{Z}[X]$, entón o seu produto $P(X)Q(X)$ tamén é primitivo. En particular, un polinomio primitivo $f(X) \in \mathbb{Z}[X]$ é irredutible se, e soamente se, $f(X)$ é irredutible en $\mathbb{Q}[X]$.

Demostración. O produto de dous polinomios con coeficientes enteiros tamén ten coeficientes enteiros. Pomos $P(X) = \sum_{i=0}^m a_i X^i$ e $Q(X) = \sum_{i=0}^n b_i X^i$. Supoñamos que $P(X)Q(X)$ non é primitivo, polo que existe un primo p que divide todos os coeficientes. Sexa a_r o primeiro coeficiente de $P(X)$ que non é divisible por p (é dicir, $a_0, a_1, \dots, a_{r-1}$ son múltiplos de p, pero a_r non) e sexa b_s o primeiro coeficiente de $Q(X)$ que non é divisible por p. Entón, temos que o coeficiente con X^{r+s} en $P(X)Q(X)$ é $\sum_{i+j=r+s} a_i b_j$; porén, se $i<r$ ou se $j<s$, o produto $a_i b_j$ é múltiplo de p. Temos unha suma na que todos os sumandos salvo un, $a_r b_s$, son múltiplos de p. Polo tanto, ese coeficiente non pode ser múltiplo de p, o que é unha contradición.

Sexa $f(X) \in \mathbb{Z}[X]$ de maneira que admite unha factorización en $\mathbb{Q}[X]$ da forma $f(X) = g(X)h(X)$, onde tanto $g(X)$ como $h(X)$ teñen grao maior ou igual que 1. Podemos supor que ningún dos polinomios $g(X)$ ou $h(X)$ ten todos os coeficientes múltiplos dun mesmo número primo p: se todos os coeficientes de $g(X)$ fosen múltiplos dun primo p, ese primo tería que aparecer nalgún denominador de $h(X)$, xa que en caso contrario $f(X)$ non sería primitivo; nese caso, podemos cambiar a factorización de $f(X)$ a $f(X) = (g(X)/p)(ph(X))$, e iterar o proceso ata que ningún dos factores de $g(X)$ ou $h(X)$ fose divisible por ningún primo.

Sexan agora r e s os menores enteiros positivos de maneira que $rg(X)$ e $sh(X)$ están en $\mathbb{Z}[X]$. En particular, temos que $(rs)f(X) = (rg(X)) \cdot (sh(X))$. Se $rs = 1$, entón $r = s = 1$ e a factorización é automaticamente en $\mathbb{Z}[X]$. En caso contrario, sexa p un divisor de rs. Procedendo como no parágrafo anterior, se nin $rg(X)$ nin $sh(X)$ son múltiplos de p, chegamos a unha contradición. Supoñamos entón que $rg(X)$ é múltiplo de p. Se r fose múltiplo de p, (r/p) seguiría cumprindo que $(r/p) \cdot g(X) \in \mathbb{Z}[X]$, o que sería unha contradición coa definición de r; senón, todos os coeficientes de $g(X)$ serían múltiplos de p, e supuxemos que iso non ocorría. Polo tanto, ten que suceder que $rs = 1$.

A outra implicación é evidente, xa que calquera factorización dun polinomio primitivo en $\mathbb{Z}[X]$ dá lugar a unha factorización en $\mathbb{Q}[X]$. $\square$

Este resultado é especialmente útil cando se combina con outros resultados de irredutibilidade en $\mathbb{Z}[X]$, como o chamado *criterio de irredutibilidade de Eisenstein*, no que a idea principal é reducir os coeficientes módulo un primo p.

Proposición 1.7. Sexa p un primo. Sexa $f(X) \in \mathbb{Z}[X]$ un polinomio mónico da forma $f(X) = X^n + a_{n-1}X^{n-1} \dots + a_1 X + a_0$, de maneira que p divide $a_0, a_1, \dots, a_{n-1}$, pero p^2 non divide a_0. Entón $f(X)$ é irredutible.

Demostración. Imos proceder por redución ao absurdo, e supor que $f(X) = g(X)h(X)$, onde $g(X)$ e $h(X)$ son polinomios mónicos non constantes. Escribimos $\overline{f(X)}$ para referirnos á redución de $f(X)$ módulo p, de maneira que $\overline{f(X)} \in (\mathbb{Z}/p\mathbb{Z})[X]$. Polo tanto, en $\mathbb{Z}/p\mathbb{Z})[X]$ tense a igualdade

$$X^n = \overline{f(X)} = \overline{g(X)} \cdot \overline{h(X)},$$

e polo resultado sobre a unicidade da descomposición en factores irredutibles tense que cumprir que $\overline{g(X)} = X^i$ e $\overline{h(X)} = X^j$, onde $i + j = n$ e $i, j \geq 1$. Polo tanto, $g(X) = X^i + b_{i-1}X^{i-1} + \ldots + b_0$ e $h(X) = X^j + c_{j-1}X^{j-1} + \ldots + c_0$, onde b_0 e c_0 son múltiplos de p. Como se cumpre que $a_0 = b_0c_0$, sucede que a_0 é un múltiplo de p^2, que é unha contradición. Por conseguinte, $f(X)$ é irredutible. $\square$

Exemplo. Polo criterio de Eisenstein, o polinomio mónico $p(X) = X^3+5X^2+10X+15$ é irredutible en $\mathbb{Z}[X]$ xa que 5 divide todos os coeficientes (salvo o coeficiente principal) e 5^2 non divide o termo independente. Polo lema de Gauss, o polinomio tamén é irredutible en $\mathbb{Q}[X]$.

Exemplo. O polinomio $X^4 + 3X^3 + 2X^2 + 1$ é irredutible en $\mathbb{Z}[X]$, xa que a súa redución módulo 2 correspóndese co polinomio X^4+X^3+1, que é irredutible. Isto é así xa que non ten factores de grao 1, e de ser produto de factores de grao 2 tería que ser $(X^2 + X + 1)^2 = X^4 + X^2 + 1$, dado que $X^2 + X + 1$ é o único polinomio irredutible de grao 2 en $(\mathbb{Z}/2\mathbb{Z})[X]$. Este exemplo ilustra que a técnica de reducir módulo un primo p pode ser útil tamén en casos nos que o polinomio non é Eisenstein.

O último caso no que estaremos interesados é o dos polinomios con coeficientes en $\mathbb{Z}/p\mathbb{Z}$, sendo p un número primo. Non faremos ningún resultado xeral, pero si comentaremos algúns exemplos que ilustran diferentes posibilidades.

Exemplo. O polinomio $X^4 + 3X^2 + 2 \in (\mathbb{Z}/5\mathbb{Z})[X]$ ten as raíces 2 e 3. Polo tanto, por Ruffini, vemos que

$$X^4 + 3X^2 + 2 = (X - 2)(X - 3)(X^2 + 2),$$

e o polinomio $X^2 + 2$ é irredutible xa que non ten ningunha raíz.
O polinomio $X^4 + 3X^2 + 2 \in (\mathbb{Z}/7\mathbb{Z})[X]$ factoriza como

$$X^4 + 3X^2 + 2 = (X^2 + 1)(X^2 + 2),$$

pero tanto $X^2 + 1$ como $X^2 + 2$ son irredutibles.
O polinomio $X^4 + 3X^2 + 2 \in (\mathbb{Z}/17\mathbb{Z})[X]$ ten 4 raíces: 4, 7, 10 e 13. Por conseguinte,

$$X^4 + 3X^2 + 2 = (X - 4)(X - 7)(X - 10)(X - 13).$$

Un tipo de polinomios cos que se traballará con frecuencia son os da forma $f(X) = X^n - 1 \in \mathbb{C}[X]$. Polo tanto, convén introducir a seguinte notación.

Definición 1.7. Unha raíz complexa ζ do polinomio $f(X) = X^n - 1$ chámase *raíz n-ésima da unidade.* Se n é o menor enteiro positivo para o que $\zeta^n = 1$, dise que ζ é unha *raíz primitiva n-ésima da unidade.*

Exemplo. As raíces cuartas da unidade son ± 1 e $\pm i$, xa que son as solucións de $X^4 - 1 = 0$. Entre esas, as raíces primitivas son $\pm i$.

1.5. Problemas

Problema 1.1. Determinar un polinomio $p(X) \in \mathbb{C}[X]$ de grao mínimo tal que $X^2+1 \mid p(X)$ e $X^3+1 \mid p(X)-1$.

Solución. Da primeira condición temos que $p(X) = (X^2+1)q(X)$. Da segunda, temos que X^3+1 divide $(X^2+1)q(X)-1$. Polo tanto,

$$(X^2+1)q(X) - 1 = (X^3+1)r(X).$$

Imos amosar agora tres maneiras distintas de resolver o problema a partir de aquí.

(i) A partir da igualdade anterior, podemos aplicar a identidade de Bézout (algoritmo de Euclides xeneralizado), dado que o problema consiste en encontrar polinomios $q(X)$ e $r(X)$ tales que

$$(X^2+1)q(X) - (X^3+1)r(X) = 1,$$

o cal vai ser posible dado que X^2+1 e X^3+1 non teñen ningún factor irredutible en común. Realizando a división euclidiana temos que

$$X^3+1 = (X^2+1)X + (1-X), \quad X^2+1 = (1-X)(-X-1)+2,$$

co cal

$$\begin{aligned} 2 &= X^2+1+(1-X)(X+1) \\ &= X^2+1+((X^3+1)-X(X^2+1))(X+1) \\ &= (X^3+1)(X+1)+(X^2+1)(-X^2-X+1). \end{aligned}$$

Dividindo por 2 e reordeando os termos quédanos que

$$(X^2+1)\cdot\Big(\frac{-X^2-X+1}{2}\Big) - 1 = (X^3+1)\cdot\Big(\frac{-X-1}{2}\Big).$$

Polo tanto, o polinomio buscado é

$$p(X) = \frac{-X^4-X^3-X+1}{2}.$$

Observamos tamén que $q(X)$ (e polo tanto $p(X)$) non pode ter grao menor, dado que calquera outro que cumpra a ecuación sería da forma $\frac{-X^2-X+1}{2} + a(X)\cdot(X^3+1)$, con $a(X) \in K[X]$, e se $a(X) \neq 0$ o polinomio sempre será de grao polo menos 3. Isto proba ademais que o polinomio que achamos é o único de grao 4 que cumpre a condición.

(ii) Na ecuación $(X^2+1)q(X)-1 = (X^3+1)r(X)$, o lado esquerdo ten que se anular nas raíces de $X^3+1 = (X+1)(X^2-X+1) = (X+1)(X+\xi)(X+\bar{\xi})$, con $\xi = \frac{1+\sqrt{-3}}{2}$. Impondo estas condicións, e usando que $\xi^2 = \xi - 1$, resulta que $2q(-1) = 1$, $\xi q(\xi) = 1$ e $\bar{\xi}q(\bar{\xi}) = 1$.

Dado que temos tres condicións, imos comezar buscando un polinomio de grao 2 (tres graos de liberdade). Pomos $q(X) = aX^2+bX+c$. Impondo a condición en ξ e $\bar{\xi}$, temos

$$\xi(b+c)-(a+b) = 1, \quad \bar{\xi}(b+c)-(a+b) = 1.$$

Restando as dúas ecuacións, $(b+c)(\xi-\bar{\xi})=0$, co cal vemos que $b+c=0$, e polo tanto $a+b=-1$. A condición en -1, á súa vez, dinos que $2a-2b+2c=1$. Polo tanto,

$$\begin{aligned} 2a-2b+2c &= 1 \\ b+c &= 0 \\ a+b &= -1. \end{aligned}$$

Isto dinos que $(a,b,c)=(-1/2,-1/2,1/2)$. Polo tanto, $q(X)=\frac{-X^2-X+1}{2}$ e

$$p(X)=\frac{-X^4-X^3-X+1}{2}.$$

Resulta doado comprobar que non existe ningún $q(X)$ de grao 1 que cumpra a condición.

(iii) Por último, podemos escribir

$$(aX^2+bX+c)(X^2+1)-1=(X^3+1)(aX+d)$$

e resolver o sistema de catro ecuacións, obtendo o mesmo resultado e sen usar explicitamente nin a identidade de Bézout nin propiedades sobre os números complexos. Neste caso, observamos tamén que non é posible que $q(X)$ teña grao 1. De ser así,

$$(aX+b)(X^2+1)-1=c(X^3+1).$$

De aquí, $c=a$, $b=0$, $a=0$ e $b-1=c$, que é un sistema incompatible. Polo tanto, o menor polinomio ten grao 4 e é o presentado nos parágrafos anteriores.

Problema 1.2. Se $p(X)\in\mathbb{Z}[X]$ e $p(r/s)=0$, con $(r,s)=1$, demostrar que $r-s\mid p(1)$ e $r+s\mid p(-1)$.

Solución. Imos resolver o problema sen empregar directamente o Lema de Gauss, senón adaptando a súa demostración a este caso concreto. Da condición do enunciado, podemos escribir

$$p(X)=\Big(X-\frac{r}{s}\Big)q(X)=\frac{(sX-r)q(X)}{s},$$

onde $q(X)\in\mathbb{Q}[X]$ (só podemos aplicar os resultados de división sobre polinomios en $\mathbb{Q}$, non en $\mathbb{Z}$). Sexa $M\geq 1$ o menor enteiro positivo tal que $Mq(X)\in\mathbb{Z}[X]$, e poñamos $Mq(X)=b_nX^n+b_{n-1}X^{n-1}+\ldots+b_1X+b_0$. Entón,

$$Msp(X)=(sX-r)(b_nX^n+\ldots+b_0),$$

onde todos os b_i son enteiros. Imos supoñer que $M>1$ e chegar a unha contradición; para iso, consideremos que ten un divisor primo q. Observemos que se q divide todos os b_i, entón poderiamos cambiar M por M/q, o cal contradiciría a condición de minimalidade sobre M. Do mesmo xeito q non pode dividir simultaneamente r e s, dado que son coprimos. Supoñamos que $q\nmid s$, sendo o caso no que $q\nmid r$ completamente análogo. Sexa $0\leq k\leq n$ o maior enteiro tal que $q\nmid b_k$. Se $k=n$, entón teriamos que o coeficiente en X^{n+1} é sb_n, que non é múltiplo de q, unha contradición. Máis en xeral, o coeficiente con X^{k+1} é sb_k-rb_{k+1}. Pola definición de k, temos que b_{k+1} é múltiplo de q, e como o coeficiente en X^{k+1} tamén o debe ser, sb_k ten que ser múltiplo de q. Iso, porén, non

é posible, dado que por construción tanto s como b_k son relativamente primos con q. Polo tanto, $M = 1$ e $q(X) \in \mathbb{Z}[X]$.
Agora as preguntas do enunciado son sinxelas. Avaliando en $X = 1$, temos que $s \cdot p(1) = (s-r)q(1)$, que como $q(1)$ é enteiro, é unha igualdade de números enteiros. Temos que $r - s$ divide $s \cdot p(1)$, pero polo algoritmo de Euclides $(s, r-s) = (r, s) = 1$, o cal $r - s$ divide $p(1)$. Do mesmo xeito, avaliando en $X = -1$, temos que $s \cdot p(-1) = (-s-r)q(-1)$, e novamente temos que $r + s$ divide $s \cdot p(-1)$; como $(s, r+s) = (r, s) = 1$, $r + s$ divide $p(-1)$.

Problema 1.3. Os seguintes apartados son independentes entre si.

(a) Descompoñer $X^4 + a^2 \in \mathbb{R}[X]$ en factores irredutibles.

(b) Descompoñer $(X+1)^n + (X-1)^n \in \mathbb{C}[X]$ en factores lineares.

(c) Demostrar que o polinomio $X^4 - 9X^2 - 18X - 3 \in \mathbb{Q}[X]$ é irredutible. Podemos dicir o mesmo do polinomio $X^4 - 9X^2 - 18X - 9 \in \mathbb{Q}[X]$?

(d) Demostrar que o polinomio $X^5 + X^4 - 4X^3 - 3X^2 + 3X + 1 \in \mathbb{Q}[X]$ é irredutible.

Solución. (a) Supoñamos sen perda de xeralidade que $a > 0$ (senón, cambiámolo por $-a$), e sexa $\sqrt{a}$ a raíz cadrada positiva de a. Sobre os complexos, as raíces do polinomio son $\zeta_8\sqrt{a}$, $\zeta_8^3\sqrt{a}$, $\zeta_8^5\sqrt{a}$ e $\zeta_8^7\sqrt{a}$, onde ζ_8 é unha raíz oitava primitiva da unidade. En concreto, $\zeta_8 = \frac{\sqrt{2}+i\sqrt{2}}{2}$. Agrupando os factores conxugados, resulta que

$$(X - \zeta_8\sqrt{a})(X - \zeta_8^7\sqrt{a}) = X^2 - \sqrt{2a}X + a$$

e

$$(X - \zeta_8^3\sqrt{a})(X - \zeta_8^5\sqrt{a}) = X^2 + \sqrt{2a}X + a.$$

Polo tanto, para un a xenérico,

$$X^4 + a^2 = (X^2 + \sqrt{|2a|}X + |a|)(X^2 - \sqrt{|2a|}X + |a|).$$

(b) Pondo

$$(X+1)^n = -(X-1)^n = (e^{\pi i/n})^n(X-1)^n = (e^{\pi i/n}X - e^{\pi i/n})^n,$$

temos dous números que ao elevalos a n dan o mesmo. Polo tanto, o seu cociente é unha raíz n-ésima da unidade, é dicir,

$$X + 1 = e^{2\pi ik/n}(e^{\pi i/n}X - e^{\pi i/n}), \quad k = 0, 1, \ldots, n-1.$$

Illando X, resulta

$$X = \frac{1 + e^{\pi i(2k+1)/n}}{e^{\pi i(2k+1)/n} - 1}, \quad k = 0, 1, \ldots, n-1.$$

Entón,

$$(X+1)^n + (X-1)^n = \prod_{k=0}^{n-1}\Big(X - \frac{1 + e^{\pi i(2k+1)/n}}{e^{\pi i(2k+1)/n} - 1}\Big).$$

(c) Para a primeira parte, a observación clave baséase en que un polinomio con coeficientes enteiros e relativamente primos entre si en $\mathbb{Q}[X]$ é irredutible se e soamente se é irredutible en $\mathbb{Z}[X]$. Por outro lado, se $f(X) = g(X)h(X)$, podemos escribir $\overline{f(X)}$ para a redución módulo 3, e temos que

$$\overline{f(X)} = \overline{g(X)} \cdot \overline{h(X)}.$$

Neste caso, a redución do polinomio módulo 3 é simplemente X^4, co que de existir unha factorización as correspondentes reducións serían da forma X^i, e en particular o termo independente sería 0. Iso quere dicir que o termo independente de $g(X)$ e de $h(X)$ é múltiplo de 3, e o termo independente do polinomio orixinal sería entón múltiplo de 9, que non é posible.

Para a segunda parte, é suficiente con observar que

$$X^4 - 9X^2 - 18X - 9 = (X^2 + 3X + 3)(X^2 - 3X - 3).$$

(d) Como no apartado anterior, é suficiente demostrar que $f(X)$ é irredutible sobre $\mathbb{Z}[X]$. Para iso, basta con encontrar un primo p de maneira que $\bar{f}(X)$ sexa irredutible en $\mathbb{F}_p[X]$. Imos considerar $p = 2$. Como $f(X)$ é de grao 5, se fose reducible necesariamente a súa descomposición ten que contar un factor de grao 1 ou de grao 2. É inmediato comprobar que o polinomio non é divisible nin por X nin por $X + 1$, dado que $\bar{f}(0) = \bar{f}(1) = 1$. Imos comprobar que $\bar{f}(X)$ tampouco é divisible polo único polinomio irredutible de grao 2, $X^2 + X + 1$; se o fose,

$$X^5 + X^4 + X^2 + X + 1 = (X^2 + X + 1)(X^3 + aX^2 + bX + c).$$

Igualando termos, vemos que $a = 0$, $b = 1$ e $c = 0$, pero iso é incompatible coa igualdade de termos independentes. Por tanto, $\bar{f}$ é irredutible en $\mathbb{F}_2[X]$ e tamén en $\mathbb{Q}[X]$.

Problema 1.4. Un número $\alpha \in \mathbb{C}$ chámase alxébrico se existe un polinomio mónico $p(X) \in \mathbb{Q}[X]$ de maneira que $p(\alpha) = 0$.

(a) Probar que os números $\sqrt{2}, \sqrt{3} + 1, \sqrt[3]{5} - 3, \sqrt{2} + \sqrt{3}$ son alxébricos e encontrar un polinomio con coeficientes racionais do cal sexan ceros.

(b) Demostrar que para todo número alxébrico α existe un único polinomio mónico de grao mínimo $p(X) \in \mathbb{Q}[X]$ de maneira que $p(\alpha) = 0$. Demostrar que calquera outro polinomio $q(X)$ tal que $q(\alpha) = 0$ cumpre que $q(X) = p(X)a(X)$, para algún $a(X) \in \mathbb{Q}[X]$. Para os números do apartado (a), cal é ese polinomio $p(X)$?

Solución. (a) No caso de $\sqrt{2}$ consideramos $X^2 - 2$. Para $\alpha = \sqrt{3} + 1$, pomos $\alpha - 1 = \sqrt{3}$ e elevamos ao cadrado, obtendo $\alpha^2 - 2\alpha - 2 = 0$, co cal o polinomio é $X^2 - 2X - 2$. Do mesmo xeito, se $\beta = \sqrt[3]{5} - 3$, $\beta + 3 = \sqrt[3]{5}$ e elevando ao cubo $\beta^3 + 9\beta^2 + 27\beta + 22 = 0$, co cal o polinomio é $X^3 + 9X^2 + 27X + 22 = 0$. Por último, se $\gamma = \sqrt{2} + \sqrt{3}$, pomos $\gamma - \sqrt{2} = \sqrt{3}$ e elevando ao cadrado temos $\gamma^2 - 1 = 2\sqrt{2}\gamma$; elevando novamente ao cadrado, $\gamma^4 - 10\gamma^2 + 1$, co cal o polinomio buscado é $X^4 - 10X^2 + 1$.

(b) Sexa α un número alxébrico e $p(X)$ un polinomio de grao mínimo con $p(\alpha) = 0$. Sexa $q(X)$ outro polinomio con $q(\alpha) = 0$. En $\mathbb{Q}[X]$ aplicamos o algoritmo da división, e podemos escribir $q(X) = p(X)a(X) + b(X)$, con $b(X)$ de grao menor ao grao de $p(X)$. Avaliando en $X = \alpha$ temos que $q(\alpha) = p(\alpha)a(\alpha) + b(\alpha)$, e como $q(\alpha) = p(\alpha) = 0$, ten que ser $b(\alpha) = 0$. Polo tanto, $b(X)$ sería un polinomio de grao menor a $p(X)$ que anula α, o cal non é posible.

No caso do apartado anterior, os polinomios que demos son xa os de menor grao. En case todos os casos é un comprobación rutineira ver que non hai ningún de grao menor; imos traballar o caso de $X^4 - 10X^2 + 1$, que é o menos sinxelo. Para iso, é suficiente ver que o polinomio é irredutible en $\mathbb{Z}$. Como os únicos divisores do termo independente son ± 1 e ningún é unha raíz, non pode haber factores lineais. De haber factores de grao 2, sucedería que

$$(X^4 - 10X^2 + 1) = (X^2 + aX + b)(X^2 + cX + d),$$

con $a, b, c, d \in \mathbb{Z}$. Polo tanto, ou $b = d = 1$ ou $b = d = -1$. Como o termo en X^3 é 0, $a = -c$; considerando o termo en X^2, temos que $b + d - a^2 = -10$, co cal $a^2 = 10 + b + d$. Se $b = d = 1$, $a^2 = 12$; e se $b = d = -1$, entón $a^2 = 8$. En ningún dos casos é posible por tanto ter unha factorización en $\mathbb{Z}[X]$, e iso conclúe o problema.

Problema 1.5. Consideremos o polinomio $P(X) = X^4 + 4$. É $P(X)$ irredutible en $\mathbb{Q}[X]$? Dar a factorización de $P(X)$ en $(\mathbb{Z}/5\mathbb{Z})[X]$.

Solución. En primeiro lugar observamos que o polinomio non ten raíces enteiras, xa que estas estarían entre os divisores do seu termo independente. Buscamos polo tanto factorizacións da forma

$$X^4 + 4 = (X^2 + aX + b)(X^2 + cX + d).$$

Igualando termos, quédanos que unha solución posible é

$$X^4 + 4 = (X^2 + 2x + 2)(X^2 - 2x + 2).$$

Por conseguinte, $P(X)$ non é irredutible. En $(\mathbb{Z}/5\mathbb{Z})[X]$, temos que os números 1, 2, 3 e 4 son raíces, xa que $1^4 + 4 = 5$, $2^4 + 4 = 20$, $3^4 + 4 = 85$ e $4^4 + 4 = 260$. Polo tanto,

$$X^4 + 4 = (X - 1)(X - 2)(X - 3)(X - 4).$$

Problema 1.6. Consideremos o polinomio $P(X) = X^8 - 1$. Dar a factorización de $P(X)$ en $\mathbb{C}[X]$, en $\mathbb{R}[X]$ e en $(\mathbb{Z}/3\mathbb{Z})[X]$.

Solución. No caso de $\mathbb{C}$, temos directamente que

$$X^8 - 1 = \prod_{k=0}^{7} (X - e^{\frac{2\pi i k}{8}}),$$

onde $e^{\frac{2\pi i}{8}} = \frac{1+i}{\sqrt{2}}$ é unha raíz primitiva oitava da unidade.
No caso de $\mathbb{R}$, temos dous factores de grao 1, $X - 1$ e $X + 1$, e tres factores de grao 2 que se obteñen agrupando os factores complexos correspondentes a raíces complexas conxugadas. Como $(X - i)(X + i) = X^2 + 1$ e

$$\left(X - \frac{\pm 1 + i}{\sqrt{2}}\right)\left(X - \frac{\pm 1 - i}{\sqrt{2}}\right) = X^2 \mp \sqrt{2}X + 1,$$

concluímos que

$$X^8 - 1 = (X-1)(X+1)(X^2+1)(X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1).$$

Finalmente, en $\mathbb{Z}/3\mathbb{Z}$, comezamos empregando as identidades notables e temos que

$$X^8 - 1 = (X-1)(X+1)(X^2+1)(X^4+1).$$

O polinomio X^2+1 é irredutible xa que non ten raíces. O polinomio X^4+1 tampouco ten raíces, pero cómpre analizar se factoriza ou non como produto de dous polinomios de grao 2. Impondo que

$$X^4 + 1 = (X^2 + aX + b)(X^2 + cX + d)$$

e resolvendo o sistema, vemos que $X^4+1 = (X^2+X+2)(X^2-X+2)$, polo que

$$X^8 - 1 = (X-1)(X+1)(X^2+1)(X^2+X+2)(X^2-X+2).$$

Problema 1.7. Sexa p un número primo, e consideramos o polinomio $q(X) = X^p - X$. Factorizar $q(X)$ en $(\mathbb{Z}/p\mathbb{Z})[X]$.

Solución. Polo pequeno teorema de Fermat, sabemos que $a^{p-1} \equiv 1 \pmod{p}$ para todo enteiro a relativamente primo con p. Multiplicando por a, temos que $a^p \equiv a \pmod{p}$. Esta congruencia tamén é certa para $a \equiv 0 \pmod{p}$, xa que ambos lados son congruentes con 0. Polo tanto, os elementos $0, 1, \ldots, p-1$ son todos eles raíces de $X^p - X$. Iso quere dicir que $X(X-1)\cdots(X-(p-1))$ é un divisor de $X^p - X$. Porén, se un polinomio mónico divide a outro polinomio mónico do mesmo grao, necesariamente teñen que ser iguais. Concluímos, polo tanto, que

$$X^p - X = X(X-1)\cdots(X-(p-1)).$$

Problema 1.8. Determinar se as seguintes afirmacións son verdadeiras ou falsas.

(a) Un polinomio de grao 3 en $\mathbb{Q}[X]$ é irredutible se, e soamente se, non ten ningunha raíz.

(b) Sexa n un número enteiro positivo. Un polinomio con coeficientes en $\mathbb{Z}/n\mathbb{Z}$ ten, como moito, tantas raíces como o seu grao.

(c) Un polinomio en $\mathbb{Q}[X]$ sempre descompón en produto de polinomios irredutibles de grao 1 ou de grao 2.

(d) Para todo número primo p, o número de polinomios mónicos irredutibles de grao 2 en $(\mathbb{Z}/p\mathbb{Z})[X]$ é $\frac{p(p-1)}{2}$.

(e) Existe un polinomio $p(X) \in \mathbb{Z}[X]$ de xeito que

$$X^6 = 1 + (X-1)(X-2)(X-3)(X-4)(X-5)(X-6) + 7p(X).$$

(f) Non existe ningún polinomio $P(X) \in \mathbb{Q}[X]$ de xeito que

$$\begin{cases} P(X) \equiv X^2 & \pmod{X^2 + X}, \\ P(X) \equiv 2X & \pmod{X^2 - X}. \end{cases}$$

Solución. (a) Verdadeira. Se un polinomio $f(X)$ non é irredutible, entón $f(X) = g(X)h(X)$, onde $g(X)$ e $h(X)$ son polinomios non constantes. Entón, o grao de f é a suma dos graos de g e h, e como dita suma é 3 e ambos son maioires que 0, un deles terá necesariamente grao 1.

(b) Falsa. En $\mathbb{Z}/6\mathbb{Z}$, o polinomio $(X-2)(X-3)$ ten grao 2, pero ten 4 raíces: 0, 2, 3 e 5. Outra opción é considerar, por exemplo, o polinomio $X^2 - 1$ en $(\mathbb{Z}/8\mathbb{Z}$, que ten as raíces 1, 3, 5 e 7.

(c) Falsa. O polinomio $X^3 + 2$ é irredutible porque é 2-Eisenstein.

(d) Verdadeira. Un polinomio mónico de grao 2 é da forma $X^2 + aX + b$, polo que hai $p \cdot p = p^2$ en total. Os redutibles son os que teñen unha raíz dobre, un total de p; ou dúas raíces diferentes, un total de $\binom{p}{2}$. Polo tanto, o número de mónicos irredutibles é
$$p^2 - p - \binom{p}{2} = \binom{p}{2} = \frac{p(p-1)}{2}.$$

(e) Verdadeira. Tense que, en $(\mathbb{Z}/7\mathbb{Z})[X]$, $X^6 - 1$ ten as raíces 1, 2, 3, 4, 5 e 6, polo pequeno teorema de Fermat. Polo tanto, a diferenza
$$X^6 - 1 - (X-1)(X-2)(X-3)(X-4)(X-5)(X-6)$$
é un polinomio en $7\mathbb{Z}[X]$.

(f) Falsa. A primeira condición implica que se ten $P(X) \equiv 0 \pmod{X}$ e $P(X) \equiv 1 \pmod{X+1}$; a segunda, que $P(X) \equiv 0 \pmod{X}$ e $P(X) \equiv 2 \pmod{X-1}$. Polo tanto, obtense o sistema de congruencias
$$\begin{cases} P(X) \equiv 0 & (\text{mód } X), \\ P(X) \equiv 1 & (\text{mód } X+1), \\ P(X) \equiv 2 & (\text{mód } X-1); \end{cases}$$
como os polinomios X, $X+1$ e $X-1$ son coprimos dous a dous, polo teorema chinés do resto, hai solución.

Capítulo 2

Aplicacións multilineais e determinantes

Sexa K un corpo e E un K-espazo vectorial de dimensión n con base $\mathcal{B} = \{v_1, \ldots, v_n\}$. O espazo dual de E, que denotamos como E^*, é o conxunto das aplicacións lineais $f\colon E \to K$. Á base $\mathcal{B}$ podémoslle asociar unha base do espazo dual, $\mathcal{B}^* = \{v_1^*, \ldots, v_n^*\}$, que cumpre $v_i^*(v_j) = \delta_{i,j}$, onde $\delta_{i,j}$ é a delta de Kronecker, que vale 1 se $i = j$ e 0 no resto dos casos.
O obxectivo central deste capítulo é dar unha definición conceptual do determinante, impondo que cumpra determinadas condicións relativas á multilinealidade e ás permutacións de filas ou columnas. Para iso, deberemos desenvolver os conceptos de *aplicación multilineal* e *grupo simétrico.*

2.1. Aplicacións multilineais

Imos comezar este tema traballando unha xeralización da noción de espazo dual.

Definición 2.1. Unha *aplicación p-lineal* ou *tensor p-covariante* de E é unha aplicación

$$f\colon E \times \overset{p}{\cdots} \times E \longrightarrow K$$

que é lineal en cada unha das p compoñentes.
O conxunto das aplicacións p-lineais denótase como $\mathcal{L}_p(E^p, K)$ ou simplemente $T_p(E)$. No caso $p = 1$, recuperamos o espazo dual: $T_1(E) = E^*$.

É sinxelo ver que a suma de dúas aplicacións multilineais tamén o é, e o mesmo sucede para o produto por escalares. Polo tanto, $T_p(E)$ ten estrutura de espazo vectorial. Ese é o contido da seguinte proposición.

Proposición 2.1. O conxunto dos tensores p-covariantes, $T_p(E)$, ten estrutura de espazo vectorial.

Demostración. Cómpre ver que dadas $\varphi, \psi \in T_p(E)$ e $\lambda \in K$, tanto $\varphi + \psi$ como $\lambda\varphi$ son elementos de $T_p(E)$. Imos comprobar unicamente o caso de $\varphi + \psi$, dado que o outro é totalmente análogo.

Á súa vez, ver que $\varphi+\psi$ é multilineal require ver que respecta as sumas e os produtos por escalares en cada variable. Para a primeira, observamos que

$$\begin{aligned}(\varphi+\psi)(x_1,\ldots,x_i+x_i',\ldots,x_p) &= \varphi(x_1,\ldots,x_i+x_i',\ldots,x_p)\\ &+\psi(x_1,\ldots,x_i+x_i',\ldots,x_p)\\ &=\varphi(x_1,\ldots,x_i,\ldots,x_p)+\varphi(x_1,\ldots,x_i',\ldots,x_p)\\ &+\psi(x_1,\ldots,x_i,\ldots,x_p)+\psi(x_1,\ldots,x_i',\ldots,x_p)\\ &=(\varphi+\psi)(x_1,\ldots,x_i,\ldots,x_p)\\ &+(\varphi+\psi)(x_1,\ldots,x_i',\ldots,x_p).\end{aligned}$$

A primeira igualdade séguese da definición de $\varphi+\psi$; a segunda da linealidade de φ e ψ; e a terceira da propiedade conmutativa de K e novamente da definición de $\varphi+\psi$. A segunda comprobación faise do mesmo xeito, observando que

$$\begin{aligned}(\varphi+\psi)(x_1,\ldots,\mu x_i,\ldots,x_p) &= \varphi(x_1,\ldots,\mu x_i,\ldots,x_p)+\psi(x_1,\ldots,\mu x_i,\ldots,x_p)\\ &=\mu\cdot\varphi(x_1,\ldots,x_i,\ldots,x_p)+\mu\cdot\psi(x_1,\ldots,x_i,\ldots,x_p)\\ &=\mu\cdot(\varphi+\psi)(x_1,\ldots,x_i,\ldots,x_p).\end{aligned}$$

□

Ademais, dadas dúas aplicacións multilineais, tamén é posible definir o seu produto. É o chamado produto tensorial.

Definición 2.2. Sexa $f\in T_p(E)$ e $g\in T_q(E)$, con $p,q\geq 0$. O *produto tensorial* de f e g, que denotamos por $f\otimes g$, é a aplicación

$$f\otimes g\colon E\overset{p}{\cdots}\times E\times E\overset{q}{\cdots}\times E\longrightarrow K$$

que cumpre que, se $x=(x_1,\ldots,x_p)\in E^p$ e $y=(y_1,\ldots,y_q)\in E^q$, entón $(f\otimes g)(x,y)=f(x)g(y)$.

O seguinte resultado baséase nunha comprobación rutineira.

Proposición 2.2. O produto tensorial $\varphi\otimes\psi$ é un elemento de $T_{p+q}(E)$.

Demostración. Hai que comprobar que $\varphi\otimes\psi$ respecta a suma e o produto por escalares en cada unha das variables. Faremos unicamente a comprobación para a suma, dado que a outra faise seguindo o mesmo razoamento. Sen perda de xeneralidade, imos comprobar a bilinealidade na compoñente i-ésima, supondo que $1\leq i\leq p$, sendo o caso $p+1\leq i\leq p+q$ completamente análogo.

$$\begin{aligned}(\varphi\otimes\psi)(x_1,\ldots,x_i+x_i',\ldots,x_{p+q}) &= \varphi(x_1,\ldots,x_i+x_i',\ldots,x_p)\cdot\psi(x_{p+1},\ldots,x_{p+q})\\ &=(\varphi(x_1,\ldots,x_i,\ldots,x_p)+\varphi(x_1,\ldots,x_i',\ldots,x_p))\\ &\cdot\psi(x_{p+1},\ldots,x_{p+q})\\ &=(\varphi\otimes\psi)(x_1,\ldots,x_i,\ldots,x_{p+q})\\ &+(\varphi\otimes\psi)(x_1,\ldots,x_i',\ldots,x_{p+q}).\end{aligned}$$

A primeira igualdade séguese da definición de $\varphi\otimes\psi$; a segunda da linealidade φ; e a terceira da propiedade distributiva de K e novamente da definición de $\varphi\otimes\psi$. □

Un dos exemplos máis importantes ocorre cando $f, g \in E^* = T_1(E)$. Entón, $f \otimes g\colon E \times E \to K$ é a aplicación lineal dada por $(u \otimes v)(x, y) = u(x)v(y)$, e é un tensor 2-covariante. Estes obxectos serán estudados con máis detalle en capítulos posteriores.

Observación. O produto tensorial non é conmutativo. Por exemplo, se E ten dimensión 2, podemos considerar $f = v_1^* + v_2^*$ e $g = v_1^*$. Entón, $(f \otimes g)(v_1, v_2) = 0$, pero $(g \otimes f)(v_1, v_2) = 1$. Polo tanto, $f \otimes g$ e $g \otimes f$ son dous elementos diferentes de $T_2(E)$.

Porén, o produto tensorial si é asociativo e cumpre tamén a propiedade distributiva. Imos resumir algunhas das súas propiedades, cuxa comprobación é inmediata. Sexan $f, f' \in T_p(E)$, $g, g' \in T_q(E)$, $h \in T_r(E)$ e $\lambda \in K$.

(a) $(f \otimes g) \otimes h = f \otimes (g \otimes h)$.

(b) $(\lambda f) \otimes g = \lambda(f \otimes g) = f \otimes (\lambda g)$.

(c) $f \otimes (g + g') = f \otimes g + f \otimes g'$ e $(f + f') \otimes g = f \otimes g + f' \otimes g$.

O noso seguinte obxectivo é dar unha base de $T_p(E)$. Por simplicidade, imos comezar traballando o caso no que E é de dimensión 2 e $p = 2$. Consideramos a base $\{v_1, v_2\}$ de E e imos considerar o seguinte conxunto de elemento de $T_2(E)$:

$$\mathcal{B} = \{v_1^* \otimes v_1^*, v_1^* \otimes v_2^*, v_2^* \otimes v_1^*, v_2^* \otimes v_2^*\}.$$

Para ver que $\mathcal{B}$ é base, cómpre ver que xeran o espazo $T_2(E)$ e que son linealmente independentes. Para ver que son xeradores, sexa $f \in T_2(E)$. Entón

$$f(av_1 + bv_2, cv_1 + dv_2) = acf(v_1, v_1) + adf(v_1, v_2) + bcf(v_2, v_1) + bdf(v_2, v_2).$$

Polo tanto, podemos pór

$$f = f(v_1, v_1) \cdot v_1^* \otimes v_1^* + f(v_1, v_2) \cdot v_1^* \otimes v_2^* + f(v_2, v_1) \cdot v_2^* \otimes v_1^* + f(v_2, v_2) \cdot v_2^* \otimes v_2^*.$$

Para ver que son independentes, procederemos por contradición, supondo que temos unha igualdade

$$\varphi = a_{11} \cdot v_1^* \otimes v_1^* + a_{12} \cdot v_1^* \otimes v_2^* + a_{21} \cdot v_2^* \otimes v_1^* + a_{22} \cdot v_2^* \otimes v_2^* = 0.$$

Considerando a avaliación nos diferentes vectores da base, temos que $\varphi(v_1, v_1) = a_{11} = 0$, $\varphi(v_1, v_2) = a_{12} = 0$, $\varphi(v_2, v_1) = a_{21} = 0$ e $\varphi(v_2, v_2) = a_{22} = 0$. Isto amosa que todos os coeficientes da combinación lineal son 0.
A demostración esténdese de maneira inmediata ao caso xeral.

Proposición 2.3. Sexa E un espazo vectorial de dimensión n e sexa $\{v_1, \ldots, v_n\}$ unha base súa. O espazo vectorial $T_p(E)$ ten dimensión n^p e unha base está dada polos vectores

$$\{v_{i_1}^* \otimes \overset{p}{\cdots} \otimes v_{i_p}^*\}_{1 \le i_1, \ldots, i_p \le n}.$$

Demostración. Dado un elemento $f \in T_p(E)$, podemos escribir

$$f = \sum_{1 \le i_1, \ldots, i_p \le n} f(v_{i_1}, \ldots, v_{i_p}) v_{i_1}^* \otimes \overset{p}{\cdots} \otimes v_{i_p}^*,$$

co cal o conxunto proposto xera todo o espazo. De xeito similar, supoñamos que existe unha combinación lineal igual a cero da forma

$$\varphi = \sum_{1 \leq i_1,\ldots,i_p \leq n} a_{i_1,\ldots,i_p} \cdot v_{i_1}^* \otimes \overset{p}{\cdots} \otimes v_{i_p}^* = 0.$$

Avaliando nas p-tuplas formadas por vectores da base, vemos que

$$0 = \varphi(v_{i_1}, \ldots, v_{i_p}) = a_{i_1,\ldots,i_p},$$

co cal todos os coeficientes terían que ser 0. □

O seguinte obxectivo é definir o determinante, pero o espazo $T_p(E)$ resulta demasiado grande para os nosos propósitos. Volvamos ao caso no que E ten dimensión 2 e $p = 2$. Nese caso, $T_p(E)$ ten dimensión 4. O determinante vai ser, efectivamente, unha aplicación multilineal $D \colon E \times E \to K$. Porén, ímoslle esixir, polo menos, dúas propiedades adicionais:

(a) $D(v, v) = 0$, para todo $v \in E$.

(b) $D(v, w) = -D(w, v)$, para calquera $v, w \in E$.

De feito, ambas condicións son *practicamente* equivalentes, como recolle a seguinte proposición.

Proposición 2.4. Sexa $D \in T_p(E)$ unha aplicación lineal que cumpre que $D(v, v) = 0$ para todo $v \in E$. Entón, $D(v, w) = -D(w, v)$ para calquera $v, w \in E$. O recíproco tamén é certo se o corpo ten característica diferente de 2, isto é, $2 \neq 0$ en K.

Demostración. Supoñamos que $D(v, v) = 0$ para todo $v \in E$. En particular, $D(v + w, v + w) = 0$. Desenvolvendo a expresión temos que

$$D(v, v) + D(w, w) + D(v, w) + D(w, v) = 0.$$

Como os dous primeiros sumandos son cero por hipótese, cúmprese que $D(v, w) = -D(w, v)$. Para o recíproco, supoñamos que $D(v, w) = -D(w, v)$ para todo $v, w \in E$. Collendo $v = w$, temos que $2D(v, v) = 0$, que implica que $D(v, v) = 0$ sempre que a característica sexa diferente de 2. □

Estas condicións de linealidade e antisimetría simplemente expresan que o determinante dunha matriz coas dúas columnas iguais é igual a 0, e que permutar dúas columnas cambia de signo o determinante. Veremos que o conxunto das aplicacións multilineais que cumpren estas dúas novas propiedades ten dimensión 1, co cal é suficiente dar os valores de D sobre unha base ordenada (v_1, v_2).
Para estender esta construción, cómpre introducir o grupo de permutacións (grupo simétrico) e o concepto de *signo* dunha permutación.

2.2. O grupo simétrico

Definición 2.3. Unha *permutación* de p elementos é unha aplicación bixectiva

$$s \colon \{1, \ldots, p\} \longrightarrow \{1, \ldots, p\}.$$

Ao conxunto destas permutacións chámaselle grupo simétrico e denótase como S_p.

Un elemento do grupo simétrico adoitamos representalo da forma

$$s = \begin{pmatrix} 1 & 2 & \cdots & p \\ s(1) & s(2) & \cdots & s(p) \end{pmatrix}.$$

O nome de grupo simétrico vén do feito de que efectivamente, coa composición, S_p ten estrutura de grupo. Se $\sigma, \tau \in S_p$, escribimos $\sigma\tau := \sigma \circ \tau$, e compróbase facilmente que tamén é unha permutación. A operación é asociativa, ten elemento neutro (a identidade ι) e todo elemento ten inverso.

Proposición 2.5. O grupo S_p ten $p!$ elementos.

Demostración. Sexa $\sigma \in S_p$. O elemento $\sigma(1)$ pode ser calquera dos elementos do conxunto $\{1, \ldots, p\}$. Unha vez escollido este, $\sigma(2)$ pode ser calquera dos $p-1$ elementos restantes; de xeito similar, $\sigma(3)$ non pode tomar nin o valor de $\sigma(1)$ nin o de $\sigma(2)$; e así ata chegar a $\sigma(p)$, para o cal só hai unha posibilidade. $\square$

Os elementos máis importantes do grupo simétrico son os ciclos e as transposicións.

Definición 2.4. Dados $a_1, a_2, \ldots, a_\ell$ elementos diferentes de $\{1, 2, \ldots, p\}$, denotamos como $\gamma = (a_1, a_2, \ldots, a_r) \in S_p$ á permutación definida por

$$\gamma(a_1) = a_2,\ \gamma(a_2) = a_3, \ldots,\ \gamma(a_{r-1}) = a_r,\ \gamma(a_r) = a_1,\ \gamma(a) = a \text{ se } a \neq a_i.$$

Un elemento así chámase *ciclo* de lonxitude r ou r-ciclo. A súa lonxitude denótase como $\ell(\gamma)$.

Proposición 2.6. Calquera permutación pódese escribir como produto de ciclos disxuntos de maneira única (salvo a orde dos ciclos).

Demostración. Sexa $\sigma \in S_p$, e sexa $a_1 \in \{1, \ldots, p\}$ un elemento calquera. Consideremos a sucesión dada por $a_{n+1} = \sigma(a_n)$ para todo $n \geq 1$. Como os a_i son elementos dun conxunto finito, chegará un momento no que se repitan. Sexa $a_{r+1} = \sigma(a_r)$ o primeiro que xa saíse antes. Entón ten que ser necesariamente $a_{r+1} = a_1$, xa que se $a_{r+1} = a_k$, con $2 \leq k \leq r$, entón $\sigma(a_r) = a_{r+1} = a_k = \sigma(a_{k-1})$; agora ben, como σ é inxectiva, teriamos que $a_r = a_{k-1}$, que sabemos que non é certo. Escribiremos $\gamma_1 = (a_1, a_2, \ldots, a_r)$, que cumpre $\gamma_1(a_i) = \sigma(a_i)$ para todo i. Se os a_i conteñen todos os elementos de $\{1, \ldots, p\}$, entón xa acabamos. En caso contrario, consideramos $b_1 \in A_n$ diferente dos a_i e procedemos como antes ata chegar a un ciclo $\gamma_2 = (b_1, b_2, \ldots, b_s)$. Os ciclos son disxuntos, xa que se $b_j = a_i$, entón $b_1 = \gamma_2^{1-j}(b_j) = \sigma^{1-j}(b_j) = \sigma^{1-j}(a_i) = a_{1+i-j \pmod r}$, que é unha contradición. Iterando o proceso chegará un momento no que os ciclos conterán todos os elementos de $\{1, \ldots, p\}$ e teremos unha expresión $\sigma = \gamma_1\gamma_2\cdots\gamma_k$ como produto de ciclos disxuntos.
Para ver a unicidade, poñamos $\sigma = \gamma_1'\gamma_2'\cdots\gamma_m'$, con $\gamma_1' = (u_1, u_2, \ldots, u_t)$. Podemos supoñer sen perda de xeneralidade que γ_1 contén u_1 e que $\gamma_1 = (a_1, a_2, \ldots, a_r)$, con $a_1 = u_1$. Como $\gamma_1(a_1) = \sigma(a_1) = \sigma(u_1) = \gamma_1'(u_1)$, vemos que os dous ciclos son o mesmo. Iterando o proceso, concluímos que a descomposición é única. $\square$

Exemplo. Sexa

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 3 & 6 & 9 & 8 & 2 & 5 & 1 & 4 \end{pmatrix} \in S_9.$$

A súa descomposición en ciclos dixuntos vén dada por $\sigma = (1, 7, 5, 8)(2, 3, 6)(4, 9)$, aínda que tamén se podería escribir $\sigma = (3, 6, 2)(5, 8, 1, 7)(9, 4)$.

Definición 2.5. Unha *transposición* é un elemento de S_p que intercambia dous números e deixa fixos todos os demais.

Proposición 2.7. Toda permutación descompón como produto de transposicións.

Demostración. Como toda permutación descompón como produto de ciclos, é suficiente con ver que un ciclo $\gamma = (a_1, a_2, \ldots, a_r)$ é un produto de transposicións. Iso pódese conseguir pondo, por exemplo,

$$\gamma = (a_1, a_2)(a_2, a_3) \cdots (a_{r-1}, a_r).$$

$\square$

A seguinte proposición establece que, aínda que a descomposición en produto de transposicións non ten por que ser única, si o é a paridade do número de transposicións.

Proposición 2.8. Sexan $\sigma = \sigma_1\sigma_2 \cdots \sigma_r = \tau_1\tau_2 \cdots \tau_s$ dúas descomposicións dunha permutación en produto de transposicións. Entón r e s teñen a mesma paridade.

Demostración. Dada unha permutación σ, escribiremos $N(\sigma)$ para o número de ciclos disxuntos nos que descompón, que é un enteiro ben determinado (isto inclúe os ciclos de lonxitude un). Para demostrar o resultado, facemos a seguinte observación: se τ é unha transposición, entón $N(\tau\sigma) = N(\sigma) \pm 1$. Para iso, pomos $\tau = (a, b)$ e cómpre distinguir dúas opcións, segundo a e b estean no mesmo ciclo ou non. Se están no mesmo ciclo, entón

$$\tau\sigma = (a, b)(a, \ldots, x, b, \ldots, y)\gamma_2 \cdots \gamma_k = (a, \ldots, x)(b, \ldots, y)\gamma_2 \cdots \gamma_k,$$

e o número de ciclos aumenta en un; en cambio, se non están no mesmo ciclo,

$$\tau\sigma = (a, b)(a, \ldots, x)(b, \ldots, y)\gamma_3 \cdots \gamma_k = (a, \ldots, x, b, \ldots, y)\gamma_3 \cdots \gamma_k,$$

e o número redúcese en un.
Con esta observación, podemos probar a proposición. Como σ_r é unha transposición, $N(\sigma_r) = p - 1$. Polo tanto,

$$N(\sigma_1 \cdots \sigma_r) \equiv p - r \pmod{2},$$

o cal mostra que $p-r$ e $p-s$ teñen a mesma paridade, e iso é suficiente para concluír. $\square$

Exemplo. Consideremos a permutación $\sigma = (1, 7, 5, 8)(2, 3, 6)(4, 9)$ do exemplo anterior. Entón,

$$(1, 5)\sigma = (1, 7)(5, 8)(2, 3, 6)(4, 9);$$

en cambio, cando consideramos dous elementos en ciclos diferentes,

$$(1, 2)\sigma = (1, 7, 5, 8, 2, 3, 6)(4, 9).$$

Definición 2.6. Unha permutación dise *par* ou *impar* segundo a paridade do número de transposicións nas que descompón. O signo dunha permutación σ é $+1$ se σ é par e -1 se σ é impar. Denotarase como $\operatorname{sgn}(\sigma)$.

Outro concepto de interese para o estudo das permutacións é a *orde*.

Definición 2.7. Sexa σ unha permutación. A *orde* de σ é o menor enteiro positivo r tal que σ^r é a identidade.

Exemplo. O ciclo $(1, 2)$ ten orde 2 e o ciclo $(1, 2, 3)$ ten orde 3. Máis en xeral, o ciclo $(a_1, \ldots, a_r)$ ten orde r.
Se $\sigma = \gamma_1 \cdots \gamma_t$ é unha descomposición de σ en produto de ciclos disxuntos de lonxitudes $\ell_1, \ldots, \ell_t$, entón a orde de σ é o mínimo común múltiplo dos ℓ_i. No exemplo anterior, no que $\sigma = (1, 7, 5, 8)(2, 3, 6)(7, 8)$, temos que a orde é $\operatorname{lcm}(4, 3, 2) = 12$.

2.3. Aplicacións multilineais antisimétricas

Despois da discusión sobre o grupo simétrico, podemos volver á discusión sobre aplicacións multilineais. Dado un elemento $\sigma \in S_p$ e $f \in T_p(E)$ escribiremos $\sigma \cdot f$ para referirnos ao elemento caracterizado por

$$(\sigma \cdot f)(x_1, \dots, x_p) = f(x_{\sigma_1}, \dots, x_{\sigma(p)}).$$

Definición 2.8. Sexa $f \in T_p(E)$. Dise que f é *simétrico* se, para toda permutación $\sigma \in S_p$ se ten que $\sigma \cdot f = f$, é dicir,

$$f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = f(x_1, \dots, x_p).$$

Do mesmo xeito, dise que f é *antisimétrico* se, para toda permutación $\sigma \in S_p$ se ten que $\sigma \cdot f = \operatorname{sgn}(\sigma) f$, é dicir,

$$f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = \operatorname{sgn}(\sigma) f(x_1, \dots, x_p).$$

Chamamos $S_p(E) \subset T_p(E)$ ao conxunto dos tensores p-covariantes simétricos e $A_p(E) \subset T_p(E)$ ao conxunto dos tensores p-covariantes antisimétricos.

Proposición 2.9. Supoñamos que K é un corpo de característica diferente de 2 e que E é un K-espazo vectorial. Se $f \in A_p(E)$ e $v_1, \dots, v_p \in E$, cúmprense as seguintes propiedades:

(a) $f(v_1, \dots, v_i, \dots, v_j, \dots, v_p) = -f(v_1, \dots, v_j, \dots, v_i, \dots, v_p)$.

(b) Se un vector $v_i = 0$, entón $f(v_1, \dots, v_i, \dots, v_p) = 0$.

(c) Se $v_i = v_j$, entón $f(v_1, \dots, v_i, \dots, v_j, \dots, v_p) = 0$.

(d) Se $v_j = \sum_{i \neq j} a_i v_i$, entón $f(v_1, \dots, v_j, \dots, v_p) = 0$.

Demostración. A primeira propiedade é consecuencia de aplicar a definición de forma alternada cunha transposición. A segunda séguese da linealidade de f, mentres que a terceira é consecuencia directa de aplicar a propiedade (a). Finalmente, (d) demóstrase combinando a linealidade de f coa propiedade (c). □

Cando un tensor f cumpre a propiedade (c), dise que é *alternado*. Como xa vimos, se a característica do corpo é distinta de 2, ambos conceptos son equivalentes; polo tanto, en diferentes contextos falaremos de tensores antisimétricos e alternados como conceptos equivalentes. Ademais, alternado implica antisimétrico, polo que é habitual traballar con esa condición.
Se E ten dimensión n, imos estudar as aplicacións multilineais alternadas de $T_n(E)$, isto é, as aplicacións multilineais $E \times \overset{n}{\cdots} \times E \to K$ que cumpren a condición sobre o signo presentada na definición anterior.

Proposición 2.10. O espazo $A_n(E)$ ten dimensión 1, isto é, dada unha aplicación $D \in A_n(E)$, esta queda determinada polos valores que toma sobre unha base ordenada $(v_1, \dots, v_n)$ de E.

Demostración. Imos comezar vendo que unha aplicación $D \in A_n(E)$ queda determinada polos seus valores sobre unha base ordenada. Para iso, consideremos vectores $(u_1, \dots, u_n)$, con $u_i = \sum_{j=1}^n a_{ij} v_j$. Entón,

$$\begin{aligned} D(u_1, \dots, u_n) &= D\left(\sum_{j_1=1}^n a_{1j_1} v_{j_1}, \dots, \sum_{j_n=1}^n a_{nj_n} v_{j_n}\right) \\ &= \sum_{j_1,\dots,j_n=1}^n D(a_{1j_1} v_{j_1}, \dots, a_{nj_n} v_{j_n}) \\ &= \sum_{j_1,\dots,j_n=1}^n a_{1j_1} \cdots a_{nj_n} D(v_{j_1}, \dots, v_{j_n}). \end{aligned}$$

Aínda que os subíndices $j_1, \dots, j_n$ poden tomar calquera valor en $\{1, \dots, n\}$, o sumando sempre se anulará se dous deles son iguais. Polo tanto, podemos sumar sobre as permutacións de S_n, e obtemos a igualdade

$$D(u_1, \dots, u_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} D(v_1, \dots, v_n).$$

Polo tanto, de existir algunha aplicación $D \in A_n(E)$ con $D(v_1, \dots, v_n) = k$, para $k \in K$, a única posibilidade sería

$$D(u_1, \dots, u_n) = k \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Chamámoslle D_k a esta aplicación. Polo de agora, estableceuse que hai, como moito, unha aplicación multilineal antisimétrica que toma o valor k sobre a base. Queda por comprobar que esta aplicación cumpre tódalas propiedades que se esixen.
Imos empezar vendo que $D_k \in A_n(E)$, o que é unha comprobación inmediata vendo que respecta a suma en cada coordenada, o produto por escalares e que aplicar unha permutación sigma multiplica o valor por $\operatorname{sgn}(\sigma)$. Comezaremos comprobando que respecta a suma, considerando que na coordenada i-ésima pomos $u_i + u_i'$, onde $u_i' = \sum_{j=1}^n b_{ij} v_j$:

$$D(u_1, \dots, u_i + u_i', \dots, u_n) = k \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots (a_{i\sigma(i)} + b_{i\sigma(i)}) \cdots a_{n\sigma(n)},$$

que é precisamente $D(u_1, \dots, u_i, \dots, u_n) + D(u_1, \dots, u_i', \dots, u_n)$. A comprobación para o produto por escalares é a mesma. Finalmente, temos que

$$\begin{aligned} (\tau D)(u_1, \dots, u_n) &= D(u_{\tau(1)}, \dots, u_{\tau(n)}) \\ &= k \cdot \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\tau(1)\sigma(1)} \cdots a_{\tau(n)\sigma(n)} \\ &= k \cdot \operatorname{sgn}(\tau) \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma\tau^{-1}) a_{1\sigma\tau^{-1}(1)} \cdots a_{n\sigma\tau^{-1}(n)} \\ &= \operatorname{sgn}(\tau) D(u_1, \dots, u_n). \end{aligned}$$

Polo tanto, probamos que

$$K \longrightarrow A_n(E), \quad k \mapsto D_k$$

é un isomorfismo de espazos vectoriais: non pode haber dúas aplicacións en $A_n(E)$ para as cales a avaliación en $(v_1, \dots, v_n)$ dea k (é dicir, a aplicación é inxectiva), pero si construímos explicitamente unha que funciona, a que chamamos D_k (a aplicación é entón sobrexectiva). Polo tanto $A_n(E)$ ten dimensión 1, como queriamos ver. $\square$

Ao valor $D(u_1, \dots, u_n)$ construído na proba do resultado anterior chamarémoslle valor do determinante dos vectores $u_1, \dots, u_n$ na base $\mathcal{B}_v = \{v_1, \dots, v_n\}$. Da demostración, séguese que o resultado de aplicar unha aplicación multilineal non nula $D \in A_n(E)$ sobre n vectores é distinto de 0 se e soamente se estes forman unha base.

2.4. Determinantes

Nesta sección, imos definir primeiro o determinante de n vectores, e logo estendemos de xeito natural ese concepto ao caso dunha matriz ou dun endomorfismo.

Definición 2.9. Sexa E un K-espazo vectorial de dimensión n e base $\mathcal{B}_v = \{v_1, \dots, v_n\}$. O *determinante* é a única aplicación n-multilineal alternada que vale 1 ao avaliala en $(v_1, \dots, v_n)$.
De xeito equivalente, se $u_1, \dots, u_n$ son n vectores de E, con $u_i = \sum_{j=1}^n a_{ij} v_j$, o seu determinante na base $\mathcal{B}_v$ é

$$\det_{(v_i)}(u_1, \dots, u_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Observamos tamén que a aplicación

$$A_n(E) \longrightarrow K, \quad D \mapsto D(u_1, \dots, u_n)$$

é lineal e bixectiva, dando lugar a un isomorfismo de espazos vectoriais de dimensión 1. Séguese, do feito de que $D \in A_n(E)$ estea determinada polos valores que toma sobre unha base ordenada, que o determinante de n vectores é 0 se e soamente se estes son linealmente dependentes. O determinante tamén cumpre o resto de propiedades que estudamos para as aplicacións multilineais alternadas.
Ao longo desta sección, sexa $A = (a_{ij}) \in \mathcal{M}_n(K)$ unha matriz $n \times n$ con entradas nun corpo K.

Definición 2.10. O *determinante* de A, $\det(A)$, defínese como

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Sexa $E = K^n$; escribindo $a_i = \sum_{j=1}^n a_{ji} v_j$, temos que $\det A = \det_{(v_i)}(a_1, \dots, a_n)$, isto é, o determinante dunha matriz coincide co determinante dos n vectores columna con respecto á base canónica. Observamos que sería posible ter introducido directamente o determinante dunha matriz deste xeito, sen a discusión previa sobre aplicacións alternadas, pero pensamos que este acercamento permite entender mellor o porqué da definición.
Imos definir por último o determinante dun endomorfismo $f\colon E \to E$ de xeito que non dependa da elección dunha base. Para iso, recordamos que a aplicación lineal f induce unha aplicación dual f^* de xeito que, dada $\omega\colon E \to K$, pódeselle asociar $f^*(\omega)\colon E \to K$ definida por $f^*(\omega) = \omega \circ f$.
De xeito análogo, para todo tensor n-covariante alternado D, a aplicación

$$f^*(D)\colon E \times \overset{n}{\cdots} \times E \longrightarrow K \quad (u_1, \dots, u_n) \mapsto D(f(u_1), \dots, f(u_n))$$

é un tensor n-covariante alternado. Temos polo tanto unha aplicación

$$f^*\colon A_n(E) \longrightarrow A_n(E) \quad D \mapsto f^*(D)$$

que é lineal, e que de feito para $n = 1$ coincide coa aplicación dual. Como $A_n(E)$ é un espazo vectorial de dimensión 1, f^* é un múltiplo da identidade, isto é $f^* = d \cdot \mathbb{I}_{A_n(E)}$.

Definición 2.11. O *determinante* do endomorfismo f é a razón d que fai que a igualdade $f^* = d \cdot \mathbb{I}_{A_n(E)}$ sexa certa, isto é, $f^* = (\det f) \cdot \mathrm{Id}_{A_n(E)}$.

Fixemos $\mathcal{B} = \{v_1, \ldots, v_n\}$ unha base de E e un tensor n-covariante alternado $D \neq 0$. Temos que $f^* = (\det f)\mathbb{I}_n$, polo que $f^*(D) = (\det f)D$, e polo tanto

$$f^*(D)(v_1, \ldots, v_n) = D(f(v_1), \ldots, f(v_n)) = \det_{(v_i)}(f(v_1) \ldots, f(v_n))D(v_1, \ldots, v_n).$$

Observamos que no último paso empregamos a mesma idea que ao demostrar que $D(u_1, \ldots, u_n)$ determina unha aplicación multilineal alternada. De aquí conclúese que $\det f = \det_{(v_i)}(f(v_1), \ldots, f(v_n))$, e como este cálculo funciona independentemente da base escollida, deducimos que o determinante do endomorfismo coincide co determinante da matriz en calquera base.

2.5. Propiedades dos determinantes

Hai algunhas propiedades habituais dos determinantes que son evidentes pola súa definición como forma bilineal alternada.

(a) O determinante obtido ao multiplicar unha columna dunha matriz por un escalar $\lambda \in K$ é o determinante inicial multiplicado por λ.

(b) O determinante non cambia cando a unha columna lle sumamos un múltiplo doutra columna.

(c) Ao permutar dúas columnas o determinante cambia de signo.

Imos comezar probando unha proposición doada, pero que non resulta completamente inmediata a partir da construción.

Proposición 2.11. Sexa A^t a matriz trasposta de A. Entón, $\det A = \det A^t$.

Demostración. O resultado é consecuencia da seguinte cadea de igualdades:

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \cdots a_{\sigma^{-1}(n)n} \\
&= \sum_{\tau \in S_n} \mathrm{sgn}(\tau) a_{\tau(1)1} a_{\tau(2)2} \cdots a_{\tau(n)n} \\
&= \det A^t
\end{aligned},$$

onde a segunda igualdade séguese de reordenar o produto en orde crecente de columna e a terceira é consecuencia de $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$. □

A definición que demos non é demasiado eficaz para o cálculo de determinantes, e é máis habitual usar a chamada regra de Laplace. Para presentala, sexa $A^{(r,s)} \in \mathcal{M}_{n-1}(K)$ a matriz obtida eliminando a fila r-ésima e a columna s-ésima. Sexa tamén $a^*_{r,s} = (-1)^{r+s} \det(A^{(r,s)})$.

Proposición 2.12. Tense que

$$\sum_{k=1}^{n} a_{ik} a^*_{jk} = \begin{cases} \det A & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases}$$

Demostración. Imos comezar co caso $i = j$. Sexa $\sigma_t \in S_n$ o ciclo dado por $(t, t+1, \ldots, n)$, onde $1 \leq t \leq n$ (en concreto, $\sigma_t(t) = t+1$, $\sigma_t(t+1) = t+2$ e así ata $\sigma_t(n) = t$). Entón, $\operatorname{sgn}(\sigma_t) = (-1)^{n-t}$. Esta notación permítenos escribir $A^{(r,s)} = (a_{ij}^{(r,s)})$, onde $a_{ij}^{(r,s)} = a_{\sigma_r(i)\sigma_s(j)}$.
Facemos a seguinte observación. Identificando $S_{n-1} \simeq \{\sigma \in S_n \mid \sigma(n) = n\}$, hai unha bixección

$$\{\sigma \in S_n \mid \sigma(i) = k\} \xrightarrow{\sim} S_{n-1}, \quad \sigma \mapsto \tau = \sigma_k^{-1}\sigma\sigma_i.$$

A aplicación está ben definida porque $\tau(n) = \sigma_k^{-1}(\sigma(\sigma_i(n))) = \sigma_k^{-1}(\sigma(i)) = \sigma_k^{-1}(k) = n$; por outro lado, $\sigma = \sigma_k\tau\sigma_i^{-1}$ é a inversa da aplicación, co cal é efectivamente unha bixección.
Observamos agora que

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\
&= \sum_{k=1}^{n} a_{ik} \sum_{\substack{\sigma \in S_n \\ \sigma(i)=k}} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \ldots a_{i-1\sigma(i-1)} a_{i+1\sigma(i+1)} \cdots a_{n\sigma(n)} \\
&= \sum_{k=1}^{n} a_{ik} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\sigma_k\tau\sigma_i^{-1}) a_{1\sigma_k\tau\sigma_i^{-1}(1)} \cdots a_{n\sigma_k\tau\sigma_i^{-1}(n)} \\
&= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) a_{\sigma_i(1)\sigma_k\tau(1)} \cdots a_{\sigma_i(n-1)\sigma_k\tau(n-1)} \\
&= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \sum_{\tau \in S_{n-1}} \operatorname{sgn}(\tau) a_{1\tau(1)}^{(i,k)} \cdots a_{n-1\tau(n-1)}^{(i,k)} \\
&= \sum_{k=1}^{n} a_{ik} a_{ik}^{*}.
\end{aligned}$$

Na terceira igualdade empregouse a substitución $\sigma = \sigma_k\tau\sigma_i^{-1}$; na cuarta, que o signo é multiplicativo e que, polo tanto, $\operatorname{sgn}(\sigma_k\tau\sigma_i^{-1}) = (-1)^{i+k}\operatorname{sgn}(\tau)$; por último, a quinta é consecuencia directa da definición das matrices $A^{(i,k)}$.
Por outra banda, se $i \neq j$, entón sexa $B = (b_{ij})$ a matriz que é igual a A pero reemprazando a fila j-ésima pola i-ésima. Entón, polo apartado anterior,

$$0 = \det B = \sum_{k=1}^{n} b_{jk} b_{jk}^{*} = \sum_{k=1}^{n} a_{ik} a_{jk}^{*},$$

onde usamos que sobre as entradas que non están na fila j-ésima as matrices A e B coinciden. □

Exemplo. Imos considerar un exemplo da regra de Laplace para o cálculo dun determinante:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3 + 12 - 9 = 0;$$

en particular, poderiamos ter observado que o determinante era 0 sen facer ningún cálculo xa que a suma da primeira e da terceira columna coincide co dobre da segunda columna (é dicir, hai unha relación de dependencia lineal entre as columnas).

Proposición 2.13. Se f e g son dous endomorfismos de E e $\mathbb{I}_E$ é o endomorfismo identidade, cúmprese que $\det(g \circ f) = \det f \cdot \det g$ e $\det \mathbb{I}_E = 1$. Do mesmo xeito, dadas $A, B \in \mathcal{M}_n(K)$, con $\mathbb{I}_n$ a matriz identidade, temos que $\det AB = \det A \cdot \det B$ e $\det \mathbb{I}_n = 1$.

Demostración. Sexa $D \in A_n(E)$ e sexan $w_1, \dots, w_n$ vectores arbitrarios de E. Entón

$$\begin{aligned}(g \circ f)^*(D)(w_1, \dots, w_n) &= D(g(f(w_1)), \dots, g(f(w_n))) \\ &= g^*(D)(f(w_1), \dots, f(w_n)) \\ &= f^*(g^*(D))(w_1, \dots, w_n) = (f^* \circ g^*)(D)(w_1, \dots, w_n).\end{aligned}$$

De aquí concluímos que $(g \circ f)^* = f^* \circ g^*$, xa que $(g \circ f)^* = \det(g \circ f) \cdot \mathrm{Id}_{A_n(E)}$ e $f^* \circ g^* = (\det f) \cdot (\det g) \cdot \mathrm{Id}_{A_n(E)}$.
No caso do endomorfismo identidade, temos que calcular o valor de $\mathbb{I}_E^*(D)(w_1, \dots, w_n)$, pero é claro que coincide con $D(w_1, \dots, w_n) = \mathbb{I}_{A_n(E)}(D)(w_1, \dots, w_n)$. Polo tanto, $\mathrm{Id}_E^* = \mathrm{Id}_{A_n(E)}$, e usando a definición de determinante temos o resultado do enunciado.
Unha vez temos o resultado demostrado para endomorfismos, é inmediato que tamén se cumpre para matrices, dado que o determinante dun endomorfismo coincide co determinante da matriz independentemente da base. □

A modo de notación para capítulos seguintes, necesitamos facer a seguinte definición.

Definición 2.12. O conxunto das matrices $n \times n$ con determinante diferente de 0 e entradas en K denótase como $\mathrm{GL}_n(K)$ e chámase *grupo lineal* (de orde n). Trátase dun grupo coa operación dada polo produto de matrices. Cando $n = 1$, podemos pór $\mathrm{GL}_1(K) = K^\times$.

2.6. Rangos de matrices e sistemas de ecuacións

Sexa A unha matriz de tamaño $m \times n$ e k un enteiro que cumpre $0 \leq k \leq \min\{m, n\}$. Un menor de orde k de A é o determinante dunha matriz $k \times k$ obtida eliminando $m - k$ filas e $n - k$ columnas de A. Recordamos que o *rango* de A é o maior número de filas (ou columnas) que son linealmente independentes. Imos enunciar e probar o seguinte resultado unicamente no caso de matrices cadradas, aínda que admite unha extensión natural a matrices non cadradas.

Proposición 2.14. O rango dunha matriz cadrada $A \in \mathcal{M}_n(K)$ é o máximo das ordes dos menores non nulos de A.

Demostración. Sexa r o rango da matriz, e sexan $j_1, \dots, j_p$ as columnas coas que se formou o menor, que identificamos con vectores $a_{j_1}, \dots, a_{j_p}$. Se $p > r$, os vectores son linealmente dependentes e polo tanto un deles é combinación lineal do resto; isto implica que o determinante correspondente é cero.
Imos probar agora que hai un menor de orde r con determinante non nulo. Para iso, consideramos r columnas que sexan linealmente independentes, $a_{j_1}, \dots, a_{j_r}$, que existen xa que o rango é r. Observamos agora que é posible escoller $n - r$ vectores da base canónica, $e_{j_{r+1}}, \dots, e_{j_n}$, de maneira que $\{a_{j_1}, \dots, a_{j_r}, e_{j_{r+1}}, \dots, e_{j_n}\}$ formen unha base (isto é consecuencia do feito que r vectores linealmente independentes só nos poden permitir escribir como combinación lineal deles un máximo de r vectores dunha base). Polo tanto, o determinante dos vectores $\{a_{i_1}, \dots, a_{i_r}, e_{i_{r+1}}, \dots, e_{i_n}\}$ é distinto de cero. Ademais, en cada unha das $n - r$ columnas do final todos os elementos son cero, salvo

un deles que vale 1. Se aplicamos a regra de Laplace ao cálculo deste determinante temos que coincide, salvo signo, co determinante $r \times r$ obtido ao eliminar das primeiras r columnas as filas nas que algunha das $n - r$ columnas finais tiña un 1. □

Imos agora enunciar e demostrar a regra de Cramer, que permite atopar as solucións duns sistema de ecuacións.

Proposición 2.15. Todo sistema lineal $Ax = b$ de n ecuacións e n incógnitas, con $A \in \mathcal{M}_n(K)$ invertible, é compatible determinado. As compoñentes x_i da súa solución $(x_1, \ldots, x_n) \in K^n$ están dadas pola fórmula

$$x_i = \frac{\det A_i}{\det A},$$

onde A_i é a matriz A coa columna i-ésima cambiada polo vector b.

Demostración. O feito de que o sistema sexa compatible determinado vén do feito de que a igualdade $Ax = b$ é equivalente a $x = A^{-1}b$, co cal a única solución é precisamente $A^{-1}b$. Se $x = (x_1, \ldots, x_n)$ e $a_1, \ldots, a_n$ son as columnas de A, temos que o vector b se escribe como $x_1 a_1 + \ldots + x_n a_n = b$. Polo tanto,

$$\begin{aligned}
\det(A_i) &= \det(a_1, \ldots, a_{i-1}, \sum_{j=1}^{n} x_j a_j, a_{i+1}, \ldots, a_n) \\
&= \sum_{j=1}^{n} x_j \det(a_1, \ldots, a_{i-1}, a_j, a_{i+1}, \ldots, a_n) \\
&= x_i \det(A),
\end{aligned}$$

onde usamos a definición de A_i e as propiedades dos determinantes. □

Observación. Desde un punto de vista computacional, o cálculo dos $n+1$ determinantes que se precisan para resolver un sistema por Cramer é moito máis custoso que aplicar o método de Gauss.

Polo xeral, un dos nosos obxectos de interese van ser os sistemas lineais homoxéneos, é dicir, nos que $b = 0$. Porén, especialmente no estudo da xeometría afín, ten sentido considerar o problema xeral no que queremos resolver o sistema matricial $Ax = b$, onde $A \in \mathcal{M}_{m \times n}$, $x \in \mathcal{M}_{n \times 1}$ e $b \in \mathcal{M}_{m \times 1}$. En termos de operadores lineais, podemos considerar que A é a matriz dunha aplicación lineal

$$f \colon K^n \longrightarrow K^m, \quad x \mapsto f(x).$$

Entón, a condición necesaria é suficiente para que o sistema teña algunha solución é que $b \in \operatorname{im}(f)$, isto é, que para unha base $\{v_1, \ldots, v_m\}$ se cumpra

$$\langle f(v_1), \ldots, f(v_n), b \rangle = \langle f(v_1), \ldots, f(v_n) \rangle.$$

Se sabemos que o sistema ten algunha solución, é dicir, $b \in \operatorname{im}(f)$, podemos considerar a aplicación $\tilde{f}$ do primeiro teorema de isomorfismo:

$$\tilde{f} \colon K^m / \ker(f) \longrightarrow \operatorname{im}(f), \quad [x] \mapsto [f(x)].$$

Se $b \in \operatorname{im}(f)$ e $f(x_0) = b$, calquera outra solución de $f(x) = b$ virá dada por $x = x_0 + y$, onde $y \in \ker(f)$, isto é, o conxunto de solucións é $x_0 + \ker(f)$.

2.7. Problemas

Grupo simétrico.

Problema 2.1.

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 8 & 9 & 4 & 5 & 2 & 1 & 6 \end{pmatrix} \in S_9,$$
$$\sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 1 & 6 & 8 & 2 & 7 & 5 \end{pmatrix} \in S_8$$
$$\sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 8 & 10 & 5 & 2 & 4 & 9 & 3 & 1 & 6 & 7 \end{pmatrix} \in S_{10}.$$

Calcular a descomposición en ciclos disxuntos, unha descomposición en produtos de transposicións, a orde e o signo. Calculade tamén σ_1^{1000}, σ_2^{250} e σ_3^{611}.

Solución. Comezamos recordando que a orde dunha transposición σ é o menor enteiro positivo n tal que $\sigma^n = \iota$, onde ι é a identidade. A continuación, para cada unha das permutacións amosamos as descomposicións en ciclos disxuntos e en produto de transposicións, así como a orde e o signo.

- $\sigma_1 = (1,3,8)(2,7)(4,9,6,5) = (1,3)(3,8)(2,7)(4,9)(9,6)(6,5)$. Orde 12. Signo $+1$.
- $\sigma_2 = (1,3)(2,4,6)(5,8)(7) = (1,3)(2,4)(4,6)(5,8)$. Orde 6. Signo $+1$.
- $\sigma_3 = (1,8)(2,10,7,3,5,4)(6,9) = (1,8)(2,10)(10,7)(7,3)(3,5)(5,4)(6,9)$. Orde 6. Signo -1.

Finalmente, calculamos as potencias que nos piden.

- Como $\sigma_1^{12} = \mathrm{Id}$ e $1000 = 83 \cdot 12 + 4$, temos que
$$\sigma_1^{1000} = \sigma_1^{83\cdot 12} \cdot \sigma_1^4 = \sigma_1^4 = (1,3,8)^4(2,7)^4(4,9,6,5)^4 = (1,3,8).$$
- Como $\sigma_2^6 = \mathrm{Id}$ e $250 = 41 \cdot 6 + 4$, temos que
$$\sigma_2^{250} = \sigma_2^{41\cdot 6} \cdot \sigma_2^4 = \sigma_2^4 = (1,3)^4(2,4,6)^4(5,8)^4 = (2,4,6).$$
- Como $\sigma_3^6 = \mathrm{Id}$ e $611 = 101 \cdot 6 + 5$, temos que
$$\sigma_3^{611} = \sigma_3^{101\cdot 6} \cdot \sigma_3^5 = \sigma_3^5 = (1,8)(2,4,5,3,7,10)(6,9).$$

Problema 2.2. Demostrar as seguintes propiedades do signo.

(a) Se ι é a identidade, entón $\operatorname{sgn}(\iota) = 1$.

(b) Se τ é unha transposición, $\operatorname{sgn}(\tau) = -1$.

(c) Se $\gamma = (a_1, \ldots, a_m)$ é un ciclo de lonxitude m, $\operatorname{sgn}(\gamma) = (-1)^{m-1}$

(d) $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$.

(e) $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$.

(f) Fixada unha transposición ε, toda permutación impar pódese escribir como o produto dunha permutación par por ε.

Solución. (a) A identidade descompón como produto de 0 transposicións, que é un número par, así que o signo é $+1$.

(b) Unha transposición é o produto dunha transposición, que é un número impar, así que o signo é -1.

(c) Observamos que $\gamma = (a_1, a_2)(a_2, a_3)\cdots(a_{m-1}, a_m)$, co cal γ é o produto de $m-1$ transposicións.

(d) Se $\sigma = \gamma_1 \cdots \gamma_r$ e $\tau = \delta_1 \cdots \delta_s$, sendo γ_i e δ_j transposicións, entón a composición pódese escribir como $\sigma\tau = \gamma_1 \cdots \gamma_r \delta_1 \cdots \delta_s$. Polo tanto, $r + s$ é par se e só se r e s teñen a mesma paridade, ou o que é o mesmo, $\mathrm{sgn}(\sigma\tau)$ é $+1$ se e soamente se $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\tau)$.

(e) Polo apartado anterior, $\mathrm{sgn}(\sigma\sigma^{-1}) = \mathrm{sgn}(\iota) = 1$.

(f) Sexa τ unha permutación impar, e sexa $\sigma = \tau\varepsilon$. Polo apartado (d), σ é par, e cúmprese que $\sigma\varepsilon = \tau\varepsilon^2 = \tau$.

Problema 2.3. Sexa $n \geq 1$ un enteiro. Demostrar que toda permutación impar $\sigma \in S_n$ se pode escribir como produto dunha transposición e de ciclos de lonxitude tres. É única esta descomposición?

Solución. Unha permutación impar é produto dun número impar de transposicións. Polo tanto, temos que

$$\sigma = \tau_1 \tau_2 \cdots \tau_{2k-1} \tau_{2k} \tau_{2k+1},$$

onde as τ_i son transposicións. O produto de dúas transposicións cun elemento en común son ciclos de lonxitude tres, xa que $(a, b)(b, c) = (a, b, c)$. Se son disxuntas, entón

$$(a, b)(c, d) = (a, b)(b, c)(b, c)(c, d) = (a, b, c)(b, c, d).$$

Se aparellamos as τ_i temos un produto de ciclos de lonxitude tres que deixa unha soa, demostrando así o resultado.
A descomposición non é única. Por exemplo, pódese coller

$$(1, 2, 3, 4) = (1, 2)(2, 3, 4) = (1, 2, 3)(3, 4).$$

Aplicacións multilineais.

Problema 2.4. Sexa $E = \mathbb{R}^3$. Probar que $f : E \times E \longrightarrow \mathbb{R}$ a aplicación definida por

$$f(x, y) = 2x_1y_3 - 3x_2y_3 + x_2y_1 - x_3y_2$$

é multilineal (dito doutra maneira, é un tensor 2-covariante). Expresalo na base $\{e_i^* \otimes e_j^*\}_{1 \leq i,j \leq 3}$, onde $\mathcal{B}_e = \{e_1, e_2, e_3\}$ é a base canónica de E e $\mathcal{B}_e^* = \{e_1^*, e_2^*, e_3^*\}$ é a súa base dual. Encontrar a matriz de f na base $\mathcal{B}_e$, definida como a que ten na entrada (i, j) o valor $f(e_i, e_j)$.

Solución. Para ver que f é multilineal, cómpre ver que respecta a suma e o produto por escalares en cada unha das variables.

- Comprobación de $f(x+x',y)=f(x,y)+f(x',y)$:

$$\begin{aligned} f(x+x',y) &= 2(x_1+x_1')y_3-3(x_2+x_2')y_3+(x_2+x_2')y_1-(x_3+x_3')y_2 \\ &= 2x_1y_3-3x_2y_3+x_2y_1-x_3y_2+2x_1'y_3-3x_2'y_3+x_2'y_1-x_3'y_2 \\ &= f(x,y)+f(x',y). \end{aligned}$$

- Comprobación de $f(\lambda x,y)=\lambda f(x,y)$:

$$f(\lambda x,y)=2(\lambda x_1)y_3-3(\lambda x_2)y_3+(\lambda x_2)y_1-(\lambda x_3)y_2=\lambda f(x,y).$$

- A comprobación de $f(x,y+y')=f(x,y)+f(x,y')$ é análoga á primeira.
- A comprobación de $f(x,\lambda y)=\lambda f(x,y)$ é análoga á segunda.

Na base canónica temos que

$$f=2e_1^*\otimes e_3^*-3e_2^*\otimes e_3^*+e_2^*\otimes e_1^*-e_3^*\otimes e_2^*,$$

o cal se comproba vendo que a avaliación en calquera par (e_i,e_j) coincide coa dada no enunciado.
A matriz dunha forma bilineal é a que ten por entrada (i,j) o resultado de facer $f(e_i,e_j)$. Neste caso temos

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

Problema 2.5. Sexan $u=\{u_1,\dots,u_n\}$ e $v=\{v_1,\dots,v_n\}$ bases do espazo vectorial E con $v_j=\sum_i s_j^i u_i$. Se $u^*=(u_1^*,\dots,u_n^*)$ e $v^*=(v_1^*,\dots,v_n^*)$ son as bases duais respectivas e $f\in T_2(E)$ é tal que

$$f=\sum_{i,j}\lambda_{ij}u_i^*\otimes u_j^*=\sum_{i,j}\mu_{ij}v_i^*\otimes v_j^*,$$

entón $\mu_{ij}=\sum_{k,\ell}s_i^k s_j^\ell\lambda_{k\ell}$.

Solución. Se $v_j=\sum_i s_j^i u_i$, temos que $u_j^*=\sum_i s_i^j v_i^*$. Polo tanto,

$$\begin{aligned} f &= \sum_{k,\ell}\lambda_{k\ell}u_k^*\otimes u_\ell^* \\ &= \sum_{k,\ell}\lambda_{k\ell}\Big(\sum_i s_i^k v_i^*\Big)\otimes\Big(\sum_j s_j^\ell v_j^*\Big) \\ &= \sum_{i,j}\Big(\sum_{k,\ell}\lambda_{k\ell}s_i^k s_j^\ell\Big)v_i^*\otimes v_j^*. \end{aligned}$$

Igualando os coeficientes en $v_i^*\otimes v_j^*$ para cada parella (i,j) tense a igualdade do enunciado.

Problema 2.6. Sexa $e=\{e_1,e_2\}$ unha base de $\mathbb{R}^2$ e $e^*=\{e_1^*,e_2^*\}$ a súa base dual. Consideramos a aplicación multilineal $\varphi:\mathbb{R}^2\times\mathbb{R}^2\times\mathbb{R}^2\longrightarrow\mathbb{R}$ definida por

$$\varphi=e_1^*\otimes e_1^*\otimes e_1^*+e_1^*\otimes e_1^*\otimes e_2^*-e_1^*\otimes e_2^*\otimes e_1^*-e_2^*\otimes e_1^*\otimes e_1^*+e_2^*\otimes e_2^*\otimes e_1^*+e_2^*\otimes e_2^*\otimes e_2^*.$$

Encontrar $\varphi((x_1,x_2),(y_1,y_2),(z_1,z_2))$. É φ un tensor simétrico?

Solución. Temos que

$$\begin{aligned}\varphi((x_1,x_2),(y_1,y_2),(z_1,z_2)) &= x_1y_1z_1 + x_1y_1z_2 - x_1y_2z_1\\ &\quad - x_2y_1z_1 + x_2y_2z_1 + x_2y_2z_2.\end{aligned}$$

En particular, φ non é un tensor simétrico. Se o fose, considerando por exemplo a transposición $(2,3)$ teriamos que $(2,3)\varphi = \varphi$, é dicir,

$$\varphi((x_1,x_2),(z_1,z_2),(y_1,y_2)) = \varphi((x_1,x_2),(y_1,y_2),(z_1,z_2)).$$

Isto non é certo; por exemplo, avaliando en $((1,0),(1,0),(0,1))$ temos que φ dá $+1$, mentres que a avaliación en $(2,3)\varphi$ dá -1.

Problema 2.7. Probar que se $\{e_1,e_2,e_3\}$ é unha base do espazo vectorial E, entón o tensor

$$\varphi = e_1^*\otimes e_1^* + 2e_1^*\otimes e_2^* - e_1^*\otimes e_3^* + 2e_2^*\otimes e_1^* + 3e_2^*\otimes e_3^* + 3e_3^*\otimes e_2^* - e_3^*\otimes e_1^* \in T_2(E)$$

é simétrico e o tensor $\eta = e_1^*\otimes e_2^* + e_1^*\otimes e_3^* - e_2^*\otimes e_1^* - e_3^*\otimes e_1^*$ é antisimétrico.

Solución. No caso de φ, temos que comprobar que é invariante pola acción de cada un dos dous elementos de S_2. No caso da identidade é inmediato. Imos analizar agora o caso da transposición $(1,2)$. Comezamos observando que

$$\varphi((y_1,y_2,y_3),(x_1,x_2,x_3)) = x_1y_1 + 2x_1y_2 - x_1y_3 + 2x_2y_1 + 3x_2y_3 + 3x_3y_2 - x_3y_1.$$

Aplicando agora a transposición $(1,2)$, vemos que

$$\begin{aligned}(1,2)\varphi((x_1,x_2,x_3),(y_1,y_2,y_3)) &= \varphi((y_1,y_2,y_3),(x_1,x_2,x_3))\\ &= y_1x_1 + 2y_1x_2 - y_1x_3 + 2y_2x_1 + 3y_2x_3 + 3y_3x_2 - y_3x_1.\end{aligned}$$

En particular, $(1,2)\varphi = \varphi$, como queriamos ver.
No caso de η, procedendo da mesma maneira temos que

$$(1,2)\eta = e_2^*\otimes e_1^* + e_3^*\otimes e_1^* - e_1^*\otimes e_2^* - e_1^*\otimes e_3^* = \eta.$$

É importante ter en conta que se $\psi = u_1^*\otimes\ldots\otimes u_p^*$, non é certo que $\sigma\psi = u_{\sigma(1)}^*\otimes\ldots\otimes u_{\sigma(p)}^*$; o que se cumpre é

$$\sigma\psi = u_{\sigma^{-1}(1)}^*\otimes\ldots\otimes u_{\sigma^{-1}(p)}^*.$$

Como para unha transposición τ se cumpre que $\tau^{-1} = \tau$, este fenómeno non se observa aquí.
Observación. Para demostrar que é unha aplicación bilineal é simétrica (resp. antisimétrica) é suficiente con ver que a matriz asociada a ela é simétrica (resp. antisimétrica).

Determinantes.

Problema 2.8. Calcular os seguintes determinantes:

$$\begin{vmatrix}1&2&3&4&5\\2&3&4&5&6\\3&4&5&6&7\\4&5&6&7&8\\5&6&7&8&9\end{vmatrix}, \qquad \begin{vmatrix}1&2&3&\cdots&n\\-1&0&3&\ldots&n\\-1&-2&0&\ldots&n\\\vdots&\vdots&\vdots&&\vdots\\-1&-2&-3&\ldots&0\end{vmatrix}.$$

Solución. No caso da primeira matriz, podemos restarlle a cada fila a anterior, comezando polo final. Quedan entón catro filas nas que todos os elementos son iguais a 1, polo que o determinante é 0.
No caso da segunda matriz, sumámoslle a primeira fila a cada unha das outras; deste xeito, todos os elementos da diagonal inferior son iguais a cero (é unha matriz triangular superior), e os elementos da diagonal son os números naturais desde 1 ata n; o determinante é por tanto $n!$, sexa cal sexa o valor de n.

Problema 2.9. Encontrar o determinante da matriz $A = (a_{ij}) \in \mathcal{M}_n(K)$ con coeficientes $a_{ij} = |i - j|$.

Solución. Podemos restar de cada fila a seguinte, comezando pola primeira e seguindo ata chegar á penúltima, é dicir,

$$\begin{vmatrix} 0 & 1 & 2 & 3 & \dots & n-1 \\ 1 & 0 & 1 & 2 & \dots & n-2 \\ 2 & 1 & 0 & 1 & \dots & n-3 \\ 3 & 2 & 1 & 0 & \dots & n-4 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ n-1 & n-2 & n-3 & n-4 & \dots & 0 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 1 & 1 & \dots & 1 \\ -1 & -1 & 1 & 1 & \dots & 1 \\ -1 & -1 & -1 & 1 & \dots & 1 \\ -1 & -1 & -1 & -1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ n-1 & n-2 & n-3 & n-4 & \dots & 0 \end{vmatrix}.$$

Amosamos dúas posibles maneiras de concluír. Unha delas é sumarlle a cada columna a última, obtendo que o determinante que se quere calcular é igual a

$$\begin{vmatrix} 0 & 2 & 2 & 2 & \dots & 2 & 1 \\ 0 & 0 & 2 & 2 & \dots & 2 & 1 \\ 0 & 0 & 0 & 2 & \dots & 2 & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ n-1 & n-2 & n-3 & n-4 & \dots & 1 & 0 \end{vmatrix}.$$

Desenvolvendo pola primeira columna (é dicir, aplicando a regra de Laplace), o determinante é igual a

$$(-1)^{n+1}(n-1) \begin{vmatrix} 2 & 2 & 2 & \dots & 2 & 1 \\ 0 & 2 & 2 & \dots & 2 & 1 \\ 0 & 0 & 2 & \dots & 2 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & 1 \end{vmatrix} = (-1)^{n+1}(n-1)2^{n-2}.$$

Alternativamente, súmaselle a primeira columna a todas as demais. Nese caso, obtemos unha matriz triangular inferior que ten diagonal $(-1, -2, \dots, -2, n-1)$, polo que o determinante é $(-1)^{n-1}(n-1)2^{n-2}$.

Problema 2.10. Encontrar o determinante da matriz $A = (a_{ij}) \in \mathcal{M}_n(K)$ con coeficientes

$$a_{ij} = \begin{cases} (-1)^{i-j} & \text{se } i \neq j \\ 2 & \text{se } i = j. \end{cases}$$

Solución. A matriz escríbese como

$$\begin{vmatrix} 2 & -1 & 1 & -1 & \cdots & \mp 1 & \pm 1 \\ -1 & 2 & -1 & 1 & \cdots & \pm 1 & \mp 1 \\ 1 & -1 & 2 & -1 & \cdots & \mp 1 & \pm 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \mp 1 & \pm 1 & \mp 1 & \pm 1 & \cdots & 2 & -1 \\ \pm 1 & \mp 1 & \pm 1 & \mp 1 & \dots & -1 & 2 \end{vmatrix}.$$

Sumando a cada fila a anterior, comezando polo final, tense que o determinante é igual a

$$\begin{vmatrix} 2 & -1 & 1 & -1 & \cdots & \mp 1 & \pm 1 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \end{vmatrix},$$

onde $\pm 1 = (-1)^{n-1}$. Se agora lle restamos a cada columna a seguinte, comezando pola penúltima, quédanos

$$\begin{vmatrix} n+1 & -(n-1) & n-2 & -(n-3) & \cdots & \mp 2 & \pm 1 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{vmatrix} = n+1.$$

Alternativamente, mediante transformacións elementais, podemos chegar a unha matriz tridiagonal, na que na diagonal principal todos os elementos son 2 e nas outras dúas diagonais todos os elementos son 1. Se lle chamamos a_n ao valor do determinante e desenvolvemos pola primeira fila, chegamos de xeito inmediato a que

$$a_n = 2a_{n-1} - a_{n-2},$$

e a partir de aí podemos demostrar por indución que $a_n = n+1$, posto que $a_1 = 2$ e $a_2 = 3$.

Problema 2.11. Sexa K un corpo de característica diferente de 2. Demostrar que toda matriz cadrada antisimétrica (é dicir, que cumpre que $A^t = -A$) de tamaño n impar ten determinante 0.

Solución. Usando a condición do enunciado e as propiedades dos determinantes, temos que

$$(-1)^n \det(A) = \det(-A) = \det(A^t) = \det(A),$$

onde a primeira igualdade é consecuencia da multilinealidade do determinante. Se n é impar, $-\det(A) = \det(A)$, o que implica que $\det(A) = 0$.
É importante observar que o resultado só é certo cando de $2\det(A) = 0$ se pode concluír que $\det(A) = 0$, isto é, cando 2 é diferente de 0; iso non sucede por exemplo en $\mathbb{Z}/2\mathbb{Z}$.

Problema 2.12. Dados $a, b \in \mathbb{R}$, consideramos a matriz $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ de xeito que $a_{ii} = a$ e $a_{ij} = b$ se $i \neq j$. Achar o valor de $\det(A)$ e discutir cal é o rango de A segundo os valores de a e b.

Solución. Restando a última fila á penúltima, a penúltima á antepenúltima, e así ata chegar a restarlle a segunda á primeira, temos que o valor do determinante non cambia e que

$$\begin{vmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & & \vdots \\ b & b & \dots & a \end{vmatrix} = \begin{vmatrix} a & b & \cdots & b \\ b-a & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a-b \end{vmatrix},$$

onde a última matriz ten, salvo na primeira fila, $a-b$ na diagonal principal e $b-a$ na diagonal inmediatamente inferior á principal. Se agora lle sumamos á penúltima columna a última, á antepenúltima a penúltima, e así ata chegar a sumarlle á primeira columna a segunda, temos que

$$\begin{vmatrix} a & b & \cdots & b \\ b-a & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a-b \end{vmatrix} = \begin{vmatrix} a+(n-1)b & (n-1)b & \cdots & b \\ 0 & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a-b \end{vmatrix}.$$

Esta última matriz é diagonal superior, polo que o seu determinante é o produto dos elementos da diagonal, isto é, $(a+(n-1)b)(a-b)^{n-1}$.

- Se $a \neq b$ e $a \neq -(n-1)b$, o rango é n, xa que o determinante é diferente de cero.
- Se $a \neq b$ e $a = -(n-1)b$, o rango é $n-1$, xa que as tres últimas filas son independentes. Isto pódese ver porque o determinante obtido ao eliminar a primeira fila e a primeira columna é $(a-b)^{n-2}(a+(n-2)b)$.
- Se $a = b$ e $a \neq -(n-1)b$, o rango é 1, xa que todas as filas son iguais, pero diferentes de cero.
- Se $a = b$ e $a = -(n-1)b$, tense que $a = b = 0$ e o rango é 0.

Problema 2.13. Sexan $A, B \in \mathcal{M}_3(\mathbb{R})$ de xeito que

$$\det(A) = \det(B) = \det(A+B) = \det(A-B) = 0.$$

Sexan $\lambda, \mu \in \mathbb{R}$. Demostrar que $\det(\lambda A + \mu B) = 0$. É certo o resultado se $A, B \in \mathcal{M}_4(\mathbb{R})$?

Solución. Como A e B son matrices fixadas, podemos ver $\det(\lambda A + \mu B)$ como un polinomio de grao 3 nas variables λ e μ, polo que, en particular,

$$\det(\lambda A + \mu B) = \alpha\lambda^3 + \beta\lambda^2\mu + \gamma\lambda\mu^2 + \delta\mu^3.$$

Cando $(\lambda, \mu) = (1, 0)$, sabemos que $\det(\lambda A + \mu B) = 0$, polo que necesariamente $\alpha = 0$. Analogamente, como $\det(B) = 0$, tense que $\delta = 0$. Pondo agora $(\lambda, \mu) = (1, 1)$, quédanos que

$$0 = \det(A+B) = \beta + \gamma,$$

e pondo $(\lambda, \mu) = (1, -1)$,

$$0 = \det(A-B) = -\beta + \gamma.$$

Sumando as dúas ecuacións, quédanos que $\gamma = 0$ e, polo tanto, $\beta = 0$. Isto demostra que $\det(\lambda A + \mu B) = 0$ para calquera elección de λ e de μ.

En dimensión 4 o resultado non é certo. Sexan

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

Tense que $\det(A) = \det(B) = \det(A+B) = \det(A-B) = 0$, pero a matriz

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

ten determinante -6.

Problema 2.14. Comprobar a identidade seguinte, coñecida como *determinante de Vandermonde*:

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{i>j} (a_i - a_j).$$

Usar este resultado para demostrar que se $t_1, \ldots, t_n$ son elementos dun corpo K, diferentes de cero e todos diferentes entre si, entón, para todo $m \in \mathbb{Z}$, o conxunto de vectores $\{(t_1^k, \ldots, t_n^k) \text{ con } m \leq k \leq m+n-1\}$ é unha base de K^n.

Solución. Imos discutir dúas maneiras de abordar o problema. A primeira é por indución, supondo o resultado certo ata matrices de tamaño $n-1$ (os casos $n=1$ ou $n=2$ son triviais). Agora, a cada columna restámoslle a anterior multiplicada por a_1, indo de dereita a esquerda; deste xeito, quédanos

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & a_2(a_2 - a_1) & \cdots & a_2^{n-2}(a_2 - a_1) \\ 1 & a_3 - a_1 & a_3(a_3 - a_1) & \cdots & a_3^{n-2}(a_3 - a_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_1 & a_n(a_n - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix} = \prod_{i>j} (a_i - a_j).$$

Podemos aplicar a regra de Laplace e desenvolver pola primeira fila, e logo en cada fila usamos a multilinealidade do determinante para sacar fóra o factor $a_i - a_1$. Queda polo tanto

$$(a_2 - a_1)(a_3 - a_1) \cdots (a_n - a_1) \begin{vmatrix} 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-2} \end{vmatrix}.$$

Usando agora a hipótese de indución, temos o resultado desexado.

Outra alternativa para resolver o problema é observar que cando $a_1 = a_i$, para algún $2 \leq i \leq n$, hai dúas filas iguais entón o determinante é 0. Ao desenvolver o determinante usando a definición, temos que é un polinomio de grao $n-1$ en a_n, co cal considerando as outras variables fixadas ten como moito $n-1$ raíces, que se corresponden cos factores

$(a_n-a_1)(a_n-a_2)\cdots(a_n-a_{n-1})$. Ao dividir por ese termo, temos agora un polinomio de grao $n-2$ en a_{n-1}, que nos permite extraer os factores $(a_{n-1}-a_1)(a_{n-1}-a_2)\cdots(a_{n-1}-a_{n-2})$, e así sucesivamente. Polo tanto, concluímos que o determinante é igual a $C\cdot\prod_{i>j}(a_i-a_j)$, onde Q é un polinomio. Agora ben, vistos como polinomios en varias variables, tanto o determinante como $\prod_{i>j}(a_i-a_j)$ teñen grao $\frac{n(n-1)}{2}$, polo que C ten que ser unha constante en K. A constante é un xa que o produto de todas as entradas diagonais do determinante é $a_2a_3^2\cdots a_n^{n-1}$, que é tamén o termo que se obtén ao coller o primeiro sumando de cada factor na expansión do produto.
Observación. Un pode estar tentado de pensar que para polinomios en varias variables se cumpre algunha xeneralización do teorema fundamental da álxebra que afirme que todo polinomio (ou polo menos todo polinomio simétrico) se pode escribir como produto de factores de grao 1. Por exemplo, en casos como X^2+Y^2 isto é certo: $X^2+Y^2=(X+iY)(X-iY)$. En cambio, se consideramos $X^2+Y^2-Z^2$, o resultado non é verdade: para cada par (Y,Z), o polinomio terá dúas raíces sobre $\mathbb{C}$, pero non existen expresión polinómicas $p(Y,Z)$ e $q(Y,Z)$ que cumpran $X^2+Y^2-Z^2=(X-p(Y,Z))(X-q(Y,Z))$. Se tiveramos considerado $X^2+Y^2+2XY-Z^2$, tense que $X^2+Y^2+2XY-Z^2=(X+Y+Z)(X+Y-Z)$; desde un punto de vista xeométrico, o feito de que un polinomio en varias variables non factorice quere dicir que o obxecto que representa é irredutible, é dicir, non ten *variedades alxébricas* (conxuntos representados por polinomios) de dimensión menor contidas dentro del.

Sobre a aplicación proposta, para calcular o determinante da matriz que ten por columnas as coordenadas dos vectores, comézase sacando factor común de cada fila o número t_i^m. O que queda é o determinante de Vandermonde dos números $t_1,\dots,t_n$, que polo tanto é igual a $\prod_{i=1}^n t_i^m\prod_{i>j}(t_i-t_j)$; está claro que este número é diferente de 0.

Capítulo 3

Estrutura das aplicacións lineais

Sexa K un corpo. Dada unha aplicación lineal $f\colon E \to F$ entre dous K-espazos vectoriais, pódense escoller bases dos dous espazos de maneira que a matriz correspondente sexa moi sinxela. Por exemplo, se E e F teñen a mesma dimensión podemos facer que sexa unha matriz diagonal con uns e ceros. 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 de dimensión n. Ao longo desta parte, imos traballar tres cuestións fundamentais.

- **Diagonalización.** Caracterizar os endomorfismos que admiten unha matriz diagonal. Xeometricamente, iso quere dicir que hai n direccións linealmente independentes nas que o endomorfismo se corresponde coa multiplicación por un escalar.

- **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.

En calquera caso, non todos os endomorfismos admiten forma de Jordan, polo que quedarán aínda casos por cubrir, algúns dos cales serán tratados na parte final do capítulo.

3.1. Diagonalización

Vectores propios e valores propios

Definición 3.1. Sexa $f\colon E \to E$ 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 poder haber moitos vectores propios que o teñen como valor propio.

Definición 3.2. Dado un endomorfismo $f \in \mathrm{End}(E)$ e un escalar $\lambda \in K$, considérase o endomorfismo $f_\lambda := f - \lambda \,\mathrm{Id} \in \mathrm{End}(E)$. O seu núcleo é

$$E_\lambda := \{v \in E \mid f(v) = \lambda v\} = \ker(f - \lambda \,\mathrm{Id}),$$

que é un subespazo vectorial de E formado polo vector cero e todos os vectores propios de f con valor propio λ, en caso de haber algún. Cando λ é un valor propio, falamos do *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 valor propio, e $(-2, 3)$ é un xerador do subespazo de vectores propios correspondente.

Proposición 3.1. 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 $f_{\lambda_r} = f - \lambda_r \,\mathrm{Id}$ a cada lado da igualdade 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 que os vectores v_i, para $1 \leq i \leq r - 1$, son independentes, e como os coeficientes $\lambda_i - \lambda_r$ son diferentes de 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 todos os 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 $v \in E$ como suma dos diferentes subespazos, se existe, é única, é dicir, se

$$v = v_1 + \ldots + v_r = w_1 + \ldots + w_r \text{ con } v_i, w_i \in E_{\lambda_i},$$

entón $v_i = w_i$ para todo i. Polo tanto, se $\lambda_1, \ldots, \lambda_r$ son valores propios dun endomorfismo $f \in \mathrm{End}(E)$ e para un certo 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 conxuntos $\cup_{i=1}^{r} \mathcal{B}_i$ é unha familia de vectores linealmente independentes.

Polinomio característico

Definición 3.3. O *polinomio característico* dunha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_n(K)$ é o polinomio

$$\operatorname{Char}(A; X) = \det(A - X \operatorname{Id}) = \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - X & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - X \end{vmatrix}.$$

O polinomio característico dun endomorfismo $f \in \operatorname{End}(E)$ 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 $\det(X \operatorname{Id} - A)$, 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, $\operatorname{Tr}(A)$, é a suma dos elementos da súa diagonal.

Proposición 3.2. Sexa f un endomorfismo e A a súa matriz asociada nunha base fixada. O polinomio característico cumpre as seguintes propiedades.

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

(b) $\operatorname{Char}(f; X)$ ten grao n e o coeficiente principal é $(-1)^n$.

(c) O coeficiente de X^{n-1} en $\operatorname{Char}(f; X)$ é $(-1)^{n-1} \operatorname{Tr}(A)$.

(d) O termo independente de $\operatorname{Char}(f; X)$ 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, o polinomio característico na base $\mathcal{B}'$ cumpre que

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

Como non depende da base, poremos simplemente $\operatorname{Char}(f; X)$.

(b) O coeficiente principal é o produto dos coeficientes que acompañan a X 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 corresponde a coller a suma de todos os 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 3.3. 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. $\square$

Definición 3.4. 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.

Convén ter en conta, á hora de realizar os cálculos, que a dimensión de E_λ coincide con $n - \mathrm{rango}(f - \lambda\,\mathrm{Id})$.

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

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

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, resolvemos

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

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 previo sobre determinantes.

Lema 3.1. Sexa $A = (a_{ij}) \in \mathcal{M}_n(K)$ unha matriz da forma

$$A = \left(\begin{array}{c|c} X & Y \\ \hline 0 & Z \end{array}\right),$$

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 $\det(A) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}$. 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_r coas permutacións de $\{1,\ldots,r\}$ e S_{n-r} coas permutacións de $\{r+1,\ldots,n\}$. Entón,

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

□

Proposición 3.4. As multiplicidades alxébrica e xeométrica dun valor propio λ cumpren as desigualdades

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

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

$$A = \left(\begin{array}{c|c} \lambda\cdot\mathbb{I}_k & A_1 \\ \hline 0_{(n-k)\times k} & A_2 \end{array}\right),$$

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

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 3.5. 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 $P^{-1}AP$ é diagonal.

Proposición 3.5. 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:

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

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 $\{v_{1,1},\dots,v_{1,m_1}\}$ unha base ordenada correspondente ao primeiro vector propio, λ_1; $\{v_{2,1},\dots,v_{2,m_2}\}$, 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 $\langle v_{k,1},\dots,v_{k,m_k}\rangle$ 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 (ou dito doutro xeito, unha combinación lineal dos vectores $v_{i,j}$, con $i \geq 2$, non pode ser un vector propio de valor propio λ_1).
Alternativamente, podemos demostrar o resultado directamente sen empregar as proposicións anteriores: se houbera outro vector $v = \sum_{i=1}^{r}\sum_{j=1}^{m_i} x_{ij}v_{ij} \in E_{\lambda_k}$, 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 finalmente 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. □

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

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

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

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 todas as 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 $E_{-i} = \langle (0,1,-i) \rangle$. Polo tanto,
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix}^{-1}.$$

- *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. Observamos que non tería sido preciso calcular os xeradores de $\ker(B-\mathbb{I})$ e sería suficiente observar que, como $B - \mathbb{I}$ ten rango 2, o núcleo ten dimensión 1.

- *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

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

Exemplo. Consideramos a matriz $A = \begin{pmatrix} 4 & 3 \\ -2 & -1 \end{pmatrix}$. Entón, para calcular A^n, observamos que os valores propios de A son $\lambda_1 = 2$ e $\lambda_2 = 1$. O subespazo propio para o valor propio 2 está xerado por $(3,-2)$, e para o valor propio 1 está xerado por $(-1,1)$. Polo tanto,

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

Por conseguinte,

$$A^n = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 \cdot 2^n - 2 & 3 \cdot 2^n - 3 \\ -2 \cdot 2^n + 2 & -2 \cdot 2^n + 3 \end{pmatrix}.$$

Nos exercicios explicarase como empregar esta mesma idea para o cálculo da raíz cadrada dunha matriz (ou, máis en xeral, doutras funcións como a exponencial, o seno ou o coseno).

3.2. Polinomio mínimo e subespazos invariantes

Polinomio mínimo

Na sección anterior, vimos que cando un endomorfismo f diagonaliza con valores propios $\lambda_1, \ldots, \lambda_r$, entón o espazo E é a suma directa

$$E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r} = \ker f_{\lambda_1} \oplus \ldots \oplus \ker f_{\lambda_r}.$$

Estudaremos agora outros obxectos que nos permitirán considerar descomposicións análogas en situacións nas que non se cumpran as condicións de diagonalización.

Para ese propósito, cómpre introducir primeiro certa notación. Sexa f un endomorfismo. Podemos considerar as súas potencias $f^0 = \mathrm{Id}$, $f^1 = f$, $f^2 = f \circ f$ e así sucesivamente, tendo $f^r = f \circ f^{r-1}$ para calquera n.

Definición 3.6. Sexa K un corpo. Unha *álxebra* F sobre K é un espazo vectorial cunha operación bilineal $F \times F \to F$ (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* $\Phi\colon F \to F$ 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 3.7. Sexa f un endomorfismo. A aplicación *avaliación en* f é o morfismo dado por

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

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

Proposición 3.6. O conxunto de polinomios anuladores diferentes de cero é non baleiro. Ademais, existe un único polinomio $\mathrm{Min}(f;X)$ non nulo, mónico e de grao mínimo no núcleo de Φ_f. Ademais, 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 todas as 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 todos os 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. $\square$

A demostración pódese acurtar empregando o feito de que $K[X]$ é o que se coñece como un *dominio de ideais principais.* A modo de notación, 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 $f = \mathrm{Id}$, entón cúmprese que $\mathrm{Min}(f;X) = X - 1$. 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 cuxas matrices na base canónica son

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

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; en moitas ocasións, vainos interesar impor unha condición menos restritiva que a anulación do endomorfismo f, e vainos chegar que se anule ao avalialo nun vector ou nun subespazo vectorial. Polo tanto, podemos considerar a aplicación

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

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. Enunciamos agora a seguinte proposición, que tamén se pode formular de xeito análogo no caso de $\mathrm{Min}(f;X)$.

Proposición 3.7. Sexa $\mathrm{Min}(f, v; X) = X^s + \ldots + a_1 X + a_0$ 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 $p(f)(v) = 0$. Iso é contraditorio coa definición de $\mathrm{Min}(f, v; X)$.
Se $t = s$, $\mathrm{Min}(f, v; f)(v) = a_0 v + a_1 f(v) + \ldots + f^s(v) = 0$, co cal os vectores son linealmente dependentes. Para $t > s$ procedemos por indución, usando como hipótese que $f^{t-1}(v) \in \langle v, f(v), \ldots, f^{s-1}(v)\rangle$. De aquí temos entón que

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

e a conclusión do enunciado é certa. □

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 $v_1 = (1, 0)$, temos que $X - 2$ é anulador xa que

$$(B - 2\mathbb{I})v_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Do mesmo xeito, $X - 3$ é un anulador para o vector $v_2 = (0, 1)$. 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).

Subespazos invariantes

Nas primeiras seccións deste capítulo estudamos os vectores propios. Un vector propio v dun endormofismo f cumpre que $F = \langle v \rangle$ é un subespazo vectorial de dimensión 1 que cumpre que $f(F) \subset F$, é dicir, a imaxe do subespazo está contida dentro del. En moitas ocasións non será posible dar unha descomposición en subespazos de dimensión 1 que sexan invariantes por f, pero si poderemos falar máis en xeral de subespazos (de dimensión posiblemente maior) que se quedan invariantes polo endomorfismo f.

Definición 3.9. 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 $f|F$

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

e que se acostuma chamar *restrición de f a F*.

Ás veces enténdese que a restrición a un subespazo é simplemente a aplicación $f \colon F \to E$ que envía v a $f(v)$; para non introducir notación innecesaria, e porque será o máis habitual no noso contexto, aquí só imos empregar esa notación para os subespazos invariantes.
Como comentamos antes, o subespazo xerado por un vector propio é un subespazo invariante de dimensión 1, ou máis en xeral r vectores propios linealmente independentes xeran un subespazo invariante de dimensión r (en particular, este sempre é o caso se os valores propios son distintos). Outro caso que terá especial relevancia ao estudar a forma de Jordan será o xerado polas sucesivas aplicacións dun endomorfismo f a un vector, isto é, $\{v, f(v), \ldots, f^{s-1}(v)\}$, sendo s o grado do polinomio mínimo $\mathrm{Min}(f, v; X)$.

Proposición 3.8. 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 $\gcd(\mathrm{Min}(f|F; X), \mathrm{Min}(f|G; X)) = 1$, 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 $\gcd(p(X), q(X)) = 1$, $E = \ker(p(f)) \oplus \ker(q(f))$.

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 $q(f)$, temos que

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

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, todas as 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. Por exemplo, se o endomorfismo f está dado pola matriz

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

temos que $\mathrm{Min}(f;X) = (X-2)(X-3)$ e entón $E = \ker(f_2) \oplus \ker(f_3)$, onde f_2 e f_3 refírense á restrición de f aos subespazos propios para os valores propios 2 e 3, respectivamente.

O subespazo $\ker(p(f))$ é simplemente o conxunto de vectores v tales que $p(f)(v) = 0$; en particular, $p(f)$ sempre é un múltiplo de $\mathrm{Min}(f,v;X)$.

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 3.2. Sexa $\mathrm{Min}(f;X) = p(X)q(X)$, con $\gcd(p(X), q(X)) = 1$ e $p(X)$ e $q(X)$ mónicos. Logo, 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 son os polinomios mínimos das restricións a $\ker(p(f))$ e $\ker(q(f))$, respectivamente. 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. Polo tanto, $p(X) = a(X)r(X)$ e $q(X) = b(X)s(X)$, con $a(X)$ e $b(X)$ polinomios mónicos. Como $p(X)q(X) = r(X)s(X)$, conclúese que $a(X) = b(X) = 1$ e tense que $r(X) = p(X)$ e $s(X) = q(X)$. □

Podemos enunciar agora o primeiro teorema de descomposición.

Proposición 3.9 (Primeiro teorema de descomposición)**.** 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, entón 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 única e cúmprese que $E_i = \ker(m_i(f)^{n_i})$ para todo $i = 1, \ldots, r$.

Demostración. Dos resultados anteriores, a existencia da descomposición está clara. Supoñamos agora 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

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

□

Exemplo. Consideremos os endomorfismos que teñen por matrices na base canónica

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

Non é difícil ver que $\mathrm{Min}(A;X) = (X-2)(X-3)$ e $\mathrm{Min}(B;X) = (X-2)^2(X-3)$. O resultado anterior dinos entón que o endomorfismo dado por A dá lugar á descomposición $E = \ker(A - 2\mathbb{I}) \oplus \ker(A - 3\mathbb{I})$, mentres que o dado por B orixina a descomposición $E = \ker((A - 2\mathbb{I})^2) \oplus \ker(A - 3\mathbb{I})$.

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 3.10. 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 todos os $\mathrm{Min}(f, v; X)$. $\square$

Proposición 3.11. Un endomorfismo $f \in \mathrm{End}(E)$ 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 todos os λ_i diferentes entre si. Entón, polo primeiro teorema de descomposición, $E = E_1 \oplus \ldots \oplus E_r$, con $E_i = \ker(f - \lambda_i \,\mathrm{Id})$ 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 $f|E_i$ 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. $\square$

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 non repetidos.

Teorema de Cayley–Hamilton

Polo de agora, vimos que se $f \in \mathrm{End}(E)$ diagonaliza, o espazo E é a suma directa dos subespazos de vectores propios e que o polinomio mínimo descompón en factores lineais, un por cada factor $X - \lambda_i$, sendo λ_i os vectores propios. En xeral, nos casos nos que o endomorfismo non diagonalice, non é obvio como achar o polinomio mínimo. Nesta sección explórase esa cuestión, atopando un polinomio anulador que polo tanto nos permite buscar o mínimo entre os seus divisores.

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

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

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

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

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 3.12. Sexa λ un escalar e $f \in \operatorname{End}(E)$. Entón, as seguintes tres propiedades son equivalentes.

(i) λ é un valor propio de f.

(ii) λ é un cero de $\operatorname{Min}(f; X)$

(iii) λ é un cero de $\operatorname{Char}(f; X)$.

Demostración. A única propiedade que non demostramos polo de agora é que se λ é unha raíz de $\operatorname{Min}(f; X)$, entón é un valor propio (e polo tanto un cero do polinomio característico).
Se λ é un cero de $\operatorname{Min}(f; X)$, entón $\operatorname{Min}(f; X) = (X - \lambda)m(X)$, 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

$$(f - \lambda \operatorname{Id})(w) = (f - \lambda \operatorname{Id})m(f)(v) = \operatorname{Min}(f; f)(v) = 0.$$

Polo tanto, λ é un valor propio.

□

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. Porén, convén observar que este resultado é suficiente para establecer o teorema de Cayley–Hamilton sobre os complexos ou sobre outros corpos nos que un polinomio descompón completamente en factores lineais.

Proposición 3.13. Sexa f un endomorfismo que cumpre que tanto o seu polinomio característico como o seu polinomio mínimo descompoñen 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 f é unha matriz con bloques diagonais, de maneira que se $\mathrm{Char}(f|E_i;X)$ 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, como $(X-\lambda_i)^{n_i}$ é o polinomio mínimo da restrición de f a E_i, hai un $v_i \in E_i$ tal que $(f-\lambda_i)^{n_i}(v_i) = 0$ pero $(f-\lambda_i)^{n_i-1}(v_i) \neq 0$. Entón o polinomio mínimo de v_i é $(X-\lambda_i)^{n_i}$ 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. □

Este resultado é suficiente para achar o polinomio mínimo en moitos casos.

Exemplo. Sexa

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

a representación matricial dun endomorfismo sobre os complexos. Sexa $v_1 = (1,0,0)$ e $v_2 = (0,0,1)$. Como v_1 e v_2 son vectores propios de valores propios 2 e 3, temos que $(X-2)$ e $(X-3)$ son factores do polinomio mínimo. Pero como o endomorfismo non diagonaliza (a multiplicidade alxébrica de 2 non coincide coa xeométrica), o polinomio mínimo ten que ter grao como pouco 3. O polinomio característico é $(X-2)^2(X-3)$, polo tanto este ten que ser tamén o mínimo.

Antes de pasar ao resultado principal da sección, precisamos do seguinte resultado sobre determinantes.

Lema 3.3. Sexan $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$ elementos de K. Entón, dada a matriz

$$A = \left(\begin{array}{cccc|c} 0 & 0 & \ldots & 0 & -\lambda_0 \\ \hline 1 & 0 & \ldots & 0 & -\lambda_1 \\ 0 & 1 & \ldots & 0 & -\lambda_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\lambda_{n-1} \end{array}\right),$$

cúmprese que $\mathrm{Char}(A;X) = (-1)^n(X^n + \lambda_{n-1}X^{n-1} + \ldots + \lambda_1 X + \lambda_0)$.

Demostración. O resultado demóstrase por indución en n. Supoñamos que se cumpre para matrices $(n-1)\times(n-1)$, e sexa $q(X) = (-1)^{n-1}(X^{n-1} + \lambda_{n-1}X^{n-2} + \ldots + \lambda_1)$. Desenvolvendo pola primeira fila, temos que

$$\begin{aligned} \mathrm{Char}(A;X) &= (-X)q(X) + (-1)^{n+1}(-\lambda_0) \\ &= (-1)^n(X^n + \lambda_{n-1}X^{n-1} + \ldots + \lambda_1 X + \lambda_0). \end{aligned}$$

□

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

Teorema 3.1 (Cayley–Hamilton)**.** Toda matriz cadrada anula o seu polinomio característico. É dicir, se $f \in \mathrm{End}(E)$, entón $\mathrm{Char}(f; f) = 0$.

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 $f \in \mathrm{End}(E)$ de maneira que $v, f(v), f^2(v), \ldots, f^{n-1}(v)$ é unha base de E. Entón $\mathrm{Char}(f; f) = 0$.

(ii) Se A é unha matriz por bloques $\left(\begin{array}{c|c} M & N \\ \hline 0 & P \end{array}\right)$ de maneira que M e P cumpren Cayley–Hamilton, entón A tamén o cumpre.

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

$$\left(\begin{array}{cccc|c} 0 & 0 & \ldots & 0 & -\lambda_0 \\ \hline 1 & 0 & \ldots & 0 & -\lambda_1 \\ 0 & 1 & \ldots & 0 & -\lambda_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\lambda_{n-1} \end{array}\right).$$

Nesta base temos que o polinomio característico é $\mathrm{Char}(f; X) = (-1)^n \sum_{i=0}^{n} \lambda_i X^i$. Vexamos que $\mathrm{Char}(f; f)$ se anula en todos os vectores da base $\{v, f(v), \ldots, f^{n-1}(v)\}$. Para iso, observamos que

$$\begin{aligned} \mathrm{Char}(f; f)(f^k(v)) &= \Big(\sum_{i=0}^{n} \lambda_i f^i\Big)(f^k(v)) \\ &= \Big(\sum_{i=0}^{n} \lambda_i f^{k+i}\Big)(v) \\ &= f^k\Big(\sum_{i=0}^{n} \lambda_i f^i\Big)(v) \\ &= f^k(0) = 0, \end{aligned}$$

onde se empregou que $\Big(\sum_{i=0}^{n} \lambda_i f^i\Big)(v) = 0$ para todo v. Polo tanto, o polinomio característico anula todos os elementos da base e é un polinomio anulador do endomorfismo. A segunda propiedade é consecuencia directa do Lema 3.1, que afirma que o determinante de A é $\det(M)\cdot\det(P)$. Máis concretamente, $\mathrm{Char}(A; X) = \mathrm{Char}(M; X)\cdot\mathrm{Char}(P; X)$. Entón

$$\begin{aligned} \mathrm{Char}(A; A) &= \mathrm{Char}(M; A)\,\mathrm{Char}(P; A) \\ &= \left(\begin{array}{c|c} \mathrm{Char}(M; M) & * \\ \hline 0 & \mathrm{Char}(M; P) \end{array}\right) \left(\begin{array}{c|c} \mathrm{Char}(P; M) & * \\ \hline 0 & \mathrm{Char}(P; P) \end{array}\right) \\ &= \left(\begin{array}{c|c} 0 & * \\ \hline 0 & \mathrm{Char}(M; P) \end{array}\right) \left(\begin{array}{c|c} \mathrm{Char}(P; M) & * \\ \hline 0 & 0 \end{array}\right) \\ &= \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array}\right). \end{aligned}$$

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 maior r tal que $v, f(v), \dots, f^{r-1}(v)$ son linealmente independentes, pero hai unha relación de dependencia lineal entre $v, f(v), \dots, f^r(v)$. Como $v \neq 0$, necesariamente $r \geq 1$. Se $r = n$, concluímos pola primeira parte. Senón, completamos $\{v, f(v), \dots, f^{r-1}(v)\}$ a unha base de E, na cal a matriz asociada é unha matriz por bloques $A = \left(\begin{array}{c|c} M & N \\ 0 & P \end{array}\right)$, con M un bloque $r \times r$ e P un bloque $(n-r) \times (n-r)$. Pola hipótese de indución, tanto M como P cumpren o teorema de Cayley–Hamilton, polo que podemos concluír. $\square$

Exemplo. Consideramos a matriz

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

que non ten valores propios racionais. Temos que $\mathrm{Char}(A; X) = X^2 + 2X + 2$, o que quere dicir que $A^2 + 2A + 2\mathbb{I} = 0$.

Un dos aspectos importantes do teorema de Cayley–Hamilton é que nos dá un polinomio que anula a matriz; polo tanto, o polinomio mínimo pódese buscar entre os seus divisores. Por outro lado, o uso do teorema pódese empregar para achar a matriz inversa. No exemplo anterior, como $A^2 + 2A + 2\mathbb{I} = 0$, tense $A(A + 2\mathbb{I}) = -2\mathbb{I}$, polo que

$$\mathbb{A}^{-1} = -\frac{1}{2}(A + 2\mathbb{I}).$$

3.3. 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, que permite atopar bases *canónicas* para un endomorfismo que non ten suficientes valores propios, pero no que o polinomio característico aínda ten tantas raíces como o seu grao.

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.

Vectores propios xeneralizados

Antes de comezar, observamos que, polo primeiro teorema de descomposición, é suficiente tratar por separado cada un dos subespazos asociados aos diferentes valores propios dun endomorfismo. Recordemos que pomos $f_\lambda = f - \lambda\mathbb{I}$. Se o polinomio mínimo é o produto de factores da forma $(X - \lambda_i)^{m_i}$, podemos considerar de forma independente cada un dos núcleos $\ker(f_\lambda)^{m_i}$. Máis concretamente, se $\mathrm{Char}(f; X) = (-1)^n \prod_{i=1}^t (X - \lambda)^{m_i}$, entón

$$E = \bigoplus_\lambda \ker((f - \lambda \,\mathrm{Id})^{m_i}),$$

posto que polo teorema de Cayley–Hamilton o polinomio característico é un múltiplo do polinomio mínimo. O obxectivo desta sección é estudar, para cada λ, o subespazo $\ker((f - \lambda \,\mathrm{Id})^{m_i})$.
Para introducir a forma de Jordan, imos comezar estendendo a definición de vector propio a través da seguinte xeralización.

Definición 3.10. 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 $f_\lambda(v_r) = v_{r-1}, \ldots, f_\lambda(v_2) = v_1$ 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.
Observamos tamén que a matriz na base (v_3, v_2, v_1) está dada por

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

É dicir, é unha matriz con entradas na diagonal e con uns na diagonal superior.

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 3.14. 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 $x_1 v_1 + \ldots + x_\ell v_\ell = 0$. Aplicando f_λ, vemos que

$$f_\lambda(x_1 v_1 + \ldots + x_\ell v_\ell) = x_2 v_1 + \ldots + x_\ell v_{\ell-1}.$$

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 $f_\mu = f_\lambda + (\lambda - \mu)\,\mathrm{Id}$, 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 en $\mathcal{M}_k(\lambda)$ da forma

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix},$$

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

Definición 3.11. Unha *base de Jordan* para un endomorfismo f é unha base formada pola unión de ciclos de Jordan para este endomorfismo.

A matriz de Jordan dun endomorfismo será, polo tanto, unha matriz por bloques na que cada peza é un ciclo de Jordan:

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & & & & \\ & J_{k_2}(\lambda_2) & & & 0 \\ & & J_{k_3}(\lambda_3) & & \\ & 0 & & \ddots & \\ & & & & J_{k_\ell}(\lambda_\ell) \end{pmatrix}$$

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 3.15. Sexa $g \in \operatorname{End}(E)$. 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 $\varphi\colon \ker(g^{k+1})/\ker(g^k) \to \ker(g^k)/\ker(g^{k-1})$ 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 $g^k(v) = g^{k-1}(g(v)) = 0$ 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 3.4. Sexa λ un valor propio de $f \in \operatorname{End}(E)$. 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 todos os 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 todos os 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 $k \le k'$ ciclos que se obteñen eliminando o último vector de todos os 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 a aplicación $\ker(f_\lambda^{j+1})/\ker(f_\lambda^j) \to \ker(f_\lambda^j)/\ker(f_\lambda^{j-1})$ inducida por f_λ é inxectiva e polo tanto preserva a condición de independencia lineal.
Pola hipótese de indución, todos os coeficientes desta combinación lineal son cero, e polo tanto todos os coeficientes $x_{i,j}$ da combinación lineal inicial con $j > 1$ tamén son cero. □

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

Teorema 3.2. 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. En primeiro lugar, observamos que se pode estudar cada subespazo por separado, buscando bases de Jordan cuxa lonxitude coincida coa multiplicidade alxébrica do valor propio; unha base de Jordan nunca pode ter tamaño maior, porque se ese fose o caso, ao considerarmos o polinomio característico da forma de Jordan, teriamos que a multiplicidade alxébrica é máis grande do permitido.
Hai unha cadea de subespazos $\ker(f_\lambda) \subset \ker(f_\lambda^2) \subset \dots$ que nalgún momento estabiliza (de feito, con expoñente $\le m$). Sexa $\ell \le m$ 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 $\ell - 1$. Isto é posible porque f_λ induce unha aplicación inxectiva entre os cocientes correspondentes.
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 $\ker(f_\lambda^{j-1})/\ker(f_\lambda^{j-2})$ porque a aplicación inducida por f_λ, $\ker(f_\lambda^j)/\ker(f_\lambda^{j-1}) \to \ker(f_\lambda^{j-1})/\ker(f_\lambda^{j-2})$, é inxectiva. Logo, 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. □

Exemplo. Imos ilustrar as ideas do lema e do teorema nun exemplo. Sexa f un endomorfismo tal que $\mathrm{Char}(f;X) = (X-2)^{18}$. Entón, sabemos que existirá un $1 \leq m \leq 18$ tal que $(f - 2\,\mathrm{Id})^m = 0$; o menor m con esa propiedade é o expoñente de $(X-2)$ no polinomio mínimo. Supoñamos que ao calcular as dimensións dos núcleos obtemos os seguintes valores:

- $\dim \ker(f - 2\,\mathrm{Id}) = 7$.
- $\dim \ker(f - 2\,\mathrm{Id})^2 = 13$.
- $\dim \ker(f - 2\,\mathrm{Id})^3 = 16$.
- $\dim \ker(f - 2\,\mathrm{Id})^4 = 18$.

Polo tanto, o polinomio mínimo é $(X-2)^4$. Para facer a matriz de Jordan, consideramos dous vectores que sexan linealmente independentes en $\ker(f - 2\,\mathrm{Id})^4/\ker(f - 2\,\mathrm{Id})^3$, e chamámoslles $v_{1,4}$ e $v_{2,4}$.
A continuación, consideramos $v_{1,3} = f_2(v_{1,4})$ e $v_{2,3} = f_2(v_{2,4})$. Estes vectores son linealmente independentes en $\ker(f - 2\,\mathrm{Id})^3/\ker(f - 2\,\mathrm{Id})^2$, e completámolos cun vector $v_{3,3}$ para ter unha base do espazo. Alternativamente, $\{v_{1,4}, v_{2,4}, v_{1,3}, v_{2,3}, v_{3,3}\}$ é unha base de $E/\ker(f - 2\,\mathrm{Id})^2$.
Definimos $v_{1,2}, v_{2,2}, v_{3,2}$ como as imaxes de $v_{1,3}, v_{2,3}, v_{3,3}$ por f_2, e completamos con vectores $v_{4,2}, v_{5,2}, v_{6,2}$ para ter unha base de $\ker(f - 2\,\mathrm{Id})^2/\ker(f - 2\,\mathrm{Id})$.
Por último, para cada $i = 1, \ldots, 6$, sexa $v_{i,1} = f_2(v_{i,2})$, e consideramos un novo vector $v_{7,1}$ que forme xunto cos seis anteriores unha base de $\ker(f - 2\,\mathrm{Id})$.
Desta maneira, os ciclos de Jordan podémolos interpretar en forma da seguinte táboa.

$v_{1,4}$	$v_{2,4}$					
$v_{1,3}$	$v_{2,3}$	$v_{3,3}$				
$v_{1,2}$	$v_{2,2}$	$v_{3,2}$	$v_{4,2}$	$v_{5,2}$	$v_{6,2}$	
$v_{1,1}$	$v_{2,1}$	$v_{3,1}$	$v_{4,1}$	$v_{5,1}$	$v_{6,1}$	$v_{7,1}$

A forma de Jordan, polo tanto, constará de 2 bloques de tamaño 4, $J_{2,4}$, un bloque de tamaño 3, $J_{2,3}$, 3 bloques de tamaño 2, $J_{2,2}$, e un bloque de tamaño 1, $J_{2,1}$:

$$J_{2,4} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix},\ J_{2,3} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},\ J_{2,2} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},\ J_{2,1} = (2).$$

Máis en xeral, calquera táboa de Jordan terá as seguintes propiedades.

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

Da demostración do teorema, temos o seguinte corolario.

Corolario 3.1. Sexa $f \in \mathrm{End}(E)$ de maneira que $\mathrm{Char}(f;X)$ descompón completamente, e sexa λ un valor propio de f. Entón, a multiplicidade de λ como raíz de $\mathrm{Min}(f;X)$ é igual ao menor enteiro positivo k tal que $\ker(f-\lambda\,\mathrm{Id})^k = \ker(f-\lambda\,\mathrm{Id})^{k+1}$. Alternativamente, k é o menor enteiro positivo de maneira que a dimensión de $\ker(f - \lambda\,\mathrm{Id})^k$ é igual a $m_a(\lambda)$.

Exemplo. Se na descomposición en ciclos de Jordan hai un único ciclo de Jordan de valor propio λ, o expoñente de λ como raíz do polinomio mínimo é precisamente a multiplicidade alxébrica.

Máis concretamente, observamos que unha base de vectores propios xeralizados para un valor propio λ de multiplicidade alxébrica m_λ ten cardinal m_λ. Para demostralo, sexa F_λ o subespazo invariante xerado por todos os vectores propios xeralizados de valor propio λ. O polinomio característico de $f|F_\lambda$ divide ao de f; como o primeiro é $(-1)^{\dim(F)}(X-\lambda)^{\dim(F)}$, temos que $\dim(F_\lambda) \leq m_\lambda$. Como esta desigualdade se cumpre para calquera valor propio λ, polo primeiro teorema de descomposición

$$E = \bigoplus_\lambda \ker((f-\lambda)^{m_\lambda});$$

tense que

$$\dim E = \sum_\lambda \dim F_\lambda \leq \sum_\lambda m_\lambda = \dim E,$$

e polo tanto todas as desigualdades teñen que ser igualdades.
Unha primeira consecuencia da existencia de bases de Jordan é o seguinte resultado.

Proposición 3.16. 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, salvo a orde das caixas.

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. $\square$

Exemplos de cálculo de matrices de Jordan

O procedemento para o cálculo da base de Jordan dun endomorfismo f será sempre o seguinte.

1. Calcular o polinomio de Jordan do endomorfismo f e achar as súas raíces.

2. Para cada raíz, achamos o menor enteiro positivo k tal que a dimensión do núcleo de $(f - \lambda\,\mathrm{Id})^k$ sexa igual á multiplicidade alxébrica.

3. Para ese k, consideramos vectores de E de maneira que formen unha base de $\ker(f - \lambda\,\mathrm{Id})^k / \ker(f - \lambda\,\mathrm{Id})^{k-1}$. Iso dános os vectores do piso k na táboa de Jordan.

4. Para cada i entre $k-1$ e 1, procedemos do seguinte xeito: calculamos a imaxe por f_λ dos vectores no piso $i+1$ e completamos con tantos vectores de E como faga falta ata ter unha base de $\ker(f - \lambda\,\mathrm{Id})^i / \ker(f - \lambda\,\mathrm{Id})^{i-1}$.

Comezamos considerando a matriz

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

Neste caso, $\mathrm{Char}(A; X) = -X(X-1)^2$, co cal os valores propios son 0 e 1. O valor propio 0 ten multiplicidade alxébrica 1, co cal é suficiente con achar $\ker(A) = \langle(0,-1,2)\rangle$. En cambio, $\ker(A - \mathbb{I}) = \langle(1,-1,5)\rangle$, e polo tanto, o 1 ten multiplicidade xeométrica 1. Isto automaticamente dinos que o polinomio mínimo é $\mathrm{Min}(A; X) = X(X-1)^2$, posto que ten que ser un divisor del coas mesmas raíces e ao mesmo tempo non pode ter todas as raíces simples xa que non diagonaliza.
Achamos agora $\ker(A - \mathbb{I})^2$, e é suficiente coller un vector que estea nese núcleo pero non no de $A - \mathbb{I}$. Sexa $v_3 = (1,2,0)$. Entón $v_1 = (A - \mathbb{I})v_2 = (1,-1,5)$. Polo tanto, unha descomposición da matriz é

$$\begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ -1 & -1 & 2 \\ 2 & 5 & 0 \end{pmatrix} \left(\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right) \begin{pmatrix} 0 & 1 & 1 \\ -1 & -1 & 2 \\ 2 & 5 & 0 \end{pmatrix}^{-1}.$$

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 $\dim\ker(B - \mathbb{I})^3 = 5$. O polinomio mínimo é entón $(X-1)^3$. 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 $v_4 = (B-\mathbb{I})v_5 = (2,0,1,3,0)$. Observamos que a elección de v_5 é correcta porque é linealmente independente con v_2 no cociente, é dicir,

$$v_5 = av_2 + bv_1 + cv_4$$

non ten solucións (aquí, collemos v_1 e v_4 como base de $\ker(B-\mathbb{I})$) posto que tanto para v_2, v_1 e v_4 a quinta compoñente é 0 e iso non é o caso para v_5.
Polo tanto,

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & -1 & -4 \\ 1 & 1 & 2 & -1 & -3 \\ 1 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|cc} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \begin{pmatrix} 1 & 0 & 1 & 2 & 3 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1}.$$

A táboa de Jordan cos vectores por ciclos ten a forma seguinte.

v_3	
v_2	v_5
v_1	v_4

O seguinte corolario é inmediato a partir dos resultados anteriores sobre a construción das matrices de Jordan.

Corolario 3.2. Sexa $f \in \mathrm{End}(E)$ un endomorfismo que admite forma de Jordan e sexa λ un valor propio de f. Sexa k o menor enteiro positivo tal que $\ker((f-\lambda\,\mathrm{Id})^k) = \ker((f-\lambda\,\mathrm{Id})^{k+1})$. Entón, k é o tamaño do maior bloque de Jordan de f.

3.4. Aplicacións e extensións da forma de Jordan

Imos explorar agora dúas aplicacións habituais da forma de Jordan: a resolucións de recorrencias e o cálculo da exponencial dunha matriz, que se emprega na resolución de ecuacións diferenciais ordinarias.

Cálculo de subespazos invariantes

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, polo que nos imos restrinxir ao caso dun valor propio λ. O seguinte resultado establece que a forma de Jordan dun subespazo invariante está contida na forma de Jordan de E.

Proposición 3.17. Sexa $f \in \mathrm{End}(E)$ con polinomio característico $\mathrm{Char}(f;X) = (\lambda - X)^n$ e sexa $F \subset E$ un subespazo invariante por F. Sexa $(v_{i,j})$, con $1 \le i \le r$ e $1 \le j \le \ell_i$ unha base de Jordan para F. Entón, a base de Jordan de F pode estenderse a unha base de Jordan de E, é dicir, existe unha base de Jordan de E da forma $(u_{i,j})$, con $1 \le i \le s$, $1 \le j \le m_i$, $r \le s$, $\ell_i \le m_i$ e $u_{i,j} = v_{i,j}$ para $1 \le i \le r$ e $1 \le j \le \ell_i$. (Na descomposición en ciclos de Jordan de E non estamos requirindo agora que os ciclos estean ordenados pola súa lonxitude.)

Demostración. Comezamos observando que o número de bloques de tamaño maior ou igual que k na restrición a un subespazo F é menor ou igual que os que existen en E. Para ver iso, observamos que, chamando $c_k = \dim\ker(f - \lambda\,\mathrm{Id})^k|F - \dim\ker(f - \lambda\,\mathrm{Id})^{k-1}|F$, temos

$$\begin{aligned} c_k &= \dim(\ker(f - \lambda\,\mathrm{Id})^k|F - \ker(f - \lambda\,\mathrm{Id})^{k-1}|F) \\ &\leq \dim(\ker(f - \lambda\,\mathrm{Id})^k - \ker(f - \lambda\,\mathrm{Id})^{k-1}) \\ &= \dim\ker(f - \lambda\,\mathrm{Id})^k - \dim\ker(f - \lambda\,\mathrm{Id})^{k-1}. \end{aligned}$$

Isto demostra que a forma de Jordan da restrición a F está contida na forma de Jordan de E (alternativamente, que no piso j da restrición a F hai como moito tantos vectores como na forma de Jordan de E).
Ademais, dada unha base de Jordan dun subespazo, esta pódese completar iterativamente a unha base de Jordan de todo o espazo. Para iso, imos ciclo a ciclo: se o ciclo k ten lonxitude s na restrición ao subespazo, miramos se o vector $v_{i,s}$ ten unha preimaxe por f_λ; se a ten, chamámoslle a ese vector $v_{i,s+1}$ e procedemos así mentres sexa posible atopar unha preimaxe. Iso dános uns cantos ciclos de Jordan completos, e a partir de aí procedemos do xeito habitual completando as bases de $\ker(f - \lambda\,\mathrm{Id})^j / \ker(f - \lambda\,\mathrm{Id})^{j-1}$ empezando polo piso superior para obter xeradores e producindo novos ciclos desa maneira. A posibilidade de completar pódese entender en termos da primeira observación da proba, pola cal o número de vectores nun piso concreto sempre é menor ou igual ao considerarmos a restrición ao subespazo. □

No caso no que $\mathrm{Min}(f;X) = p(X)q(X)$, con $\gcd(p(X), q(X)) = 1$, temos polo primeiro teorema de descomposición que

$$E = \ker(p(f)) \oplus \ker(q(f)).$$

Se $F \subset E$ é un subespazo invariante, temos que $\mathrm{Min}(f|F;X) = p'(X)q'(X)$, onde $p'(X) \mid p(X)$ e $q'(X) \mid q(X)$. Nese caso,

$$F = \ker(p'(f)|F) \oplus \ker(q'(f)|F).$$

Polo tanto, para achar os subespazos invariantes de E chega con calcular os de $\ker(p(f))$ e os de $\ker(q(f))$. Iso fai que a situación sexa especialmente sinxela cando hai un único bloque de Jordan para cada valor propio. Porén, cando hai máis dun, a situación vólvese máis complicada, como exploraremos nos exercicios.

Exemplo. Ao igual que antes, sexa f un endomorfismo tal que $\mathrm{Char}(f;X) = (X-2)^{18}$, con

- $\dim\ker(f - 2\,\mathrm{Id}) = 7$.
- $\dim\ker(f - 2\,\mathrm{Id})^2 = 13$.
- $\dim\ker(f - 2\,\mathrm{Id})^3 = 16$.
- $\dim\ker(f - 2\,\mathrm{Id})^4 = 18$.

Consideremos agora o subespazo que ten por descomposición en ciclos

$v_{1,3}$				
$v_{1,2}$	$v_{2,2}$	$v_{3,2}$		
$v_{1,1}$	$v_{2,1}$	$v_{3,1}$	$v_{4,1}$	$v_{5,1}$

Podemos completar cada un deses ciclos obtendo imaxes por f_λ mentres sexa posible, e chegamos a

$v_{1,4}$	$v_{2,4}$			
$v_{1,3}$	$v_{2,3}$	$v_{3,3}$		
$v_{1,2}$	$v_{2,2}$	$v_{3,2}$	$v_{4,2}$	$v_{5,2}$
$v_{1,1}$	$v_{2,1}$	$v_{3,1}$	$v_{4,1}$	$v_{5,1}$

Agora temos que $\ker(f-\lambda\,\mathrm{Id})^4/\ker(f-\lambda\,\mathrm{Id})^3$ xa ten dimensión 2, co que non hai que engadir novos vectores ao piso 4; o mesmo pasa co piso 3; en cambio $\ker(f-\lambda\,\mathrm{Id})^2/\ker(f-\lambda\,\mathrm{Id})$ ten dimensión 6, mentres que no piso 2 só hai 5 vectores. Completamos a base cun novo vector $v_{6,2}$, e definimos $v_{6,1} := f_\lambda(v_{6,2})$. Finalmente, como $\ker(f-\lambda\,\mathrm{Id})$ ten agora 6 vectores, completamos a base cun último vector $v_{7,1}$.

Exemplo. No caso do endomorfismo xa traballado e dado pola matriz

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

temos unicamente 6 espazos invariantes.

- O subespazo trivial que consta unicamente do $\{0\}$.
- Dous subespazos de dimensión 1, os xerados por cada un dos vectores propios.
- Dous subespazos de dimensión 2, o xerado polos dous vectores propios e o xerado polo ciclo de Jordan de tamño 2.
- O subespazo de dimensión 3 igual ao espazo total.

No caso do endomorfismo do exemplo anterior representado pola matriz

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

temos 12 subespazos invariantes que se obteñen a partir da forma de Jordan. Sexa $\{v_1, v_2, v_3, v_4, v_5\}$ a base de Jordan construída con anterioridade.

- O subespazo trivial que consta unicamente do $\{0\}$.
- Dous subespazos de dimensión 1, $\langle v_1\rangle$ e $\langle v_4\rangle$.
- Tres subespazos de dimensión 2, $\langle v_1, v_2\rangle$, $\langle v_4, v_5\rangle$ e $\langle v_1, v_4\rangle$.
- Tres subespazos de dimensión 3, $\langle v_1, v_4, v_5\rangle$, $\langle v_1, v_2, v_4\rangle$ e $\langle v_1, v_2, v_3\rangle$.
- Dous subespazos de dimensión 4, $\langle v_1, v_2, v_4, v_5\rangle$ e $\langle v_1, v_2, v_3, v_4\rangle$.
- O subespazo de dimensión 5 igual ao espazo total.

Porén, estes non son os únicos. Por exemplo, no caso de dimensión 1, podemos coller o xerado por calquera vector da forma $\lambda v_1 + \mu v_4$, donde $(\lambda, \mu) \neq (0, 0)$.

Nos exercicios, veremos como atopar todos os subespazos invariantes en situacións nas que hai máis dun ciclo de Jordan para un mesmo valor propio. Non presentamos ningún método xeral, senón que o máis doado pasa por unha análise individual de cada situación.
Outro exemplo habitual de aplicación da forma de Jordan é o cálculo de potencias n-ésimas. Para iso, como sucedía no caso da diagonalización, tense que é suficiente saber calcular as potencias dos bloques de Jordan. Neste caso, podemos ver, por exemplo por indución que

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^{n-1} \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}.$$

En xeral, se temos un bloque $r \times r$ na diagonal que está i posicións por riba da diagonal principal teremos $\binom{n}{i}\lambda^{n-i}$.

Resolución de recorrencias

Definición 3.12. Unha *recorrencia lineal homoxénea* é unha sucesión na que se dan os termos $a_0, a_1, \ldots, a_{k-1}$ e se cumpre unha ecuación da forma

$$a_{n+k} + \alpha_{k-1}a_{n+k-1} + \ldots + \alpha_1 a_{n+1} + \alpha_0 a_n = 0, \qquad \text{para todo } n \geq 0.$$

Podemos representar esta recorrencia mediante a igualdade matricial

$$\begin{pmatrix} a_{n+k} \\ a_{n+k-1} \\ a_{n+k-2} \\ \vdots \\ a_{n+1} \end{pmatrix} = \left(\begin{array}{cccc|c} -\alpha_{k-1} & -\alpha_{k-2} & \ldots & -\alpha_1 & -\alpha_0 \\ \hline 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{array}\right) \begin{pmatrix} a_{n+k-1} \\ a_{n+k-2} \\ a_{n+k-3} \\ \vdots \\ a_n \end{pmatrix}.$$

Sexa A a matriz de coeficientes. Procedendo como no caso da demostración de Cayley–Hamilton, tense que

$$\mathrm{Char}(A; X) = (-1)^n (X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0).$$

Polo tanto, ao facer A^n, a matriz de Jordan terá na diagonal as raíces do polinomio $X^n + a_{n-1}X^{n-1} + \ldots + a_0$, e o tamaño do maior bloque de Jordan dunha raíz λ estará limitado superiormente pola multiplicidade de λ como raíz do polinomio. Máis concretamente, tense que se las raíces do polinomio son $\lambda_1, \ldots, \lambda_r$ con multiplicidades $m_1, \ldots, m_r$, respectivamente, entón

$$a_n = (a_{1,1} + \ldots + a_{1,m_1}X^{m_1-1})\lambda_1^n + \ldots + (a_{r,1} + \ldots + a_{r,m_r}X^{m_r-1})\lambda_r^n.$$

Os coeficientes $a_{i,j}$ obtéñense ao impoñer as condicións iniciais.

Exemplo. Consideramos a recorrencia dada por $a_0 = 4$, $a_1 = 0$, $a_2 = 5$ e $a_{n+3} = 3a_{n+1} + 2a_n$ para todo $n \geq 3$. Entón, temos que

$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^n \begin{pmatrix} a_2 \\ a_1 \\ a_0 \end{pmatrix}.$$

Facendo a forma de Jordan da matriz obtense que

$$\begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 2 & 4 \\ 1 & -1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & -1 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \begin{pmatrix} 1/9 & 2/9 & -8/9 \\ 1/3 & -1/3 & -2/3 \\ 1/9 & 2/9 & 1/9 \end{pmatrix}.$$

Polo tanto,

$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} -1 & 2 & 4 \\ 1 & -1 & 2 \\ -1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cc|c} (-1)^n & n(-1)^{n-1} & 0 \\ 0 & (-1)^n & 0 \\ \hline 0 & 0 & 2^n \end{array}\right) \begin{pmatrix} 1/9 & 2/9 & -8/9 \\ 1/3 & -1/3 & -2/3 \\ 1/9 & 2/9 & 1/9 \end{pmatrix} \begin{pmatrix} 5 \\ 0 \\ 4 \end{pmatrix}.$$

Operando, vemos que

$$a_n = (-n+3)(-1)^n + 2^n.$$

Esta mesma técnica pódese empregar para resolver ecuacións non homoxéneas, nas que ademais dos termos lineais temos contribucións da forma $p_\lambda(n)\lambda^n$, onde $\lambda \in K$ é un escalar e $p_\lambda(X) \in K[X]$ é un polinomio. Imos explicar o procedemento cun exemplo.

Exemplo. Consideramos a recurrencia dada por $a_0 = 0$ e $a_{n+1} = a_n + n2^n$ para todo $n \geq 1$. Entón, temos que

$$\begin{pmatrix} a_{n+1} \\ 2^{n+1} \\ (n+1)2^{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} a_n \\ 2^n \\ n2^n \end{pmatrix}.$$

Polo tanto,

$$\begin{pmatrix} a_{n+1} \\ 2^{n+1} \\ (n+1)2^{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix}^n \begin{pmatrix} a_1 \\ 2^1 \\ 1 \cdot 2^1 \end{pmatrix}.$$

Facendo a forma de Jordan da matriz obtense

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix}^n = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix} \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right) \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Procedendo como no exemplo anterior obtemos que

$$a_n = n2^n - 2^{n+1} + 2.$$

A exponencial dunha matriz

Pasamos agora ao estudo das aplicacións á resolución de ecuacións diferenciais ordinarias, o que require definir a exponencial dunha matriz. No que resta de sección, sexa $K = \mathbb{R}$ ou $K = \mathbb{C}$, aínda que máis en xeral podemos traballar con calquera corpo no que teñamos unha topoloxía axeitada que nos permita falar de converxencia.

Definición 3.13. Sexa $A \in \mathcal{M}_n(K)$ unha matriz. A *exponencial* de A, que se denota como e^A ou $\exp(A)$, é a matriz de $\mathcal{M}_n(K)$ dada pola serie de potencias

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k,$$

onde A^0 é a matriz identidade. A serie sempre converxe, polo que a exponencial está ben definida.

Resulta doado comprobar que se $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, entón $e^A = \mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})$. No caso das matrices de Jordan, a situación é máis sutil. Por exemplo, sexa

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

un bloque de Jordan de valor propio λ e tamaño 4. Entón, tendo en conta que

$$J^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \binom{n}{3}\lambda^{n-3} \\ 0 & \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & 0 & \lambda^n \end{pmatrix},$$

tense que

$$e^J = e^\lambda \begin{pmatrix} 1 & 1 & \frac{1}{2!} & \frac{1}{3!} \\ 0 & 1 & 1 & \frac{1}{2!} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A aplicación principal do cálculo da matriz exponencial é a resolución de ecuacións diferenciais ordinarias. No caso de dimensión 1, a solución de

$$x' = \lambda x, \quad x(0) = x_0$$

é da forma $x(t) = x_0 e^{\lambda t}$. Isto mesmo segue a ser certo en dimensión arbitraria. Representamos por $\mathbf{x}$ o vector de solucións, que temos que interpretar como unha n-tupla da forma $(x_1(t), \ldots, x_n(t))$, e pomos $\mathbf{x}'$ para referirnos á súa derivada. Sexa $A \in \mathcal{M}_n(\mathbb{R})$. Dado un sistema de ecuacións diferenciais ordinarias da forma

$$\mathbf{x}' = A\mathbf{x}, \quad x(0) = x_0,$$

a súa única solución é $\mathbf{x}(t) = e^{tA}x_0$.

Exemplo. Consideramos o sistema de ecuacións

$$\begin{cases} x' &= x + y \\ y' &= y \\ z' &= -z \end{cases},$$

coas condicións iniciais $(x(0), y(0), z(0)) = (1, 1, 1)$. A matriz do sistema é

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

que xa está en forma de Jordan. Polo tanto, a solución do sistema é

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} e^t & te^t & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (t+1)e^t \\ e^t \\ e^{-t} \end{pmatrix}.$$

Forma de Jordan real

Imos considerar por último o caso no que $K = \mathbb{R}$ e o polinomio característico non descompón completamente. Sexa E un $\mathbb{R}$-espazo vectorial e $f \in \mathrm{End}(E)$. Nese caso, sabemos que se λ é un valor propio con multiplicidade alxébrica $m_a(\lambda)$, entón $\bar{\lambda}$ tamén é un valor propio e $m_a(\bar{\lambda}) = m_a(\lambda)$. Non só iso: se consideramos unha base formada por ciclos de Jordan para λ, podemos considerar os conxugados de cada vector e obter así unha base de ciclos de Jordan para $\bar{\lambda}$. Iso é unha consecuencia directa do feito de que o endomorfismo f está definido sobre $\mathbb{R}$, polo que $\overline{(f - \lambda\,\mathrm{Id})^k(v)} = (f - \bar{\lambda}\,\mathrm{Id})^k(\bar{v})$.

Proposición 3.18. Sexa $\{v_1, \ldots, v_r\}$ un ciclo de Jordan para $f \in \mathrm{End}(E)$ de valor propio λ. Tense que $\{\bar{v}_1, \ldots, \bar{v}_r\}$ é un ciclo de Jordan para $f \in \mathrm{End}(E)$ de valor propio $\bar{\lambda}$. Ademais, sexan $\Re(\lambda)$ e $\Im(\lambda)$ as partes real e imaxinaria de λ, respectivamente, $\Re(\lambda) = \frac{\lambda+\bar{\lambda}}{2}$ e $\Im(\lambda) = \frac{\lambda-\bar{\lambda}}{2i}$, e sexan

$$C_\lambda = \begin{pmatrix} \Re(\lambda) & \Im(\lambda) \\ -\Im(\lambda) & \Re(\lambda) \end{pmatrix}, \quad \mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad 0_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Para $i = 1, \ldots, r$, sexa $w_{2i-1} = \frac{v_i+\bar{v}_i}{2}$ e $w_{2i} = \frac{v_i-\bar{v}_i}{2i}$. Entón, a matriz de f na base $\{w_1, \ldots, w_{2r}$ é da forma

$$\begin{pmatrix} C_\lambda & \mathbb{I}_2 & \ldots & 0_2 & 0_2 \\ 0_2 & C_\lambda & \ldots & 0_2 & 0_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_2 & 0_2 & \ldots & C_\lambda & \mathbb{I}_2 \\ 0_2 & 0_2 & \ldots & 0_2 & C_\lambda \end{pmatrix}.$$

Demostración. Imos comezar calculando a imaxe de w_1 e w_2 por f. Para iso, observamos que $v_1 = w_1 + iw_2$ e $\bar{v}_1 = w_1 - iw_2$. Entón,

$$f(w_1) = f\Big(\frac{v_1 + \bar{v}_1}{2}\Big) = \frac{\lambda_1 v_1 + \bar{\lambda}_1 \bar{v}_1}{2} = \Re(\lambda)w_1 - \Im(\lambda)w_2$$
$$f(w_2) = f\Big(\frac{v_1 - \bar{v}_1}{2}\Big) = \frac{\lambda_1 v_1 - \bar{\lambda}_1 \bar{v}_1}{2} = \Im(\lambda)w_1 + \Re(\lambda)w_2.$$

Imos agora estudar a imaxe de w_{2i-1} e w_{2i}, con $2 \leq i \leq r$. Nese caso,

$$f(w_{2i-1}) = f\Big(\frac{v_i + \bar{v}_i}{2}\Big) = \Re(\lambda)w_{2i-1} - \Im(\lambda)w_{2i} + w_{2i-3}$$
$$f(w_{2i}) = f\Big(\frac{v_i - \bar{v}_i}{2i}\Big) = \Im(\lambda)w_{2i-1} + \Re(\lambda)w_{2i} + w_{2i-2}.$$

□

Exemplo. Consideramos a matriz con entradas reais

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

Os seus valores propios son $\lambda_1 = 1+i$ e $\lambda_2 = 1-i$. Os vectores propios correspondentes son $v_1 = (-2, 1+i)$ e $v_2 = (-2, 1-i)$. Polo resultado anterior, na base dada por $w_1 = (-2, 1)$ e $w_2 = (0, 1)$, a matriz de Jordan é

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Exemplo. Consideremos a matriz con entradas reais

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

Os valores propios son $\lambda_1 = i$ e $\lambda_2 = -i$, ambos con multiplicidade alxébrica 2 e multiplicidade xeométrica 1. Sexa (v_1, v_2) un ciclo de Jordan para i e sexa (v_3, v_4) o ciclo correspondente para $-i$. Collendo $v_1 = (i, 0, 1, 0)$, $v_2 = (0, 1+i, 1, 2i)$, $v_3 = (-i, 0, 1, 0)$ e $v_4 = (0, 1-i, 1, -2i)$ temos que a matriz nesa base é

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

Podemos polo tanto coller $w_1 = \Re(v_1) = (0, 0, 1, 0)$, $w_2 = \Im(v_1) = (1, 0, 0, 0)$, $w_3 = \Re(v_2) = (0, 1, 1, 0)$ e $w_4 = \Im(v_2) = (0, 1, 0, 2)$. Polo tanto, obtemos nesa base a forma de Jordan real:

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

3.5. Problemas

Vectores propios e diagonalización.

Problema 3.1. Achar os valores propios racionais, reais e complexos das seguintes matrices:

$$\begin{pmatrix} 3 & -13 & 10 \\ 1 & -3 & 10 \\ 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & -15 & -17 \\ 1 & -7 & -6 \\ -1 & 5 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & -7 & 1 \\ 4 & 0 & 1 \\ -4 & 28 & -4 \end{pmatrix}.$$

Solución. No primeiro caso, o polinomio característico é

$$(X^2 - 9)(-X + 3) + 10 + 3(3 - X) = -(X - 1)(X - 4)(X + 2),$$

polo que os valores propios son 1, 4 e -2, os tres racionais (e polo tanto tamén reais e complexos).
No segundo, o polinomio característico é

$$-X^3 - 3X^2 - 3X + 2 = -(X + 2)(X^2 + X + 1).$$

A única raíz racional e real e $X = -2$; por outro lado, o polinomio $X^2 + X + 1$ ten dúas raíces complexas, que son $X = \frac{-1 \pm \sqrt{-3}}{2}$. Polo tanto, a matriz ten tres valores propios complexos e unicamente un racional ou real.
No terceiro caso, o polinomio característico é

$$-X^3 - 3X^2 = -X^2(X + 3),$$

polo tanto os valores propios (xa sexa sobre os racionais, os reais ou complexos) son o -3 e o 0 (este último dobre).

Problema 3.2. Encontrar os valores e vectores propios das matrices seguintes e dicir se son ou non diagonalizables.

$$\begin{pmatrix} 2 & 0 & 0 \\ -3 & -1 & 3 \\ 3 & 3 & -1 \end{pmatrix}, \quad \begin{pmatrix} -2 & 0 & 1 \\ 4 & 2 & -1 \\ 1 & 0 & -2 \end{pmatrix}, \quad \begin{pmatrix} -2 & 20 & 4 \\ 0 & -3 & 0 \\ -1 & 7 & 2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 2 & 0 \\ -1 & 0 & 1 \\ 0 & -2 & 0 \end{pmatrix}.$$

Solución. O polinomio característico da primeira matriz, á que chamaremos A, é $\operatorname{Char}(A;X) = -(X-2)^2(X+4)$. O subespazo de vectores propios correspondente ao valor propio -4 é $\langle(0,1,-1)\rangle$; o correspondente ao valor propio 2 é $\langle(1,0,1),(0,1,1)\rangle$. Como as dúas multiplicidades alxébricas coinciden coas xeométricas, a matriz diagonaliza.
Para a segunda matriz, á que chamamos B, temos que $\operatorname{Char}(B;X) = -X^3+12X-16 = -(X+3)(X+1)(X-2)$. Como todos os valores propios teñen multiplicidade alxébrica 1, a matriz diagonaliza. O subespazo de vectores propios correspondente a $X=-3$ é $\langle(1,-1,-1)\rangle$; o correspondente a $X=-1$ é $\langle(1,-1,1)\rangle$; e o correspondente a $X=2$ é $\langle(0,1,0)\rangle$.
Se C é a terceira matriz tense que $\operatorname{Char}(C;X) = -X^3-3X^2 = -X^2(X+3)$. A multiplicidade xeométrica do valor propio 0 é 1, e o seu subespazo propio é o xerado polo vector $(2,0,1)$; polo tanto, a matriz non é diagonalizable. O subespazo de vectores propios correspondente a $X=-3$ é $\langle(8,-1,3)\rangle$.
Finalmente, se D é a cuarta matriz, $\operatorname{Char}(D;X) = -X^3-4X = -X(X^2+4)$. Os valores propios son 0 e as raíces de $X^2+4=0$, en caso de existir. Polo tanto, para que a matriz diagonalice cómpre que -4 teña raíz cadrada. En particular, D non diagonaliza nin sobre os racionais nin sobre os reais. O subespazo de vectores propios asociado a $X=0$ é $\langle(1,0,1)\rangle$. Sobre $\mathbb{C}$, temos que o subespazo de vectores propios asociado a $X=2i$ (resp. $X=-2i$) é $\langle(-1,-i,1)\rangle$ (resp. $\langle -1,i,1\rangle$).

Problema 3.3. Encontrar a matriz na base canónica dun endomorfismo $f \in \operatorname{End}(\mathbb{R}^3)$ tal que:

(i) $f(e_1)+f(e_3) = e_1+e_3$.

(ii) $f(e_1)+f(e_2) = e_1+e_2$.

(iii) Calquera vector non nulo na intersección $U\cap V$, onde $U = \{(x,y,z)\in\mathbb{R}^3 \mid z+y-2x=0\}$ e $V = \langle(2,0,3),(1,0,1)\rangle$, é un vector propio de valor propio -1.

Solución. O subespazo V pode escribirse como $V = \{(x,y,z)\in\mathbb{R}^3 \mid y=0\}$, co cal

$$U\cap V = \langle(1,0,2)\rangle.$$

Polo tanto, as condicións do enunciado afirman que $(1,0,1)$, $(1,1,0)$ e $(1,0,2)$ son vectores propios con valores propios 1, 1 e -1, respectivamente. Iso quere dicir que nesa base a matriz é diagonal, cos correspondentes valores propios na diagonal. A matriz na base canónica é simplemente

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

Problema 3.4. Sabendo que $(1,2)$ e $(1,3)$ son vectores propios dunha matriz real $A\in\mathcal{M}_2(\mathbb{R})$ de valores propios -1 e 3, respectivamente, calcular A^n para $n\geq 1$.

Solución. Usando as condicións do enunciado, podemos escribir

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

Elevando agora á potencia n, temos que

$$A^n = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} (-1)^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 \cdot 3^n + 3 \cdot (-1)^n & 3^n - (-1)^n \\ -6 \cdot 3^n + 6 \cdot (-1)^n & 3 \cdot 3^n - 2(-1)^n \end{pmatrix}.$$

Problema 3.5. Demostrar que os endomorfismos $f, g : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathcal{M}_2(\mathbb{R})$ definidos por

$$f(A) = A^t \quad \text{e} \quad g(A) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} A$$

son ambos diagonalizables e encontrar unha base de vectores propios para cada un.

Solución. Consideremos a base habitual con

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Entón, a matriz de f na base canónica vén dada por

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

que ten polinomio característico $(X-1)^3(X+1)$. Só temos que comprobar que a multiplicidade xeométrica do vector propio 1 é 3; iso é equivalente a ver que $M - \mathbb{I}$ ten rango 1; pero iso é evidente xa que

$$M - \mathbb{I} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(dúas filas iguais a cero e a terceira é igual á segunda cambiada de signo). Alternativamente, podemos ver que o polinomio mínimo non ten raíces repetidas, é dicir, é igual a $X^2 - 1$; é unha comprobación totalmente rutineira ver que $M^2 = \mathbb{I}$.
No caso de g, a matriz na base canónica é

$$N = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

cuxo polinomio característico é $X^4 - 2X^2 + 1 = (X^2-1)^2 = (X-1)^2(X+1)^2$. Temos de novo dúas maneiras de proceder. A primeira delas é vendo que tanto 1 como -1 teñen multiplicidade xeométrica 2, ou o que é o mesmo, que

$$N - \mathbb{I} = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \quad N + \mathbb{I} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

teñen ambas rango dous. En ambos casos é obvio, por ser a primeira fila un múltiplo da terceira e a segunda da cuarta. Alternativamente, podemos observar que $N^2 = \mathbb{I}$ e que polo tanto o polinomio mínimo é $X^2 - 1$.

Problema 3.6. Determinar exemplos de endomorfismos f de $\mathbb{R}^3$ que verifiquen as condicións que se especifican e, en cada caso, calcular o seu polinomio característico.

(a) Que teña os vectores propios $(1, 0, 0)$, $(1, 1, 0)$ e $(1, 0, 1)$ con valores propios 2, 1 e -1, respectivamente.

(b) Que teña núcleo $\ker(f) = \langle(0, 0, 1)\rangle$ e tal que $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = z\}$ sexa o subespazo xerado polos vectores propios de valor propio 2.

(c) Non diagonalizable e tal que $(1, 0, 0)$ e $(0, -1, 1)$ sexan vectores propios de valor propio -1.

Solución. (a) Usando as condicións do enunciado, a matriz na base canónica é

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

O polinomio característico é $\operatorname{Char}(A; X) - (X - 2)(X - 1)(X + 1)$.

(b) Aplicando as condicións do enunciado, resulta que a matriz buscada é

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

O polinomio característico é $\operatorname{Char}(B; X) = -X(X - 2)^2$.

(c) Completamos a base que contén os dous vectores do enunciado co $(0, 0, 1)$. Nesta base, a matriz é da forma

$$\begin{pmatrix} -1 & 0 & a \\ 0 & -1 & b \\ 0 & 0 & -1 \end{pmatrix},$$

e a condición para que a matriz non diagonalice é que a e b non sexan ambos cero, xa que é a única situación na que a multiplicidade xeométrica do -1 sería igual a 3. Na base canónica, a matriz é

$$C = \begin{pmatrix} -1 & a & a \\ 0 & b - 1 & b \\ 0 & 0 & -1 \end{pmatrix}.$$

O polinomio característico é $\operatorname{Char}(C; X) = -(X + 1)^3$.

Problema 3.7. Consideramos o endomorfismo $f : \mathbb{C}_3[t] \to \mathbb{C}_3[t]$ dado por

$$f(p(t)) = p(t) + p(1)(t - 3) - 2p'(1)(t - 1)$$

Encontrar os valores propios e os vectores propios de f e discutir se o endomorfismo diagonaliza.

Solución. Na base $\{1, t, t^2, t^3\}$ a matriz do endomorfismo é

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

O polinomio característico é $(X-1)^2(X+1)^2$, polo que para ver que diagonalice temos que achar a multiplicidade xeométrica dos valores propios 1 e -1.

- Comezamos con $\lambda = 1$. Nese caso, a multiplicidade xeométrica é 2 e unha base do subespazo propio está dada por $v_1 = (2, -3, 0, 1)$ e $v_2 = (1, -2, 1, 0)$.
- No caso de $\lambda = -1$, a multiplicidade xeométrica é 1 e o único vector propio é $(-1, 1, 0, 0)$.

Polo tanto, o endomorfismo f **non** diagonaliza.

Problema 3.8. Encontrar a forma diagonal dun endomorfismo $f \in \mathrm{End}(\mathbb{R}^3)$ que cumpre as seguintes catro condicións.

(i) $f(e_1) + f(e_2) + f(e_3) = (1, 1, 1)$.

(ii) $f(e_1) - f(e_2) = f(1, 1, -2)$.

(iii) $f^2 = f$.

(iv) O subespazo $\{(x, y, z) \in \mathbb{R}^3 \mid y + 2z = 0\} \cap \langle (1, 1, 0), (1, 0, 1) \rangle$ é invariante por f.

Solución. Da terceira condición sabemos que f diagonaliza e que os únicos valores propios posibles son 0 e 1. A primeira condición dinos que $v_1 = (1, 1, 1)$ é un vector propio de valor propio 1. A segunda, que $f(0, 1, -1) = 0$, polo que $v_2 = (0, 1, -1)$ é un vector propio de valor propio 0. Por último, a cuarta condición afirma que $v_3 = (1, 2, -1)$ é tamén un vector propio, que pode ter valor propio 0 ou 1.
Polo tanto, $\{v_1, v_2, v_3\}$ é unha base de vectores propios, é os valores propios correspondentes son $(1, 0, a)$, onde $a \in \{0, 1\}$.

Problema 3.9. Sexa $A \in \mathcal{M}_2(\mathbb{R})$ unha matriz con traza t. Demostrar que se t é un valor propio de A, entón $A^m = t^{m-1}A$ para todo natural $m \geq 1$.

Solución. O polinomio característico dunha matriz $A \in \mathcal{M}^2(\mathbb{R})$ é $X^2 - \mathrm{Tr}(A)X + \det(A)$. Se ten unha raíz real, ambas o son, e o polinomio descompón como $(X-t)(X-u)$. Como $t = \mathrm{Tr}(A)$ e $t + u = t$, ten que ser $u = 0$.
Se $t = u = 0$, o polinomio característico é X^2. Consideramos unha base na cal o primeiro vector é un vector propio de valor propio 0; facendo o cambio de base, temos que existe unha matriz invertible P e un real a de maneira que

$$P^{-1}AP = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}.$$

Como $\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}^2 = 0$, entón $A^m = 0$ para calquera $m \geq 2$ e a igualdade do enunciado é sempre certa.

Se $t \neq 0$, entón a matriz diagonaliza e nunha base de vectores propios temos que

$$P^{-1}AP = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}.$$

Polo tanto, da igualdade

$$P^{-1}A^mP = \begin{pmatrix} t^m & 0 \\ 0 & 0 \end{pmatrix} = t^{m-1}\begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$$

séguese que $A^m = t^{m-1}A$.

Problema 3.10. Sexa $\mathcal{B}_e = \{e_1, \ldots, e_n\}$ unha base dun espazo vectorial E e sexa $f \in \operatorname{End}(E)$ un endormofismo que non é identicamente cero de maneira que

$$f(e_1) = \ldots = f(e_n) = \sum_{i=1}^{n} a_i e_i.$$

Demostrar que f é diagonalizable se e soamente se $\sum_{i=1}^{n} a_i \neq 0$.

Solución. Consideramos a base $\mathcal{B}_v = \{v_1, \ldots, v_n\}$, con $v_1 = e_1$ e $v_i = e_i - e_1$, para todo $i = 2, \ldots, n$. Temos que

$$f(v_2) = \ldots = f(v_n) = 0,$$

mentres que

$$f(v_1) = \sum_{i=1}^{n} a_i e_i = \sum_{i=2}^{n} a_i(e_i - e_1) + \Big(\sum_{i=1}^{n} a_i\Big)e_i = \Big(\sum_{i=1}^{n} a_i\Big)v_i + \sum_{i=2}^{n} a_i v_i.$$

Polo tanto, a matriz de f na base $\mathcal{B}_v$ é

$$\begin{pmatrix} \sum_{i=1}^{n} a_i & 0 & \ldots & 0 \\ a_2 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ a_n & 0 & \ldots & 0 \end{pmatrix}.$$

O polinomio característico de f é

$$\operatorname{Char}(f; X) = (-1)^n X^{n-1}(X - \sum_{i=1}^{n} a_i),$$

polo que a multiplicidade alxébrica de 0 é

$$m_a(0) = \begin{cases} n-1 & \text{se } \sum_{i=1}^{n} a_i \neq 0 \\ n & \text{en caso contrario.} \end{cases}$$

A multiplicidade xeométrica sempre é $m_x(0) = n - 1$, xa que a matriz ten rango 1, por ter $n - 1$ columnas identicamente nulas e outra que é diferente de cero, xa que estamos supondo que o endomorfismo non é nulo. Polo tanto, o endomorfismo diagonaliza se, e soamente se, $m_a(0) = m_x(0)$, e iso é equivalente a $\sum_{i=1}^{n} a_i \neq 0$.
Alternativamente, podemos traballar na base canónica e observar que a multiplicidade xeométrica do 0 é sempre $n - 1$, polo que a alxébrica só pode ser $n - 1$ ou n, e será un valor ou outro dependendo da traza da matriz.

Polinomio mínimo e teorema de Cayley–Hamilton.

Problema 3.11. Determinar os endomorfismos f de $\mathbb{R}^3$ que teñan un único valor propio real con multiplicidade alxébrica 1, núcleo $\ker(f) = \langle(0,0,1)\rangle$ e tal que $W = \{(x,y,z) \in \mathbb{R}^3 \mid x = y + z\}$ sexa un subespazo vectorial invariante.

Solución. Consideremos unha base formada por un elemento do núcleo e por dous vectores de W, por exemplo $\{(0,0,1),(1,0,1),(1,1,0)\}$. Nesa base, a matriz é

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix},$$

e a condición de que teña un único valor propio real é que $(a-d)^2 + 4bc < 0$. Polo tanto, os endomorfismos son todos os da forma

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ a+c & b+d & 0 \end{pmatrix};$$

como queremos que só haxa un valor propio e o polinomio característico é $-X(X^2 - (a+d)X + ad - bc)$, tense que cumprir ademais que $(a-d)^2 + 4bc < 0$.

Problema 3.12. Sexa f o endomorfismo de $\mathbb{R}^3$ definido por

$$f(x,y,z) = (x+y, x+ay-z, -x+y-bz).$$

(a) Determinar os valores de a e b para os cales o vector $(1,1,0)$ é un vector propio de f e o subespazo $\{(x,y,z) \in \mathbb{R}^3 \mid y + z = 0\}$ é un subespazo invariante por f.

(b) Determinar os valores de a e b para os cales 1 é un valor propio de f. Para que valores 1 é valor propio de f con multiplicidade alxébrica 2?

Solución. (a) Como $f(1,1,0) = (2, 1+a, 0)$, o vector $(1,1,0)$ é un vector propio se e soamente se $a = 1$. Unha base do subespazo descrito no enunciado está formada polos vectores $(1,0,0)$ e $(0,1,-1)$; chega con impor que a imaxe dos vectores da base pertenza tamén ao subespazo. Por un lado, $f(1,0,0) = (1,1,-1)$ sempre pertence; por outro, $f(0,1,-1) = (1, a+1, 1+b)$, e tense que cumprir que $a + b + 2 = 0$. Isto quere dicir que $b = -3$.

(b) A condición de que 1 sexa valor propio quere dicir que $f - \mathrm{Id}$ ten un núcleo non trivial. A matriz na base canónica é

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & a-1 & -1 \\ -1 & 1 & -b-1 \end{pmatrix}.$$

O seu determinante é $b+2$, polo que 1 é valor propio se, e soamente se, $b = -2$.

Nese caso, temos que o polinomio característico é $-(2-X)(1-X)(a-X)$. A multiplicidade alxébrica de 1 será 2 se, e soamente se, $a = 1$.

Problema 3.13. Sexa $f \in \mathrm{End}(E)$ e sexa $F \subset E$ un subespazo invariante. Sexa $f|F \in \mathrm{End}(F)$ a restrición do endomorfismo ao subespazo F. Demostrar que o polinomio característico $\mathrm{Char}(f|F; X)$ divide ao polinomio característico $\mathrm{Char}(f; X)$.

Solución. Consideramos unha base de F e ampliámola a unha base de todo o espazo E. Como F é un subespazo invariante, $f(F) \subset F$, e a matriz de f con respecto a esta base é da forma $\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$, onde A é un bloque cadrado de tamaño igual á dimensión de F e D é outro bloque cadrado de dimensión igual á dun complementario de F. Polo tanto, $\operatorname{Char}(f;X) = \operatorname{Char}(A;X) \cdot \operatorname{Char}(D;X)$, e como o polinomio característico non depende da base escollida, $\operatorname{Char}(A;X) = \operatorname{Char}(f|F;X)$. Isto demostra que $\operatorname{Char}(f|F;X)$ divide $\operatorname{Char}(f;X)$.

Problema 3.14. Sexa $f \in \operatorname{End}(E)$ e sexa $F \subset E$ un subespazo invariante. Sexa $G = E/F$ o espazo vectorial cociente correspondente. Comprobar que a aplicación $\tilde{f}_F : G \to G$ definida por $\tilde{f}_F([v]) = [f(v)]$ está ben definida e é un endomorfismo de G. Demostrar que

$$\operatorname{Char}(f;X) = \operatorname{Char}(f|F;X)\operatorname{Char}(\tilde{f}_F;X).$$

Solución. En primeiro lugar, observamos que $f(F) \subset F$. Para ver que $\tilde{f}_F$ está ben definida, temos que ver que non depende da elección do representante, é dicir, se $v - v' \in F$, entón $f(v) - f(v') \in F$. Iso é así porque $f(v - v') \in F$ xa que F é invariante por definición. Por outro lado, é inmediato comprobar que $\tilde{f}_F$ é lineal, polo que é un endomorfismo. Consideramos unha base de F, $\{v_1, \ldots, v_r\}$ e ampliámola a unha base de todo o espazo E, $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$. Como F é un subespazo invariante, $f(F) \subset F$, e a matriz de f con respecto a esta base é da forma $\left(\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right)$, onde A é un bloque cadrado de tamaño igual á dimensión de F e D é outro bloque cadrado de dimensión igual á dun complementario de F. Agora ben, observamos que as clases de $\{v_{r+1}, \ldots, v_n\}$ son unha base de G e polo tanto D é unha matriz de $\tilde{f}_F$ nesa base. Iso é así porque, chamando x_{ij} ás columnas da matriz, temos que, para calquera $r + 1 \leq k \leq n$,

$$f(v_k) = \sum_{i=1}^{r} x_{ik} v_i + \sum_{i=r+1}^{n} x_{ik} v_i;$$

como a primeira suma está en F, tense que

$$\tilde{f}_F([v_k]) = \sum_{i=r+1}^{n} x_{ik}[v_i].$$

Temos entón que

$$\operatorname{Char}(f;X) = \operatorname{Char}(A;X)\operatorname{Char}(D;X) = \operatorname{Char}(f|F;X)\operatorname{Char}(\tilde{f}_F;X),$$

como queriamos ver.

Problema 3.15. Sexa f un endomorfismo dun $\mathbb{R}$-espazo vectorial E de dimensión finita. Demostrar que o endomorfismo f ten como mínimo un valor propio se e soamente se existe como mínimo un subespazo invariante por f de dimensión impar.

Solución. Para resolver este problema, imos usar que todo polinomio de $\mathbb{R}[X]$ de grao impar ten polo menos unha raíz; iso pode verse usando que as raíces complexas que non son reais veñen por pares (é dicir, se α é raíz, o seu conxugado $\bar{\alpha}$) ou aplicando o teorema de Bolzano. Polo tanto, se f ten algún valor propio ten algún vector propio e, polo tanto, ten rectas invariantes, que son subespazos de dimensión 1.

Sexa agora F un subespazo de dimensión impar invariante por f. Nunha base obtida ao ampliar unha base de F a todo o espazo E, a matriz de f é unha matriz por bloques da forma $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$, onde A é a matriz de $f|F$ na base F. Polo tanto, o polinomio característico de f é o produto dos polinomios característicos das matrices A e D. O polinomio característico de A ten grao impar e, polo tanto, ten algunha raíz real, que é un valor propio do endomorfismo f.

Problema 3.16. Sexa E un espazo vectorial de dimensión finita e sexa $f : E \longrightarrow E$ un endomorfismo que cumpre $f^2 = f$.

(a) Demostrar que f diagonaliza.

(b) Demostrar que $E = \ker(f) \oplus \mathrm{im}(f)$.

(c) Dar unha fórmula para expresar un vector v de E como suma dun do núcleo e outro da imaxe.

Solución. (a) Temos que o polinomio mínimo é un divisor de $X^2 - X = X(X-1)$, que descompón completamente e non ten raíces repetidas. Polo tanto, diagonaliza.

(b) Se $X^2 - X$ é o polinomio mínimo, polo primeiro teorema de descomposición, $E = \ker(f) \oplus \ker(f - \mathrm{Id})$. Agora ben, temos que $\ker(f - \mathrm{Id}) = \mathrm{im}(f)$. Para ver iso, se $v = f(u)$, entón

$$(f - \mathrm{Id})(v) = (f - \mathrm{Id})(f(u)) = (f^2 - f)(u) = 0;$$

reciprocamente, se $(f - \mathrm{Id})(u) = 0$, entón $u = f(u) \in \mathrm{im}(f)$. Polo tanto, $E = \ker(f) \oplus \mathrm{im}(f)$. Se o polinomio mínimo é un dos divisores de $X^2 - X$ chégase á mesma conclusión analizando cada un dos casos.

(c) É suficiente con observar que $v = (v - f(v)) + f(v)$, onde $v - f(v) \in \ker(f)$ e $f(v) \in \mathrm{im}(f)$.

Problema 3.17. Probar que unha matriz cadrada A de orde n con entradas reais e diagonalizable sobre os números complexos en calquera dos seguintes casos.

(a) Se ten inversa e esta cumpre que $A^{-1} = 2\mathbb{I}_n - A^3/2$.

(b) Se as matrices A^2 e $A^3 + A - \mathbb{I}_n$ son inversas unha da outra.

(c) Se o seu polinomio característico é igual a $-X^5 - X^3 + X^2 + 1$.

Solución. En primeiro lugar, observamos que sobre os complexos calquera polinomio descompón completamente, polo que é suficiente comprobar que o polinomio mínimo non ten raíces repetidas. Se un polinomio anulador non as ten, entón como o mínimo é un divisor, tampouco as terá.

(a) Temos que $A^4 - 4A + 2 = 0$, polo que $X^4 - 4X + 2$ é un polinomio anulador. Se tivera raíces dobres, estas serían tamén raíces da súa derivada, que é $4(X^3 - 1)$. É inmediato ver que este polinomio ten por ceros $1, \frac{1 \pm \sqrt{-3}}{2}$, é ningún deles anula $X^4 - 4X + 2$.

(b) Temos que $A^2(A^3 + A - \mathbb{I}_n) = \mathbb{I}_n$, polo que $X^5 + X^3 - X^2 - 1$ é anulador. Observamos que

$$X^5 + X^3 - X^2 - 1 = (X - 1)(X^4 + X^3 + 2X^2 + X + 1),$$

co cal hai que establecer que $X^4 + X^3 + 2X^2 + X + 1$ non ten raíces dobres. Hai agora varias maneiras de proceder. A primeira é observar que

$$X^4 + X^3 + 2X^2 + X + 1 = (X^2 + 1)(X^2 + X + 1),$$

e observar que as raíces son polo tanto $\pm i$, $\frac{1 \pm \sqrt{-3}}{2}$. A segunda é notar que se un polinomio ten raíces dobres, entón tamén son raíces da súa derivada, e usando o algoritmo de Euclides podemos ver que

$$\gcd(X^4 + X^3 + 2X^2 + X + 1, 4X^3 + 3X^2 + 4X + 1) = 1.$$

A terceira é observar que $P(X) = X^4 + X^3 + 2X^2 + X + 1$ é un polinomio *palindrómico*, é dicir, que os coeficientes son simétricos. Nestes casos podemos considerar a variable $Y = X + X^{-1}$ e escribir $P(Y) = Y^2 + Y$. As raíces son $Y = 0, 1$ e os valores de X son por tanto as raíces de $X + X^{-1} = 0$ e de $X + X^{-1} = 1$. É dicir, as raíces do polinomio son $\pm i$, $\frac{1 \pm \sqrt{-3}}{2}$.

(c) Polo teorema de Cayley–Hamilton sabemos que o polinomio característico é anulador, e vimos no apartado anterior que ese polinomio non ten raíces dobres.

Problema 3.18. Unha matriz $A \in \mathcal{M}_n(K)$ dise que é nilpotente se existe un enteiro $k \geq 1$ tal que $A^k = 0$. Sexa A unha matriz nilpotente.

(a) Demostrar que A é conxugada dalgunha matriz triangular superior estrita (é dicir, unha matriz (u_{ij}) con entradas $u_{ij} = 0$ se $i \geq j$).

(b) Demostrar que $A^n = 0$.

(c) Encontrar o polinomio característico da matriz A e, usando este, encontrar o da matriz $A + \mathbb{I}_n$.

(d) Calcular $\det(A + \mathbb{I}_n)$.

(e) Demostrar que $\det(A + B) = \det(B)$ para toda matriz $B \in \mathcal{M}_n(K)$ que conmute con A.

Solución. (a) É obvio que a condición de ser nilpotente depende da clase de conxugación e que todas as matrices nilpotentes teñen determinante 0 (de non ser así, as súas potencias serían invertibles). Procedemos entón por indución sobre o tamaño n da matriz. O caso $n = 1$ é trivial. Como o determinante é 0, existe algún vector non nulo $x \in K^n$ tal que $Ax = 0$. Completándoo a unha base de K^n, obtense unha matriz $P \in \mathrm{GL}_n(K)$ tal que $B = P^{-1}AP$ é unha matriz nilpotente coa primeira columna igual a 0. Entón,

$$B = \begin{pmatrix} 0 & b \\ 0 & C \end{pmatrix}, \quad B^k = \begin{pmatrix} 0 & bC^{k-1} \\ 0 & C^k \end{pmatrix},$$

onde $b \in \mathcal{M}_{1\times(n-1)}$ e $C \in \mathcal{M}_{n-1}$. Polo tanto, C é tamén nilpotente e pola hipótese de indución ten unha conxugada $Q^{-1}CQ$ que é triangular superior estrita. Polo tanto, conxugando B pola matriz $Q = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$ obtense unha matriz

$$Q^{-1}BQ = \begin{pmatrix} 1 & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} 0 & b \\ 0 & C \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} = \begin{pmatrix} 0 & b \\ 0 & R^{-1}BR \end{pmatrix}$$

que é triangular superior estrita.

(b) Unha matriz triangular superior estrita de tamaño n dá 0 ao multiplicala por si mesma n veces.

(c) Se usamos a matriz triangular superior estrita para calcular o polinomio característico temos que $\mathrm{Char}(A;X) = (-X)^n$. Entón,

$$\mathrm{Char}(A + \mathbb{I}_n; X) = \mathrm{Char}(A; X-1) = (1-X)^n.$$

(d) O determinante obtense avaliando o polinomio característico en $X = 0$, e polo tanto é igual a 1.

(e) Supoñamos en primeiro lugar que B ten inversa. Entón, AB^{-1} é nilpotente, xa que A e B^{-1} tamén conmutan, e polo tanto $(AB^{-1})^k = 0$. Do apartado anterior dedúcese que $\det(AB^{-1} + \mathbb{I}_n) = 1$. Entón, $\det(B) = \det(AB^{-1} + \mathbb{I}_n)\det(B) = \det(A+B)$. Se B non ten inversa e o seu determinante é 0. Se $A^k = 0$, tense que

$$(A+B)^k = \sum_{i=0}^{k} \binom{k}{i} A^{k-i}B^i = B\Big(\sum_{i=1}^{k} A^{k-i}B^{i-1}\Big),$$

e tomando determinantes dedúcese que $\det(A+B)^k = \det(B)\det(\sum \cdots) = 0$. Polo tanto, $\det(A+B) = 0$.

Problema 3.19. Sexa A unha matriz cadrada real de orden n de maneira que $A^5 = \mathbb{I}_n$. Probar que se A diagonaliza sobre $\mathbb{R}$, entón $A = \mathbb{I}_n$.

Solución. Temos que $X^5 - 1$ é un polinomio anulador da matriz. A condición necesaria é suficiente para que A diagonalice é que o seu polinomio descompoña completamente en factores lineais e ningún estea repetido. Neste caso, sabemos que $X^5 - 1$ non ten raíces múltiples, e os seus ceros son as raíces quintas da unidade. Todas elas, salvo 1, son raíces complexas. Polo tanto, non poden ser factores do seu polinomio mínimo, que ten que ser entón $X - 1$. Iso implica que $A = \mathbb{I}_n$

Problema 3.20. Sexan $f, g \in \mathrm{End}(E)$ endomorfismos dun espazo vectorial de dimensión finita. Demostrar que se v é un vector propio de $g \circ f$ de valor propio $\lambda \neq 0$, entón $f(v)$ é un vector propio de $f \circ g$ de valor propio λ. Que pasa se $\lambda = 0$? Demostrar que se $g \circ f$ é invertible e diagonalizble, entón $f \circ g$ tamén é diagonalizable.

Solución. Se $(g \circ f)(v) = \lambda v$, entón $(f \circ g)(f(v)) = f((g \circ f)(v)) = f(\lambda v) = \lambda f(v)$. Como $v \neq 0$ por ser un vector propio, $\lambda v \neq 0$, e polo tanto a identidade $g(f(v)) = \lambda v$ é suficiente para asegurar que $f(v) \neq 0$, e polo tanto $f(v)$ é un vector propio de valor propio λ.
Se $\lambda = 0$ poden pasar dúas cousas: se $f(v) \neq 0$, este vector é vector propio de $f \circ g$ de valor propio 0; en cambio, se $f(v) = 0$ entón xa non é vector propio.

Se $g \circ f$ é invertible, entón todos os seus valores propios son diferentes de cero (un endomorfismo ten o cero como valor propio se, e soamente se, ten núcleo diferente de $\{0\}$, que equivale a dicir que non é invertible). Como endomorfismo bixectivo equivale a inxectivo e sobrexectivo, tense que os dous endomorfismos f e g son bixectivos. Polo tanto, se $v_1, \dots, v_n$ é unha base de vectores propios de $g \circ f$, entón $f(v_1), \dots, f(v_n)$ é unha base de vectores propios de $f \circ g$.

Problema 3.21. Sexa $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.

(a) Calcular a matriz da aplicación lineal $f_A : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathcal{M}_2(\mathbb{R})$ dada por $X \mapsto AX - XA$. Cal é o seu rango?

(b) Cal é o polinomio mínimo de f_A?

(c) Se $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$, e $a = d$, $a' = d'$ e $b = c' = 0$, demostrar que os núcleos de f_A e f_B teñen unha recta común.

(d) Dar unha condición necesaria e suficiente para que $\ker(f_A) + \ker(f_B) = \mathcal{M}_2(\mathbb{R})$.

Solución. (a) Consideremos a base habitual con

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Entón, un cálculo doado mostra que a matriz de f_A é

$$\begin{pmatrix} 0 & -c & b & 0 \\ -b & a-d & 0 & b \\ c & 0 & d-a & -c \\ 0 & c & -b & 0 \end{pmatrix}.$$

Chamemos v_1, v_2, v_3, v_4 ás columnas da matriz. Sempre temos que $v_4 = -v_2$. Se $c \neq 0$, $v_3 = \frac{d-a}{c} v_1 - \frac{b}{c} v_2$; se $b \neq 0$, entón $v_2 = \frac{d-a}{b} v_1 - \frac{c}{b} v_3$; se $b = c = 0$, $v_1 = v_4 = 0$. Isto amosa que o rango é como moito 2.

Por outro lado, se b e c non son ambos cero, a primeira columna é non cero e é linealmente independente coa segunda ou coa terceira. Se $b = c = 0$, entón o rango é 2 se $a - d \neq 0$ e 0 se $a = d$.

Polo tanto, concluímos que o rango é 0 se $b = c = a - d = 0$ e 2 no resto dos casos.

(b) Podemos calcular o polinomio característico, e vemos que

$$\mathrm{Char}(f_A; X) = X^4 + X^2(-a^2 - d^2 + 2ad - 4bc) = X^2(X^2 + (-a^2 - d^2 + 2ad - 4bc)).$$

- Supoñamos que $2ad - a^2 - d^2 - 4bc \neq 0$, que en particular implica que non estamos na situación $b = c = a - d = 0$. Polo tanto, o factor cuadrático do polinomio ten dúas raíces diferentes e distintas de cero. A multiplicidade xeométrica de 0 é 2, que coincide coa alxébrica, polo tanto o endomorfismo diagonaliza. O polinomio mínimo é

$$X(X^2 + (-a^2 - d^2 + 2ad - 4bc)),$$

co cal o seu grao é 3. Neste caso, a dimensión do núcleo é 2 e o grao do polinomio mínimo é 3.

- Se $a = d = b - c$, temos o endomorfismo 0, polo que o polinomio mínimo é X. Neste caso, a dimensión do núcleo é 4 e o grao do polinomio mínimo é 1.
- Queda por analizar o caso no que o endomorfismo non é cero, pero $-a^2 - d^2 + 2ad - 4bc = 0$. Entón, o núcleo ten dimensión 2, e unha comprobación rutineira amosa que o cadrado da matriz nunca é 0; polo tanto, o grao do polinomio mínimo é 3. Novamente, a dimensión do núcleo é 2 e o grao do polinomio mínimo é 3.

(c) Se $a = d$ e $b = 0$, o núcleo de f_A está xerado polos vectores $(1, 0, 0, -1)$ e $(0, 0, 1, 0)$. Se $a' = d'$ e $c' = 0$, o núcleo de f_B está xerado polos vectores $(1, 0, 0, -1)$ e $(0, 1, 0, 0)$. Polo tanto, a intersección é un subespazo de dimensión 1 xerado por $(1, 0, 0, 1)$.

(d) Se un dos núcleos de f_A ou f_B é igual a todo o espazo, entón a condición cúmprese (é dicir, se $a - d = b = c = 0$ ou $a' - d' = b' = c' = 0$). Senón, os núcleos de f_A ou f_B son ambos de dimensión 1 e sempre conteñen o vector $(1, 0, 0, 1)$, polo que a fórmula de Grassmann afirma que a suma ten dimensión como moito 3. Polo tanto, a condicón do enunciado dáse se e soamente se $a - d = b = c = 0$ ou $a' - d' = b' = c' = 0$.

Problema 3.22. Sexa $f \in \mathrm{End}(\mathbb{R}^n)$. Fixado un vector $u \in \mathbb{R}^n$, supoñamos que

$$\mathcal{B} = \{u, f(u), \ldots, f^{n-1}(u)\}$$

é unha base de $\mathbb{R}^n$. Denotamos por $(a_0, a_1, \ldots, a_{n-1})$ as coordenadas de $f^n(u)$ na basde $\mathcal{B}$. Atopar os polinomios mínimo e característico de f en termos cos coeficientes a_i.

Solución. Comezamos observando que, aplicando o lema sobre determinantes empregado para demostrar Cayley–Hamilton, o polinomio característico é

$$\mathrm{Char}(f; X) = (-1)^n (X^n - a_{n-1} X^{n-1} - \ldots - a_1 X - a_0).$$

O polinomio mínimo é necesariamente un divisor do característico. Supoñamos que ten grao k, con $k < n$; pomos $\mathrm{Min}(f; X) = X^k + b_{k-1} X^{k-1} + \ldots + b_1 X + b_0$. Entón, $\mathrm{Min}(f; X)$ é anulador, polo que

$$0 = \mathrm{Min}(f; f)(u) = f^k(u) + b_{k-1} f^{k-1}(u) + b_1 f(v) + b_0 u,$$

o que querería dicir que o conxunto $\{u, f(u), \ldots, f^{k-1}(u), f^k(u)\}$ non é linealmente independente. Iso é unha contradición co feito de que $\mathcal{B}$ sexa unha base.

Problema 3.23. Sexa E un espazo vectorial e f un endomorfismo de E. Dise que f é nilpotente se existe un enteiro $n \geq 1$ de maneira que $f^n = 0$. Os seguintes apartados son independentes entre si.

(a) Demostrar que se $\mathrm{Char}(f; X) = (-X)^n$, entón f é nilpotente.

(b) Demostrar que para calquera f, $E = F \oplus G$, onde $f|F$ é un isomorfismo e $f|G$ é nilpotente.

Solución. (a) Polo teorema de Cayley–Hamilton, o polinomio característico é anulador, polo que $f^n = 0$.

(b) A idea esencial para a demostración é separar a parte correspondente ao valor propio 0 e o resto. En concreto, podemos escribir o polinomio mínimo de f como $\mathrm{Min}(f;X) = X^r \cdot P(X)$, onde $r \geq 0$ e $X \nmid P(X)$. Polo tanto, como $\gcd(X^r, P(X)) = 1$, temos que

$$E = \ker(P(f)) \oplus \ker(f^r).$$

Temos que $\mathrm{Min}(E|\ker(P(f))) = P(X)$; como $X \nmid P(X)$, o 0 non é un valor propio do endomorfismo $f|\ker(P(f))$, polo que $f|\ker(P(f))$ é invertible. Por outra banda, $\mathrm{Min}(f|\ker(f^r)) = X^r$, de onde temos que a restrición de f a $\ker(f^r)$ é nilpotente. Por conseguinte, o enunciado é certo collendo $F = \ker(P(f))$ e $G = \ker(f^r)$.

Forma de Jordan e subespazos invariantes.

Problema 3.24. Calcular bases de Jordan dos endomorfismos definidos polas seguintes matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}, \quad \begin{pmatrix} 9 & 15 & 5 \\ -1 & 1 & -1 \\ -2 & -6 & 2 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}.$$

Solución. Chamámoslles A, B e C ás matrices do enunciado.

Para a primeira, tense que $\mathrm{Char}(A;X) = -(X-2)^3$. Tense que $(A - 2\mathbb{I})^3$ é a matriz identicamente cero, polo que a forma de Jordan, J_A, consta dun bloque de tamaño 2 e dun bloque de tamaño 1:

$$J_A = \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right).$$

Unha base de $\ker(A-2\mathbb{I})$ é $\{(1,2,0),(0,0,1)\}$. Polo tanto, para dar unha base $\{v_1, v_2, v_3\}$ na que a matriz sexa J_A, collemos $v_2 \notin \ker A$, por exemplo $v_2 = (0,1,0)$; logo pomos $v_1 = (A-2\mathbb{I})v_2 = (1,2,1)$. Finalmente, completamos cun vector do núcleo, por exemplo $v_3 = (0,0,1)$.

Para a segunda, $\mathrm{Char}(B;X) = -(X-4)^3$. Igual que asaba antes, $(B-4\mathbb{I})^3$ é a matriz identicamente cero, polo que a forma de Jordan, J_B, consta dun bloque de tamaño 2 e dun bloque de tamaño 1:

$$J_B = \left(\begin{array}{cc|c} 4 & 1 & 0 \\ 0 & 4 & 0 \\ \hline 0 & 0 & 4 \end{array}\right).$$

Unha base de $\ker B$ é $\{(1,0,-1),(3,-1,0)\}$. Polo tanto, para dar unha base $\{v_1, v_2, v_3\}$ na que a matriz sexa J_B, collemos $v_2 \notin \ker B$, por exemplo $v_2 = (1,0,0)$; logo pomos $v_1 = (B-4\mathbb{I})v_2 = (5,-1,-2)$. Finalmente, completamos cun vector do núcleo, por exemplo $v_3 = (1,0,-1)$.

Para a terceira, $\mathrm{Char}(C;X) = -(X-1)^2(X+2)$. A multiplicidade xeométrica do valor propio 1 é 1, e o subespazo propio correspondente a ese valor propio é $\langle(-3,6,-20)\rangle$. A matriz de Jordan é

$$J_C = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -2 \end{array}\right).$$

Ao considerarmos $(C-\mathbb{I})^2$, temos que unha base do núcleo é $\{(-11,7,0),(-9,0,28)\}$. Polo tanto, para dar unha base de Jordan $\{v_1, v_2, v_3\}$, collemos $v_2 = (-11,7,0)$ e $v_1 = (C-\mathbb{I})v_2 = (-15,30,-100)$. Finalmente, pomos $v_3 = (0,0,1)$, que é un vector propio de valor propio -2.

Problema 3.25. Calcular bases de Jordan dos endomorfismos definidos polas seguintes matrices

$$\begin{pmatrix} 4 & 1 & 3 & 0 \\ -1 & 6 & 4 & 0 \\ 0 & 0 & 5 & 0 \\ 2 & 6 & 7 & 1 \end{pmatrix}, \quad \begin{pmatrix} -2 & 1 & -5 & 0 \\ -4 & 4 & -20 & 5 \\ 1 & 0 & 0 & 1 \\ -2 & -1 & 1 & -4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 9 & 3 & 4 \\ 0 & 4 & 1 & 2 \\ 0 & 2 & 0 & 1 \\ 0 & -8 & -2 & -4 \end{pmatrix}.$$

Solución. Sexa A a primeira matriz. O polinomio característico é $\mathrm{Char}(A; X) = (X-5)^3(X-1)$. O valor propio 1 terá multiplicidade xeométrica 1 e un vector propio asociado será o $v_1 = (0,0,0,1)$. Por outro lado, é doado comprobar que o 5 ten multiplicidade xeométrica 1. Polo tanto, teremos un único bloque de Jordan de tamaño 3, e para achar o xerador temos que coller un vector que estea no núcleo de $(A - 5\mathbb{I})^3$, pero non no de $(A - 5\mathbb{I})^2$. Collemos, por exemplo, $v_4 = (0,0,1,0)$. Entón, $v_3 = Av_4 = (3,4,0,7)$ e $v_2 = Av_3 = (1,1,0,2)$. Polo tanto, unha base de Jordan está dada polos vectores (v_1, v_2, v_3, v_4) e a forma de Jordan correspondente é

$$J_A = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{array}\right).$$

Chamando B á segunda matriz, temos que o polinomio característico é $\mathrm{Char}(B; X) = (X+2)^2(X-1)^2$. O valor propio -2 ten multiplicidade xeométrica 1. Para construír a base de Jordan collemos v_2 no núcleo de $(B + 2\mathbb{I})^2$, pero non no de $B + 2\mathbb{I}$. Por exemplo, $v_2 = (0,0,0,1)$ e $v_1 = Av_2 = (0,5,1,-2)$. Do mesmo xeito, 1 ten multiplicidade xeométrica 1. Collemos v_4 no núcleo de $(B - \mathbb{I})^2$, pero non de $B - \mathbb{I}$; nese caso, $v_4 = (0,1,0,0)$ e $v_3 = Bv_4 = (1,3,0,-1)$. Os vectores $\{v_1, v_2, v_3, v_4\}$ forman unha base de Jordan e a forma de Jordan correspondente é

$$J_B = \left(\begin{array}{cc|cc} -2 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

No terceiro caso, o polinomio característico da matriz C é $\mathrm{Char}(C; X) = X^4$. A multiplicidade xeométrica do valor propio 0 é 1, o que automaticamente quere dicir que hai un único bloque de Jordan de tamaño 4. Para achar un xerador, é suficiente con coller un vector que non estea no núcleo de C^3. Como

$$C^3 = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

podemos coller $v_4 = (0,0,0,1)$. Entón, $v_3 = Cv_4 = (4,2,1,-4)$, $v_2 = Cv_3 = (5,1,0,-2)$ e $v_1 = Cv_2 = (1,0,0,0)$. Polo tanto, unha base de Jordan está dada polos vectores (v_1, v_2, v_3, v_4) e a forma de Jordan correspondente é

$$J_C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Problema 3.26. Considérase o endomorfismo de $\mathbb{R}^4$ definida por

$$f(x, y, z, t) = (2x, 4y - 14z + 5t, y - 4z + 2t, y - 6z + 4t).$$

Encontrar os seus vectores propios e os vectores propios xeralizados, dar unha base de vectores propios xeralizados e a matriz de f nesta base.

Solución. A matriz correspondente ao endomorfismo f é

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & -14 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 1 & -6 & 4 \end{pmatrix}.$$

Cúmprese que $\mathrm{Char}(A; X) = (X-1)^2(X-2)^2$. No caso do vector propio 2 temos que $\ker(A - 2\mathbb{I})$ ten dimensión dous e está xerado polos vectores $v_3 = (1, 0, 0, 0)$ e $v_4 = (0, 2, 1, 2)$. No caso do valor propio 1, a dimensión de $\ker(A-\mathbb{I})$ é 1, e o espazo está xerado por $(0, 3, 1, 1)$. O espazo $\ker(A-\mathbb{I})^2$ ten dimensión 2 e está xerado por $(0, 1, 0, 0)$ e $(0, 0, 1, 1)$. Como o vector $v_2 = (0, 1, 0, 0)$ dá lugar a unha base de $\ker(A-\mathbb{I})^2/\ker(A-\mathbb{I})$, temos que forma xunto con $v_1 = (A-\mathbb{I})v_2 = (0, 3, 1, 1)$ unha base de Jordan. Polo tanto, $\{v_1, v_2, v_3, v_4\}$ é unha base de Jordan e a matriz nesa base é

$$\left(\begin{array}{cc|c|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right).$$

Problema 3.27. Determinar todas as posibles formas de Jordan para un endomorfismo f de polinomio característico $-(x-2)^3(x-5)^2$.

Solución. Para o valor propio 2, a forma de Jordan queda completamente determinada polas dimensións de $\ker(f - 2\,\mathrm{Id})$, $\ker(f - 2\,\mathrm{Id})^2$ e $\ker(f - 2\,\mathrm{Id})^3$. As posibilidades para estas dimensións son $(3, 3, 3)$, $(2, 3, 3)$ e $(1, 2, 3)$, que se corresponden cos bloques

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

respectivamente. Para o valor propio 5, as formas de Jordan quedan determinadas polas dimensións de $\ker(f - 5\,\mathrm{Id})$ e $\ker(f - 5\,\mathrm{Id})^2$, que poden ser $(2, 2)$ ou $(1, 2)$; os bloques correspondentes son

$$\begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix},$$

respectivamente. Polo tanto, hai 6 posibilidades, que dan lugar a matrices por bloques diagonais correspondentes a escoller unha das tres opcións para o valor propio 2 e unha das dúas opcións para o valor propio 5.

Problema 3.28. Achar, en cada un dos seguintes casos, a forma canónica de Jordan dun endomorfismo f se o seu polinomio característico é x^7.

(a) As dimensións dos núcleos de f e f^2 son 3 e 6, respectivamente.

(b) As dimensións dos núcleos de f, f^2 e f^3 son 3, 5 e 7, respectivamente.

(c) As dimensións dos núcleos de f e f^3 son 3 e 6, respectivamente.

Solución. (a) A dimensión do núcleo de f^3 ten que ser 7, porque de ser 6 a dimensión estabilizaría antes de chegar a 7, e sabemos que iso non é posible. Polo tanto, temos

$$\left(\begin{array}{ccc|cc|cc} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

é dicir, un bloque de tamaño 3 e dous bloques de tamaño 2.

(b) Neste caso temos dous bloques de tamaño 3 e un bloque de tamaño 1: temos

$$\left(\begin{array}{ccc|ccc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

(c) A dimensión do núcleo de f^2 ten que ser 5: se chamamos $3+d$ a esa dimensión, entón $d-3 \geq 6-d$, o que quere dicir que $d \geq 1{,}5$. Polo tanto, temos un bloque de tamaño 4, un bloque de tamaño 2 e un bloque de tamaño 1:

$$\left(\begin{array}{cccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Problema 3.29. Achar a forma canónica de Jordan dun endomorfismo que ten polinomio característico $-(x-2)^4(x-5)^3$ e no que os subespazos de vectores propios asociados aos valores propios 2 e 5 teñen dimensións 3 e 1, respectivamente.

Solución. Temos que $\dim\ker(f-2\,\mathrm{Id}) = 3$ e polo tanto ten que ser $\dim\ker(f-2\,\mathrm{Id})^2 = 4$, xa que a dimensión non pode estabilizar antes de chegar á multiplicidade alxébrica. Por outro lado, $\dim\ker(f-5\,\mathrm{Id}) = 1$, $\dim\ker(f-5\,\mathrm{Id})^2 = 2$ e $\dim\ker(f-5\,\mathrm{Id}) = 3$, xa que a dimensión de $\ker(f-5\,\mathrm{Id})^2$ non pode ser 3 pois iso implicaría que a aplicación

$$f_5 : \frac{\ker(f-5\,\mathrm{Id})^2}{\ker(f-5\,\mathrm{Id})} \longrightarrow \ker(f-5\,\mathrm{Id})$$

non é inxectiva. Polo tanto, a forma de Jordan consiste en tres bloques de Jordan para o vector propio 2, un de tamaño 2 e os outros de tamaño 1; e nun único bloque de

Jordan para o vector propio 5, este de tamaño 3:

$$\left(\begin{array}{cc|c|c|ccc} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right).$$

Problema 3.30. Sexa E un K-espazo vectorial de dimensión 5 e $f \in \operatorname{End}(E)$ unha aplicación lineal. Sábese que existe un $\lambda \in K$, con $\lambda \neq 0$, de maneira que se cumpren as seguintes propiedades simultaneamente: (1) $f(f - \lambda\mathbb{I}_5)^3 = 0$; (2) $\dim\ker(f - \lambda\mathbb{I}_5)^2 < \dim\ker(f - \lambda\mathbb{I}_5)^3$; (3) $\operatorname{rango}(f) \geq 4$.

(a) Determinar as posibles formas de Jordan de f e os polinomios característico e mínimo en cada caso.

(b) Se sabemos ademais que f é non inxectiva, demostrar que f^2 tamén cumpre as condicións (1), (2) e (3) do enunciado para algún escalar $\mu \in K$. Neste caso, determinar a forma de Jordan de f^2 e os polinomios característico e mínimo.

(c) Supoñamos que a matriz de f na base canónica é

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

Dar a forma de Jordan de A e atopar unha base de Jordan.

Solución. (a) Para a primeira parte imos distinguir dous casos segundo o rango de f sexa 4 ou 5. Observamos que da segunda condición λ ten que ser sempre un valor propio, pois en caso contrario as dimensións de $\ker(f - \lambda\operatorname{Id})^k$ serían sempre nulas.

- Se o rango é 4 temos que 0 é un valor propio. Nese caso, o polinomio mínimo é $\operatorname{Min}(f;X) = X(X-\lambda)^3$, xa que ese é un polinomio anulador e non é posible coller ningún divisor seu: por un lado, 0 é valor propio, e por outro $\dim\ker(f - \lambda\operatorname{Id})^2 < \dim\ker(f - \lambda\operatorname{Id})^3$. Como o expoñente de 0 no polinomio mínimo é 1, a multiplicidade alxébrica e a xeométrica coinciden, polo que $\operatorname{Char}(f;X) = -X(X-\lambda)^4$ e temos que a única opción para as dimensións dos núcleos é $\dim\ker(f-\lambda) = 2$, $\dim\ker(f-\lambda)^2 = 3$ e $\dim\ker(f-\lambda)^3 = 4$. Polo tanto, a forma de Jordan é

$$\left(\begin{array}{c|ccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array}\right).$$

- Se o rango é 5, o polinomio mínimo é $\operatorname{Min}(f;X) = (X-\lambda)^3$ e o polinomio característico é $\operatorname{Char}(f;X) = -(X-\lambda)^5$. Polo tanto, as únicas opcións para as dimensións dos núcleos son $\dim\ker(f-\lambda) = 2$, $\dim\ker(f-\lambda)^2 = 4$ e $\dim\ker(f-\lambda)^3 = 5$; ou $\dim\ker(f-\lambda) = 3$, $\dim\ker(f-\lambda)^2 = 4$ e $\dim\ker(f-\lambda)^3 = 5$. No primeiro caso, a forma de Jordan é

$$\left(\begin{array}{ccc|cc} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{array}\right),$$

mentres que no segundo é

$$\left(\begin{array}{ccc|c|c} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array}\right).$$

(b) No caso non inxectivo estamos sempre na primeira situación das descritas no epígrafe anterior, e polo tanto existe unha base $u = (u_1, \ldots, u_5)$ de E de maneira que as matrices de f e f^2 son

$$\left(\begin{array}{c|ccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda \end{array}\right) \quad \text{e} \quad \left(\begin{array}{c|ccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \lambda^2 & 2\lambda & 1 & 0 \\ 0 & 0 & \lambda^2 & 2\lambda & 0 \\ 0 & 0 & 0 & \lambda^2 & 0 \\ \hline 0 & 0 & 0 & 0 & \lambda^2 \end{array}\right),$$

respectivamente. Como $\lambda \neq 0$, o rango de f^2 é 4. Temos ademais que $\operatorname{Char}(f;X) = -X(X-\lambda^2)^4$ e os valores propios son 0 e λ^2 con multiplicidades alxébricas 1 e 4, respectivamente. Ademais, un cálculo rutineiro amosa que $\dim\ker(f^2-\lambda^2\operatorname{Id}) = 2$, $\dim\ker(f^2-\lambda^2\operatorname{Id})^2 = 3$ e $\dim\ker(f^2-\lambda^2\operatorname{Id})^3 = 4$. Polo tanto, o polinomio mínimo é $\operatorname{Min}(f;X) = X(X-\lambda^2)^3$ e temos que se cumpren as condicións do enunciado.

(c) Tense que $\operatorname{Char}(A;X) = -(X-2)^5$. Ademais, $\dim\ker(A-2\mathbb{I}_5) = 2$ e tamén $\dim\ker(A-2\mathbb{I}_5)^2 = 4$. Polo tanto, a forma de Jordan é da forma

$$\begin{array}{|c|c} \hline v_3 & \\ \hline v_2 & \multicolumn{1}{c|}{v_5} \\ \hline v_1 & \multicolumn{1}{c|}{v_4} \\ \hline \end{array}$$

Para escoller v_3, tomamos un vector que non pertenza ao núcleo de $(A-2\mathbb{I}_5)^2$, por exemplo $v_3 = (1,0,0,0,0)$. Entón, $v_2 = (A-2\mathbb{I}_5)(v_3) = (2,4,-2,-3,-2)$ e $v_1 = (A-2\mathbb{I}_5)(v_2) = (0,0,4,6,4)$. Para escoller v_5 tomamos un elemento de $\ker(f-2\operatorname{Id})^2$ de forma que $\{v_2, v_5\}$ sexa unha base de $\ker(A-2\mathbb{I}_5)^2/\ker(A-2\mathbb{I}_5)$. Por exemplo, $v_5 = (0,0,1,0,0)$, e nese caso $v_4 = (A-2\mathbb{I}_5)(v_4) = (0,0,0,0,3)$. Por conseguinte, $\{v_1, v_2, v_3, v_4, v_5\}$ é unha base de Jordan e a forma de Jordan é

$$\left(\begin{array}{ccc|cc} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{array}\right).$$

Problema 3.31. Sexa $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ o endomorfismo de $\mathbb{R}^4$ que na base canónica ten por matriz

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

Encontrar a forma de Jordan e unha base de Jordan, e encontrar todos os subespazos vectoriais invariantes por f.

Solución. O polinomio característico é $\mathrm{Char}(A; X) = X^2(X+1)^2$, polo que os valores propios son 0 e -1, ambos con multiplicidade alxébrica 2. O rango da matriz é 3, polo que a multiplicidade xeométrica de 0 é 1, e o subespazo de vectores propios está xerado por $(1, 0, -1, 0)$. Temos que

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

que ten rango 2 e ten núcleo xerado por $(1, 0, -1, 0)$ e $(-1, 6, 0, 1)$. Polo tanto, collemos $v_2 = (-1, 6, 0, 1)$ e $v_1 = Av_2 = (1, 0, -1, 0)$.
Para o valor propio -1, temos

$$A + \mathbb{I} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (A + \mathbb{I})^2 = \begin{pmatrix} 1 & 1 & 0 & -4 \\ 0 & 0 & 0 & 6 \\ -1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

O núcleo de $A+\mathbb{I}$ está xerado por $(1, -1, 0, 0)$ e o de $(A+\mathbb{I}^2)$ por $(1, -1, 0, 0)$ e $(0, 0, 1, 0)$. Collemos entón $v_4 = (0, 0, 1, 0)$ e $v_3 = (A + \mathbb{I})v_3 = (-1, 1, 0, 0)$. Entón, unha base de Jordan está dada por $\{v_1, v_2, v_3, v_4\}$ e a forma de Jordan é

$$\left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{array}\right).$$

Sobre os subespazos invariantes, temos:

- O subespazo $\{0\}$, que ten dimensión 0.
- Os subespazos $\langle v_1 \rangle$ e $\langle v_3 \rangle$, de dimensión 1 (un subespazo invariante de dimensión 1 está xerado por un vector propio).

- Os subespazos $\langle v_1, v_2\rangle$, $\langle v_3, v_4\rangle$ e $\langle v_1, v_3\rangle$, de dimensión 2, dado que un subespazo invariante de dimensión 2 está xerado ou por dous vectores propios ou por un vector propio e un xeralizado.
- Os subespazos $\langle v_1, v_2, v_3\rangle$ e $\langle v_1, v_3, v_4\rangle$, de dimensión 3, xerados polos dous vectores propios e un vector propio xeralizado.
- O subespazo total E de dimensión 4.

Explicamos agora o motivo polo que non hai máis. De haber algún outro, ao que chamaremos F, o polinomio mínimo da restrición a F é da forma $X^a(X+1)^b$, onde $0 \le a \le 2$ e $0 \le b \le 2$. Para cada elección do par (a, b), hai un único subespazo, que é o que describimos con anterioridade. En efecto, polo primeiro teorema de descomposición, é suficiente tratar cada valor propio por separado. No caso do valor propio 0, se $a = 0$, entón está claro que a restrición a F non ten ningún vector propio de valor propio 0; se $a = 1$, entón o subespazo propio xeralizado asociado a 1 consta unicamente de vectores propios, polo que novamente só hai unha opción, e é que v_1 estea, pero que non haxa ningún vector da forma $av_1 + bv_2$ con $b \neq 0$; por último, se $a = 2$, temos que tanto v_1 como v_2 pertencen a F.
Máis en xeral, se $f : E \to E$ é un endomorfismo tal que $\mathrm{Char}(f; X) = (-1)^n X^n$ e $m_x(0) = 1$, hai un único bloque de Jordan; sexa $\{v_1, \dots, v_n\}$ a base de Jordan. Se $f|F : F \to F$ é a restrición a un subespazo invariante, entón o 0 segue a ter multiplicidade xeométrica 1 e hai tamén un único bloque de Jordan. Se F ten dimensión r, sexa u un xerador do único bloque de Jordan e sexa $\{f^{r-1}(u), f^{r-2}(u), \dots, f(u), u\}$ unha base de Jordan, na que $f^r(u) = 0$. Se pomos $u = \sum_{i=1}^n a_i v_i$, entón $f^r(u) = \sum_{i=1}^{n-r} a_{r+i} v_i$, polo que $a_{r+1} = \dots = a_n = 0$. Entón, $u \in \langle v_1, \dots, v_r\rangle$, e máis en xeral

$$\langle f^{r-1}(u), f^{r-2}(u), \dots, f(u), u\rangle \subset \langle v_1, \dots, v_r\rangle.$$

Como ambos subespazos teñen a mesma dimensión, necesariamente os dous conxuntos son iguais e $F = \langle v_1, \dots, v_r\rangle$.

Problema 3.32. Sexa $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ o endomorfismo de $\mathbb{R}^3$ tal que $f(1, 0, 0) = (5, 1, -2)$, $f(1, 1, 0) = (6, 6, 0)$ e $f(1, 1, 1) = (4, 8, \alpha)$, onde $\alpha \in \mathbb{R}$ é o único valor que fai que f sexa non inxectivo.

(a) Determinar o valor de α encontrar a matriz de f na base canónica.

(b) Probar que f diagonaliza e encontrar unha base de E formada por vectores propios de f.

(c) Encontrar todos os $a, b, c \in \mathbb{R}$ tales que $a \cdot f^2 = b \cdot f + c \cdot \mathrm{Id}$, onde Id é o endomorfismo identidade.

(d) Dar un exemplo dun endomorfismo $g : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ cos mesmos valores propios que f e para o cal se cumpra $a \cdot g^2 = b \cdot g + c \cdot \mathrm{Id}$, con $a, b, c \in \mathbb{R}$, se, e soamente se, $a = b = c = 0$. (É dicir, que o único polinomio anulador de g de grao como moito dous sexa o polinomio 0).

Solución. (a) Usando as condicións do enunciado, como $(0, 1, 0) = (1, 1, 0) - (1, 0, 0)$ e $(0, 0, 1) = (1, 1, 1) - (1, 1, 0)$, temos que a matriz buscada é

$$A = \begin{pmatrix} 5 & 1 & -2 \\ 1 & 5 & 2 \\ -2 & 2 & \alpha \end{pmatrix}.$$

A aplicación será non inxectiva se e soamente se o determinante é 0. É inmediato comprobar que se ten entón que $\alpha = 2$.

(b) Procedendo do xeito habitual, $A = PDP^{-1}$, con

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

(c) O polinomio mínimo é $\mathrm{Min}(f;X) = X(X-6) = X^2 - 6X$, co cal $f^2 = 6f$. Calquera outro endomorfismo que anule f ten que ser un múltiplo do polinomio mínimo, co cal as únicas posibilidades para (a,b,c) son da forma $(\lambda, 6\lambda, 0)$, con $\lambda \in \mathbb{R}$.

(d) O endomorfismo g ten valores propios 0 e 6, este último con multiplicidade 2. Os valores propios son sempre raíz do polinomio mínimo, polo que este é un múltiplo de $X(X-6)$. Para que o polinomio mínimo teña grao maior que 2 (é dicir, 3), o endomorfismo non pode diagonalizar, polo que a multiplicidade xeométrica de 6 ten que ser 1. Isto quere dicir que a matriz de Jordan é

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 6 & 1 \\ 0 & 0 & 6 \end{pmatrix},$$

polo que é suficiente coller o endomorfismo representado por esta matriz na base canónica. Neste caso, o polinomio mínimo é $\mathrm{Min}(g;X) = X(X-6)^2$.

Problema 3.33. Sexa $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ unha aplicación lineal con polinomio mínimo $\mathrm{Min}(f;X) = X^2$. Supoñamos que F, G, H son tres subespazos vectoriais invariantes por f e que $\mathbb{R}^3 = F \oplus G$. É certo sempre que $H = (F \cap H) \oplus (G \cap H)$?

Solución. O resultado é falso. Collemos por exemplo $f(x,y,z) = (y,0,0)$, con matriz asociada

$$\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Sexa $F = [e_1, e_2]$, $G = [e_3]$ e $H = [e_2 + e_3, e_1]$. Entón, $F \cap H = [e_1]$ e $G \cap H = 0$.
De cara a complementar o resultado, podemos estudar todos os subespazos invariantes por f (imos traballar só os casos de dimensión 1 e 2, xa que obviamente o $\{0\}$ e o total sempre son invariantes). Sexa $\{v_1, v_2, v_3\}$ unha base de Jordan, que entón cumpre $f(v_2) = v_1$ e $f(v_1) = f(v_3) = 0$. Tense que un subespazo invariante de dimensión 1 é da forma $\langle \alpha v_1 + \beta v_3 \rangle$, con $(\alpha, \beta) \neq (0,0)$. Se é de dimensión 2, a restrición de f ou ben diagonaliza ou, en caso contrario, admite unha base $\{w_1, w_2\}$ de maneira que $f(w_2) = w_1$ e $f(w_1) = 0$. No primeiro caso tería que ser o subespazo $\langle v_1, v_3 \rangle$. No segundo, pomos $w_2 = av_1 + bv_2 + cv_3$; entón $f(w_2) = bv_1$, polo que $b \neq 0$. En particular, v_1 sempre pertence a un subespazo invariante de dimensión 2, e é da forma $\langle v_1, bv_2 + cv_3 \rangle$, xa que o caso $b = 0$ corresponde á situación na que a restrición diagonaliza.

Problema 3.34. Sexa E un $\mathbb{R}$-espazo vectorial e sexa $g \in \mathrm{End}(E)$ con polinomio característico $\mathrm{Char}(g;X) = X^4$ e polinomio mínimo $\mathrm{Min}(g;X) = X^3$.

(a) Demostrar que existe $v \in E$ de maneira que $g^2(v) \neq 0$.

(b) Demostrar que calquera subespazo invariante por g de dimensión maior ou igual que 2 sempre contén $g^2(v)$.

(c) Describir 4 subespazos invariantes de g de dimensión 1, 4 de dimensión 2 e 4 de dimensión 3.

Solución. (a) Polas condicións do enunciado, temos que a forma de Jordan de g consta dun bloque de Jordan de tamaño 3 e dun de tamaño 1. Polo tanto, podemos considerar un vector propio u de g, de maneira que $\{g^2(v), g(v), v, u\}$ é unha base de Jordan de g, polo que en particular $g^2(v)$ é distinto de 0. Alternativamente, podemos argumentar que se $g^2(v) = 0$ para todo $v \in E$ teríase que X^2 é un polinomio anulador, que contradice o feito de que o polinomio mínimo sexa X^3.

(b) Sexa F un subespazo invariante de dimensión 2. Entón, se $\text{Min}(g|F; X) = X$, temos que $F = \langle f^2(v), u\rangle$. Se $\text{Min}(g|F; X) = X^2$, o subespazo admite unha base de Jordan $\{w_1, w_2\}$, onde w_1 é un vector propio e w_2 un vector propio xeralizado. Se pomos

$$w_2 = af^2(v) + bf(v) + cv + du,$$

temos que $f(w_2) = bf^2(v) + cf(v)$ é un vector propio, polo que $c = 0$ e $b \neq 0$. Unha vez se ten esa condición, $f^2(w_2) = 0$. Polo tanto, podemos coller $w_2 = af^2(v) + bf(v) + du$, con $a, b, c \in K$ e $b \neq 0$, e $w_1 = bf^2(v)$. Do mesmo xeito, se G é un subespazo invariante de dimensión 3 e $\text{Min}(g|G; X) = X^2$, entón $G = \langle f^2(v), f(v), u\rangle$. Se $\text{Min}(g|X; X) = X^3$, entón G admite unha base de Jordan $\{w_1, w_2, w_3\}$. Se pomos

$$w_3 = af^2(v) + bf(v) + cv + du,$$

da condición de $w_1 = f^2(w_3) = cf^2(v)$, temos que $c \neq 0$ e w_1 sempre é un múltiplo de w_1. Finalmente, un subespazo de dimensión 4 é igual ó total e tamén contén $f^2(v)$.

(c) A forma canónica de Jordan de g é

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Chamámoslle v_1, v_2, v_3, v_4 aos vectores que forman a base.

- Subespazos invariantes de dimensión 1: podemos coller os vectores propios $\langle v_1\rangle$, $\langle v_4\rangle$, $\langle v_1 + v_4\rangle$, $\langle v_1 - v_4\rangle$. En xeral, serve calquera da forma $\langle \alpha v_1 + \beta v_4\rangle$, con $(\alpha, \beta) \neq (0, 0)$.
- Subespazos invariantes de dimensión 2: seguindo o razoamento anterior, collemos $\langle v_1, v_4\rangle$, $\langle v_1, v_2\rangle$, $\langle v_1, v_2 + v_4\rangle$, $\langle v_1, v_2 - v_4\rangle$. En xeral, serve calquera da forma $\langle v_1, \alpha v_2 + \beta v_4\rangle$, con $(\alpha, \beta) \neq (0, 0)$.
- Subespazos invariantes de dimensión 3: seguindo o razoamento anterior, collemos $\langle v_1, v_2, v_3\rangle$, $\langle v_1, v_2, v_4\rangle$, $\langle v_1, v_2, v_3 + v_4\rangle$, $\langle v_1, v_2, v_3 - v_4\rangle$. En xeral, serve calquera da forma $\langle v_1, v_2, \alpha v_3 + \beta v_4\rangle$, con $(\alpha, \beta) \neq (0, 0)$.

Problema 3.35. Sexa $f : \mathbb{R}^4 \to \mathbb{R}^4$ o endomorfismo de $\mathbb{R}^4$ que ten por matriz na base canónica

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

(a) Encontrar a forma reducida de Jordan J de f e unha base $\mathcal{B}_u$ de $\mathbb{R}^4$ para a cal a matriz do endomorfismo sexa J.

(b) Calcular A^{2024} (pódese deixar o resultado indicado en termos da inversa dunha matriz).

(c) Achar todos os subespazos vectoriais invariantes por f.

Solución. (a) Tense que $\mathrm{Char}(A; X) = X^2(X-2)^2$ e ambos valores propios teñen multiplicidade xeométrica 1. Temos que

$$\ker(A) = \langle (2,-1,0,-2), \quad \ker(A^2) = \langle (2,-1,0,-2), (1,0,1,0) \rangle.$$

Collemos entón $v_2 = (1,0,1,0)$ e $v_1 = Av_2 = (2,-1,0,-2)$. Por outro lado,

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

polo que podemos definir $v_4 = (0,0,0,1)$ e $v_3 = (A-2\mathbb{I})v_4 = (0,1,0,0)$. Entón, $\{v_1, v_2, v_3, v_4\}$ é unha base de Jordan e a forma reducida de Jordan é

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

(b) Tense que

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

Como

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}^{2024} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2^{2024} & 2024 \cdot 2^{2023} \\ 0 & 0 & 0 & 2^{2024} \end{pmatrix},$$

entón

$$A^{2024} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2^{2024} & 2024 \cdot 2^{2023} \\ 0 & 0 & 0 & 2^{2024} \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 0 & 1 \end{pmatrix}^{-1}.$$

Operando, chegamos a

$$A^{2024} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2025 \cdot 2^{2023} & 2^{2024} & -2025 \cdot 2^{2023} & 2024 \cdot 2^{2023} \\ 0 & 0 & 0 & 0 \\ 2^{2024} & 0 & -2^{2024} & 2^{2024} \end{pmatrix}.$$

(c) Hai 9 subespazos invariantes: un de dimensión 0, que é o $\{0\}$; dous de dimensión 1, os xerados por v_1 e por v_3; tres de dimensión 2, $\langle v_1, v_3\rangle$, $\langle v_1, v_2\rangle$ e $\langle v_3, v_4\rangle$; dous de dimensión 3, $\langle v_1, v_2, v_3\rangle$, $\langle v_1, v_3, v_4\rangle$; e un de dimensión 4, o total.

Para demostrar que non hai máis, consideramos a forma de Jordan da restrición da aplicación lineal ao subespazo, que unicamente pode ter valores propios 0 e 2; se o 0 é valor propio de multiplicidade alxébrica 1, o subespazo de valores propios xeralizado é $\langle v_1\rangle$, mentres que se ten multiplicidade alxébrica 2, é $\langle v_1, v_2\rangle$; o cálculo para o 2 é análogo. Polo primeiro teorema de descomposición, sabemos que podemos analizar por separado cada valor propio, polo que as posibilidades son as nove antes mencionadas, xa que hai tres opcións para cada valor propio.

Problema 3.36. Sexa $a \in \mathbb{R}$ e sexa f o endomorfismo de $E = \mathbb{R}_3[X]$ definido por

$$\begin{aligned} f(1) &= 2, \\ f(X) &= 2X + 1, \\ f(X^2) &= 2X^2 + 2a, \\ f(X^3) &= 2X^3 + X^2 - 2aX + 4a^2. \end{aligned}$$

(a) Encontrar a forma de Jordan e unha base de Jordan para o endomorfismo f.

(b) Calcular $f^{2024}(X^3)$.

(c) Sexa $F \subset E$ un subespazo de dimensión m invariante por f. Demostrar que $\mathrm{Char}(f|F; X) = (2 - X)^m$ e que F ten unha base formada por ciclos de vectores propios xeralizados de f.

(d) Dar tres subespazos $F_1, F_2, F_3 \subset V$ de dimensións 1, 2 e 3 respectivamente, que sexan invariantes por f.

Solución. (a) Traballamos na base natural $\mathcal{B} = \{1, X, X^2, X^3\}$. Un elemento de $\mathbb{R}_3[X]$ pódese representar indistintamente como un polinomio ou en termo das súas coordenadas nesta base. A matriz do endomorfismo é

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

Por tratarse dunha matriz triangular superior, vemos de xeito inmediato que $\mathrm{Char}(f; X) = (X - 2)^4$, polo que o único valor propio é 2. Temos que $\ker(f_2) = \langle(1, 0, 0, 0), (0, 2a, -1, 0)\rangle$, mentres que $\ker(f_2^2) = \mathbb{R}_3[X]$. Por conseguinte, a forma de Jordán é

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Para atopar unha base de Jordan seleccionamos dous vectores v_2 e v_4 que sexan linealmente independentes en $\ker(f_2^2)/\ker(f_2)$; por exemplo, $v_2 = (0, 1, 0, 0)$ e $v_4 = (0, 0, 0, 1)$. Nese caso, $v_1 = f_2(v_2) = (1, 0, 0, 0)$ e $v_3 = f_2(v_4) = (4a^2, -2a, 1, 0)$. Pola tanto, unha base de Jordan é

$$\{1, X, X^2 - 2aX + 4a^2, X^3\}.$$

Chamámoslle P á matriz que ten por columnas os vectores v_1, v_2, v_3 e v_4.

(b) Para calquera enteiro n non negativo temos que

$$J^n = \begin{pmatrix} 2^n & n2^{n-1} & 0 & 0 \\ 0 & 2^n & 0 & 0 \\ 0 & 0 & 2^n & n2^{n-1} \\ 0 & 0 & 0 & 2^n \end{pmatrix}.$$

Para atopar $f^{2024}(x^3)$ simplemente temos que multiplicar a matriz $PJ^{2024}P^{-1}$ polo vector columna $(0,0,0,1)$. Resulta entón que

$$f^{2024}(X^3) = 2^{2024}X^3 + 2024 \cdot 2^{2023}X^2 - 2024a2^{2024}X + 2024a^2 2^{2025}.$$

(c) É un resultado xeral que se f é un endomorfismo dun espazo vectorial que admite unha base de Jordan, entón todo subespazo $F \subset E$ invariante por f ten unha base formada por ciclos de vectores propios xeralizados de f. Ademais, o polinomio característico da restrición de f a F é un divisor do polinomio característico de f.

Imos comprobalo neste caso concreto. Sexa $(u_1, \dots, u_m)$ unha base de F e completámola a unha base de $\mathbb{R}_3[X]$, $(u_1, \dots, u_m, \dots, u_4)$. A matriz de f nesa base, á que chamaremos M, é unha matriz por bloque da forma

$$M = \begin{pmatrix} M_F & B \\ 0 & C \end{pmatrix},$$

polo que $\operatorname{Char}(M;X) = \operatorname{Char}(M_F;X)\operatorname{Char}(M,C)$. Como $\operatorname{Char}(M;X) = (X-2)^4$, necesariamente teremos que $\operatorname{Char}(M_F;X) = (-1)^m(X-2)^m$.

Finalmente, como o polinomio característico de $f|F$ descompón en factores lineais, o teorema de existencia de bases de Jordan asegura que existe unha base de Jordan do espazo F formada por ciclos de vectores propios xeralizados do endomorfismo $f|F$. Como $f|F$ é a mesma aplicación f pero restrinxida a F, estes ciclos tamén son ciclos de vectores propios xeralizados do endomorfismo f.

(d) Os subespazos invariantes de dimensión 1 están xerados por vectores propios. Por exemplo, podemos coller $F_1 = \langle v_1 \rangle$, polo que $X_2 = X_3 = X_4 = 0$. Os subespazos de dimensión 2 teñen por base ou ben un ciclo de lonxitude dous ou dous vectores propios. Por exemplo, podemos coller $F_2 = \langle v_1, v_2 \rangle$, polo que as ecuacións son $X_3 = X_4 = 0$. No caso de dimensión 3, a única opción é coller un ciclo de lonxitude 2 e un vector propio independente. Por exemplo, $F_3 = \langle v_1, v_2, v_3 \rangle$, polo que podemos coller o subespazo dado pola ecuación $X_4 = 0$.

Problema 3.37. Calcular J^n, onde $J \in \mathcal{M}_6(K)$ é a matriz diagonal por bloques que ten na diagonal

$$J_1 = (\lambda), \quad J_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad J_3^n = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Solución. É unha demostración estándard por indución:

$$J_1^n = (\lambda^n), \quad J_2^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}, \quad J_3 = \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}.$$

Polo tanto, J^n é a matriz diagonal por bloques con J_1^n, J_2^n e J_3^n.

Problema 3.38. Dar a matriz do endomorfismo $f \in \mathrm{End}(\mathbb{R}^3)$ que cumpre as seguintes cinco condicións.

(i) $\langle e_1, e_2\rangle$ e $\langle e_3\rangle$ son subespazos invariantes por f.

(ii) $\ker(f) = \langle (1, 2, 0)\rangle$.

(iii) A matriz de f na base canónica é simétrica.

(iv) $\mathrm{Tr}(f) = 6$.

(v) 1 é un valor propio de f.

Solución. Da primeira condición temos que $(0, 0, 1)$ é un vector propio, e da segunda, que $(1, 2, 0)$ é un vector propio de valor propio 0. A cuarta afirma que 1 é un valor propio, e como a suma dos valores propios é 6, o outro vector propio ten que ser 5. Imos impor estas condicións sobre a matriz de f, que sabemos que pola simetría será da forma

$$\begin{pmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & d \end{pmatrix}.$$

As ecuacións que temos son as seguintes:

$$a + 2b = 0, \quad b + 2c = 0, \quad a + c + d = 6, \quad (d - 1)(ac - a - c + 1 - b^2) = 0,$$

onde a última quere dicir que 1 é valor propio. Se $d = 1$, entón $(a, b, c) = (4, -2, 1)$, e a matriz cumpre todas as condicións. Se $d \neq 1$, ten que ser $d = 5$, xa que é un valor propio e non pode ser o 0 (xa que o subespazo propio é unha combinación lineal de e_1 e e_2). Nese caso $(a, b, c) = (0{,}2, -0{,}4, 0{,}8)$ e a matriz tamén cumpre as condicións.

Problema 3.39. Sexa $\varphi\colon \mathbb{R}^4 \to \mathbb{R}^4$ o isomorfismo que ten por matriz na base canónica

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

(a) Calcular a súa forma de Jordan e unha descomposición de Jordan da forma $A = PJP^{-1}$.

(b) Determinar o polinomio mínimo de A.

(c) Demostrar que existe un subespazo invariante F de dimensión 2 que contén tódolos vectores propios de A. Describir F en termos dos seus xeradores.

(d) Sexa $u \notin F$. Demostrar que existe un único subespazo invariante G_u de dimensión 2 que contén u.

(e) Sexa H un subespazo invariante de dimensión 3. Demostrar que $F \subset H$.

Solución. (a) Comezamos observando que a multiplicidade xeométrica do valor propio 2 é igual a 2 e que

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

Por outra banda, temos que $(A - 2\mathbb{I})^2 = 0$, polo que a forma de Jordan ten unha descomposición en ciclos da forma

v_2	v_4
v_1	v_3

Para escoller v_2 e v_4 consideramos dous vectores de xeito que as súas clases sexan unha base do espazo cociente $\mathbb{R}^4/\ker(A-2\mathbb{I})$; isto é equivalente a pedir que ao xuntalos cos dous vectores que forman unha base de $\ker(A-2\mathbb{I})$ dean lugar a unha base de $\mathbb{R}^4$. Podemos considerar, polo tanto, $v_2 = (0,1,0,0)$ e $v_4 = (0,0,1,0)$. Nese caso, $v_1 = (A-2\mathbb{I})v_2 = (2,0,0,1)$ e $v_3 = (A-2\mathbb{I})v_4 = (-1,0,0,2)$. Polo tanto, $\{v_1, v_2, v_3, v_4\}$ é unha base de Jordan e a matriz de Jordan é

$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

A matriz P é simplemente a que ten por columnas os vectores da base de Jordan:

$$P = \begin{pmatrix} 2 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 \end{pmatrix}.$$

(b) O único valor propio é o 2, e o tamaño do maior bloque de Jordan asociado a ese valor propio é 2. Iso quere dicir que $\mathrm{Min}(A; X) = (X-2)^2$.

(c) O subespazo de vectores propios correspóndese con $\ker(A-2\mathbb{I})$, que é un subespazo de $\mathbb{R}^4$ de dimensión 2 xerado por v_1 e v_3. Máis en concreto, isto dinos que

$$F = \langle v_1, v_3 \rangle = \langle (1,0,0,0), (0,0,0,1) \rangle.$$

Non pode haber outros subespazos formados unicamente por vectores propios; o seu polinomio mínimo tería que ser $X-2$, polo que tódolos vectores do subespazo teñen que cumprir que $\varphi(v) = 2v$.

(d) Sexa $\{v_1, v_2, v_3, v_4\}$ a base de Jordan do primeiro apartado. e sexa $u = \sum_{i=1}^4 a_i v_i$, de xeito que $(a_2, a_4) \neq (0,0)$. Se G_u é invariante, necesariamente contén $\varphi(u)$ e tamén

$$\tilde{u} = \varphi(u) - 2u = a_2 v_1 + a_4 v_3.$$

Como $u \notin F$, tense que $(a_2, a_4) \neq (0,0)$, polo que $\tilde{u} \neq 0$ e, ademais, $\tilde{u}$ é linealmente independente con u. Polo tanto, calquera subespazo invariante que conteña u contén tamén $\tilde{u}$. Polo tanto, o único posible subespazo invariante de dimensión 2 é

$$G_u = \langle u, \varphi(u) - 2u \rangle,$$

e unha comprobación rutineira amosa que é de feito invariante, xa que $\varphi(\tilde{u}) = 2\tilde{u}$.

(e) Como H é un subespazo invariante, podemos considerar a súa forma de Jordan, que necesariamente ten que constar dun bloque de tamaño 2 e de outro de tamaño 1. É dicir, $\mathrm{Min}(\varphi|H; X)$ divide $(X-2)^2$, e non pode ser $X-2$, porque iso querería dicir que hai un subespazo de dimensión 3 de vectores propios, que non é posible porque F ten dimensión 2. Polo tanto, $\mathrm{Min}(\varphi|H; X) = (X-2)^2$. De aquí tense que

$\dim\ker((\varphi - 2\,\mathrm{Id})|H) = 2$ e $\dim\ker((\varphi - 2\,\mathrm{Id})^2|H) = 3$. Como $\ker((\varphi - 2\,\mathrm{Id})|H)$ é de dimensión 2 e está contido en $F = \ker(\varphi - 2\,\mathrm{Id})$, ambos subespazos teñen que ser iguais, o que demostra que $F \subset H$.

Alternativamente, podemos argumentar que a forma de Jordan ten a seguinte estrutura:

<table>
<tr><td>w_2</td><td></td></tr>
<tr><td>w_1</td><td>w_3</td></tr>
</table>

Tense que $w_2 = \sum_{i=1}^4 a_i v_i$, con $(a_2, a_4) \neq (0,0)$, e $w_1 = a_2 v_1 + a_4 v_3$. Logo, cóllese $w_3 = b_1 v_1 + b_2 v_3$, de xeito que os vectores w_1 e w_3 sexan linealmente independentes. Polo tanto,

$$H = \langle w_1, w_2, w_3 \rangle = \langle v_1, v_3, a_2 v_2 + a_4 v_4 \rangle.$$

Problema 3.40. Sexa $A \in \mathcal{M}_3(\mathbb{R})$ unha matriz cadrada 3×3 que cumpre que $\mathrm{Tr}(A) = \mathrm{Tr}(A^2) = \mathrm{Tr}(A^3) = 0$, onde Tr é a traza da matriz.

(a) Demostrar que $A^3 = 0$.

(b) Cúmprese sempre que $A^2 = 0$? Se é certo, demostralo; senón, dar un contraexemplo.

Sexa agora K un corpo arbitrario e sexa $B \in \mathcal{M}_3(K)$ unha matriz cadrada 3×3 que cumpre que $\mathrm{Tr}(B) = \mathrm{Tr}(B^2) = \mathrm{Tr}(B^3) = 0$.

(c) Cúmprese sempre que $B^3 = 0$? Se é certo, demostralo; senón, dar un contraexemplo.

Solución. (a) A traza da matriz A é a suma dos seus valores propios. Estes, en principio, non teñen por que ser reais, senón que estarían definidos sobre os complexos. Por outro lado, se os valores propios de A son λ_i, con $1 \leq i \leq 3$, os valores propios de A^k, para calquera $k \geq 0$ enteiro, son λ_i^k, para $1 \leq i \leq 3$. Polo tanto, a condición do enunciado afirma que

$$\lambda_1 + \lambda_2 + \lambda_3 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 = 0.$$

Amosamos agora dúas maneiras de demostrar que $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Da condición $\lambda_3 = -\lambda_1 - \lambda_2$ tense que

$$0 = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 = 3\lambda_1\lambda_2(\lambda_1 + \lambda_2) = 0.$$

Se $\lambda_1 = 0$, entón $\lambda_2^2 + \lambda_3^2 = 2\lambda_2^2 = 0$ e conclúese que os tres valores propios son 0. Se $\lambda_2 = 0$, o razoamente é análogo; e se $\lambda_1 + \lambda_2 = 0$, logo $\lambda_3 = 0$, e tamén se ten que todos son nulos.

Alternativamente, temos que

$$\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 = \frac{(\sum_{i=1}^3 \lambda_i)^2 - \sum_{i=1}^3 \lambda_i^2}{2} = 0$$

e ademais

$$\lambda_1\lambda_2\lambda_3 = \frac{(\sum_{i=1}^3 \lambda_i)^3 + 2\sum_{i=1}^3 \lambda_i^3 - 3\sum_{i=1}^3 \lambda_i^2 \sum_{i=1}^3 \lambda_i}{6} = 0$$

Concluímos entón que λ_1, λ_2 e λ_3 son as raíces do polinomio $X^3 = 0$, co cal $\lambda_1 = \lambda_2 = \lambda_3 = 0$:

$$(X-\lambda_1)(X-\lambda_2)(X-\lambda_3) = X^3-(\lambda_1+\lambda_2+\lambda_3)X^2+(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1)X-\lambda_1\lambda_2\lambda_3.$$

Entón, como as tres raíces son reais, A admite forma de Jordan, e esta terá tres ceros na diagonal principal; podería ter, ademais, un ou dous uns na diagonal superior. Se non hai ningún bloque de Jordan de tamaño 3, entón xa sucede que $A^2 = 0$; e se hai un único bloque de Jordan de tamaño 3, entón $A^2 \neq 0$, pero $A^3 = 0$.

(b) É falso. Seguindo o razoamento do apartado anterior, é suficiente con considerar a matriz

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

para a cal

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

(c) É falso. É suficiente con coller $K = \mathbb{Z}/3\mathbb{Z}$ e a matriz identidade,

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Neste caso $B = B^2 = B^3$ e $\mathrm{Tr}(B) = 3 = 0$. A idea clave é que cando a característica de K é 2 ou 3, isto é, $2 = 0$ ou $3 = 0$, entón o razoamento do primeiro apartado non serve que coñecer a suma dos cadrados e a suma dos cubos non é suficiente para determinar univocamente as raíces do polinomio.

Forma de Jordan e subespazos invariantes.

Problema 3.41. A evolución temporal dunha poboación de árbores está descrita polo seguinte modelo matemático:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & a/5 \\ 1 & 9/5 \end{pmatrix}\begin{pmatrix} x_n \\ y_n \end{pmatrix},$$

con $a \in \mathbb{R}$ un parámetro e onde x_n e y_n representan as densidades das clases de árbores en formación e as clases de árbores formados, respectivamente.

(a) Para que valores de a é $\lambda = 1$ un valor propio da matriz?

(b) Para $a = 1$, encontrar a forma de Jordan e unha base de Jordan da matriz. Estudar o comportamento a longo prazo da poboación de árbores.

(c) Para $a = -4/5$, encontrar a forma de Jordan e unha base de Jordan da matriz.

(d) Estudar o comportamento a longo prazo da poboación para $a \geq -4/5$.

Solución. (a) Un é valor propio se, e soamente se, o determinante da matriz menos a identidade é 0:

$$\begin{pmatrix} -1/2 & a/10 \\ 1/2 & -1/10 \end{pmatrix}$$

ten determinante $(1-a)/20$, polo que ten que pasar que $a = 1$.

(b) Os valores propios son 1 e 0.4. O subespazo propio asociado ao valor propio 1 está xerado por $(1,5)$ e o asociado ao valor propio 0.4, por $(1,-1)$. Polo tanto,

$$\lim_{n\to\infty} \begin{pmatrix} 0{,}5 & 0{,}1 \\ 0{,}5 & 0{,}9 \end{pmatrix}^n = \begin{pmatrix} 1 & 1 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/6 & 1/6 \\ 5/6 & -1/6 \end{pmatrix} = \begin{pmatrix} 1/6 & 1/6 \\ 5/6 & 5/6 \end{pmatrix}.$$

Polo tanto, se (x_0, y_0) son as densidades iniciais de árbores en formación e de árbores formadas, respectivamente, tense que os valores no límite, (x,y), cumpren que $x = \frac{x_0+y_0}{6}$ e $y = \frac{5(x_0+y_0)}{6}$.

(c) Neste caso a matriz é

$$A = \begin{pmatrix} 0{,}5 & -0{,}08 \\ 0{,}5 & 0{,}9 \end{pmatrix},$$

que ten o valor propio dobre $\lambda = 0{,}7$. Collemos $v_2 = (1,0)$ e $v_1 = (A - 0{,}7\mathbb{I})v_2 = (-2,5)$, de maneira que (v_1, v_2) é unha base de Jordan. A forma reducida de Jordan é

$$J = \begin{pmatrix} 0{,}7 & 1 \\ 0 & 0{,}7 \end{pmatrix}.$$

Está claro que o límite cando n tende a infinito de J^n é a matriz 0, xa que $\lim_{n\to\infty} n\lambda^{n-1} = 0$ para calquera $|\lambda| < 1$.

(d) O polinomio característico é

$$x^2 - 1{,}4x + (0{,}45 - 0{,}05a),$$

que se $a \geq -0{,}8$ ten as raíces reais $0{,}7 \pm \sqrt{0{,}04 + 0{,}05a}$.

- Para $a = -0{,}8$ xa vimos que o límite tendía a 0 e a poboación converxía cara á extinción.
- O mesmo sucede se $-0{,}8 < a < 1$, xa que ambas raíces do polinomio característico son menores que 1 en valor absoluto.
- O caso $a = 1$ xa o estudamos (a poboación non se extingue).
- Se $1 < a < 57$, un dos valores propios é maior que 1 e o outro ten valor absoluto menor que 1; chamámoslle λ_1 e λ_2, respectivamente. Iso quere dicir que a poboación crece infinitamente na dirección do maior valor propio, mentres que na outra dirección se anula. É dicir, se a condición inicial se expresa como $c_0 = ae_1 + be_2$, onde e_1 é o vector propio de maior valor propio e e_2 é o de menor, despois de n iteracións temos que $c_n = a\lambda_1^n e_1 + b\lambda_2^n e_2$; λ_1^n tende a infinito mentres que λ_2^n tende a 0.
- Se $a \geq 57$ un dos valores propios é maior que 1 e o outro menor ou igual que -1; chamámoslle λ_1 e λ_2, respectivamente. Se a condición inicial se expresa como $c_0 = ae_1 + be_2$, onde e_1 é o vector propio de valor propio positivo e e_2 o de valor propio negativo, entón despois de n iteracións temos

que $c_n = a\lambda_1^n e_1 + b\lambda_2^n e_2$. Como $|\lambda_1| > |\lambda_2$, temos que λ_1^n é dominante e a poboación tende a infinito (supondo sempre que $a > 0$, dado que en caso contrario comezariamos con densidades negativas); isto é consistente co feito de que todas as entradas da matriz son positivas.

Problema 3.42. Encontrar todas as matrices A tales que

$$A^2 = \begin{pmatrix} 16 & 1 & 0 \\ 0 & 16 & 0 \\ 13 & -25 & 3 \end{pmatrix}.$$

Solución. O polinomio característico é $\mathrm{Char}(A^2; X) = -(X-16)^2(X-3)$. Temos que $A^2 - 16\mathbb{I}$ está xerado por $(1,0,1)$ e $(A^2 - 16\mathbb{I})^2$ por $(1,0,1)$ e $(2,1,0)$. Collemos $v_2 = (2,1,0)$ e $v_1 = (A - 16\mathbb{I})v_2 = (1,0,1)$. Por outro lado, o núcleo de $A^2 - 3\mathbb{I}$ está xerado por $(0,0,1)$. Entón,

$$A^2 = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \left(\begin{array}{cc|c} 16 & 1 & 0 \\ 0 & 16 & 0 \\ \hline 0 & 0 & 3 \end{array}\right) \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1}.$$

Para facer a raíz cadrada do bloque de Jordan, escollemos unha raíz cadrada de 16 e logo resolvemos. Se collemos a raíz positiva,

$$\begin{pmatrix} 4 & a \\ 0 & 4 \end{pmatrix}^2 = \begin{pmatrix} 16 & 1 \\ 0 & 16 \end{pmatrix},$$

e entón $8a = 1$, polo que $a = 1/8$. Se collemos a raíz negativa,

$$\begin{pmatrix} -4 & b \\ 0 & -4 \end{pmatrix}^2 = \begin{pmatrix} 16 & 1 \\ 0 & 16 \end{pmatrix},$$

e entón $-8b = 1$, polo que $b = -1/8$. As matrices posibles son entón o resultado de coller as formas de Jordan

$$\begin{pmatrix} 4 & 1/8 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix}, \begin{pmatrix} 4 & 1/8 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix}, \begin{pmatrix} -4 & -1/8 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix}, \begin{pmatrix} -4 & -1/8 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix}.$$

Se lle chamamos P á matriz que ten por columnas a base de Jordan e J_i a cada un destes catro posibles bloques de Jordan, temos que as raíces cadradas da matriz están dadas por PJ_iP^{-1}.

Capítulo 4

Formas bilineais e cuadráticas

Ao longo deste capítulo, o corpo K será sempre $\mathbb{R}$ ou $\mathbb{C}$, salvo que se indique especificamente o contrario; en particular, ao tratar con formas cuadráticas e simplécticas faremos algúns resultados de carácter xeral. A motivación para desenvolver os conceptos que imos tratar é o estudo da xeometría: en $\mathbb{R}^2$ ou en $\mathbb{R}^3$ necesitamos falar de distancias entre puntos, rectas ou planos, e tamén do ángulo que forman dous vectores ou da noción de perpendicularidade. Aquí estenderemos esas nocións a un $\mathbb{R}$-espazo vectorial ou a un $\mathbb{C}$-espazo vectorial, engadindo unha nova estrutura que permita definir os conceptos xeométricos análogos. Esta estrutura é o que se chama produto escalar (resp. hermítico no caso complexo), e os $\mathbb{R}$-espazos vectoriais (resp. $\mathbb{C}$-espazos vectoriais) nos que hai un produto escalar chámanse espazos euclidianos (resp. hermíticos). Ademais, estaremos especialmente interesados en estudar aquelas aplicacións lineais que respectan a estrutura do produto escalar, o que nos levará a introducir as nocións de aplicacións autoadxuntas e de isometrías, entre outras.

4.1. O espazo euclidiano

Definicións e primeiras propiedades

Definición 4.1. Sexa E un $\mathbb{R}$-espazo vectorial. Un *produto escalar* en E é unha aplicación bilineal $\phi\colon E \times E \to \mathbb{R}$ que ten as seguintes dúas propiedades:

(i) **Simétrica.** $\phi(u, v) = \phi(v, u)$ para todo $u, v \in E$.

(ii) **Definida positiva.** $\phi(u, u) > 0$ para todo vector $u \neq 0$.

Polo xeral escribiremos $\langle u, v\rangle$ en vez de $\phi(u, v)$. Tamén é común escribir $u \cdot v$, aínda que non usaremos esa notación.
No caso complexo, a definición é menos natural, xa que non imos usar o concepto de bilinealidade, senón o de sesquilinealidade (o prefixo *sesqui* denota unha unidade e media; por exemplo, sesquicentenario significa 150 anos).

Definición 4.2. Sexa E un $\mathbb{C}$-espazo vectorial. Unha aplicación $\phi\colon E\times E \to \mathbb{C}$ dise que é *sesquilineal* se é lineal na primeira variable, respecta a suma na segunda e $\phi(x, \lambda y) = \bar{\lambda}\phi(x, y)$, onde $\bar{\lambda}$ refírse á conxugación complexa dun número.
Dise que unha aplicación sesquilineal ϕ é un *produto hermítico* se cumpre as seguintes dúas propiedades.

(i) **Hermítica.** $\phi(u, v) = \overline{\phi(v, u)}$ para todo $u, v \in E$.

(ii) **Definida positiva.** $\phi(u,u) \in \mathbb{R}$ e $\phi(u,u) > 0$ para todo vector $u \neq 0$.

Pódese observar que na propiedade (ii), o feito de que $\phi(u,u) \in \mathbb{R}$ é consecuencia da propiedade (i), polo que o único contido novo é que $\phi(u,u) > 0$ para todo $u \neq 0$. Algúns dos exemplos máis importantes de produtos escalares son os seguintes.

- No espazo vectorial $\mathbb{R}^n$, o produto escalar habitual é

$$\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

- No espazo $\mathcal{C}([a,b],\mathbb{R})$ das funcións reais continuas no intervalo $[a,b]$, defínese

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx.$$

  Este produto escalar tamén se pode considerar en subespazos como o das funcións polinómicas.

No caso dos produtos hermíticos, os exemplos máis destacados son os seguintes.

- No espazo vectorial $\mathbb{C}^n$, o produto escalar habitual é

$$\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n.$$

- No espazo $\mathcal{C}([a,b],\mathbb{C})$ das funcións reais continuas no intervalo $[a,b]$, defínese

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\, dx.$$

  Este produto escalar tamén se pode considerar en subespazos como o das funcións polinómicas.

Definición 4.3. Un *espazo euclidiano* é un $\mathbb{R}$-espazo vectorial E cun produto escalar $\langle \cdot, \cdot \rangle$. Do mesmo xeito, un *espazo hermítico* é un $\mathbb{C}$-espazo vectorial cun produto hermítico.

É importante observar que todo subespazo vectorial dun espazo euclidiano ou hermítico é tamén un espazo vectorial ou hermítico, respectivamente. A seguinte definición ímola realizar no caso euclidiano, pero funciona exactamente igual no caso hermítico.

Definición 4.4. Sexa $(E, \langle \cdot, \cdot \rangle)$ un espazo euclidiano. Defínese a *norma* dun vector $v \in E$ como o número real $\|v\| = \sqrt{\langle v, v \rangle}$. Os vectores de norma 1 chámanse unitarios.

Nun espazo euclidiano, hai unha relación que non é completamente trivial entre o espazo dual e as formas bilineais. Sexa $v \in E$ e consideramos a aplicación $\phi_v \colon E \to K$ definida por $\phi_v(w) = \phi(v,w)$. Entón, a aplicación é lineal, polo que pertence ao espazo dual.

Proposición 4.1. A aplicación

$$\psi \colon E \longrightarrow E^*, \quad v \mapsto \phi_v$$

é un isomorfismo de espazos vectoriais.

Demostración. A linealidade é inmediata a partir da definición. Para ver a inxectividade, observamos que se $\phi_v = 0$, entón, $\langle v, w\rangle = 0$ para todo w; en particular, $\langle v, v\rangle$, co cal $v = 0$. Como as dimensión de E e E^* coinciden, ao ser inxectiva, ten que ser un isomorfismo. □

En particular, nun espazo euclidiano, calquera elemento do dual é da forma ϕ_w, para algún $w \in E$. No caso de dimensión infinita, nos chamados espazos de Hilbert, este resultado é moito menos trivial; é o que se chama o Teorema de Representación de Riesz.
Imos presentar agora a desigualdade de Cauchy–Schwarz, que relaciona o produto escalar de dous vectores co produto das súas normas.

Proposición 4.2. Para toda parella de vectores dun espazo euclidiano cúmprese a desigualdade

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|.$$

No caso hermítico, cómpre cambiar o valor absoluto do lado esquerdo pola norma do número complexo correspondente.

Demostración. Faremos a demostración no caso real, xa que o caso complexo é igual con pequenas modificacións. Se $v = 0$, ambos lados da desigualdade son 0, polo que o resultado é certo. En caso contrario, sexa $w = \frac{\langle u,v\rangle}{\langle v,v\rangle} v$. Entón,

$$\langle u - w, v\rangle = \langle u, v\rangle - \frac{\langle u, v\rangle}{\langle v, v\rangle}\langle v, v\rangle = 0.$$

Polo tanto, $\langle u - w, w\rangle = 0$. Entón,

$$\begin{aligned}\langle u, u\rangle &= \langle w + (u - w), w + (u - w)\rangle \\ &= \langle w, w\rangle + \langle u - w, u - w\rangle \\ &\geq \langle w, w\rangle = \frac{\langle u, v\rangle^2}{\langle v, v\rangle}.\end{aligned} .$$

Multiplicando ambos lados por $\langle v, v\rangle$ temos o resultado desexado.
No referente á igualdade, se os vectores son linealmente dependentes un dos dous é múltiplo do outro e é inmediato ver que se cumpre a igualdade. Para o recíproco, se $v = 0$ os vectores son dependentes; e se $v \neq 0$, as desigualdades anteriores implican que $\langle u - w, u - w\rangle = 0$, polo que $u = w$. Como w é múltiplo dos dous vectores, tamén son dependentes. □

A norma cumpre as seguintes propiedades.

Proposición 4.3. Sexa $(E, \langle\cdot, \cdot\rangle)$ un espazo vectorial euclidiano. A aplicación norma $v \mapsto \|v\| \colon E \to \mathbb{R}$ cumpre as seguintes propiedades:

(a) $\|v\| \geq 0$, con igualdade se, e soamente se, $v = 0$.

(b) $\|xv\| = |x|\|v\|$, onde $|x|$ é o valor absoluto en $\mathbb{R}$.

(c) $\|u + v\| \leq \|u\| + \|v\|$ (desigualdade triangular).

Demostración. (a) Temos que $\|v\| = 0$ se, e soamente se, $\|v\|^2 = \langle v, v\rangle = 0$, o que é equivalente a $v = 0$ polo feito de que o produto escalar é definido positivo.

(b) Temos que $\|xv\|^2 = \langle xv, xv\rangle = x^2\langle v, v\rangle$ e extraendo raíces cadradas temos a identidade do enunciado.

(c) Aplicando a desigualdade de Cauchy–Schwarz temos que

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle \\ &\leq \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2.\end{aligned}$$

Extraendo as raíces cadradas obtemos a desigualdade do enunciado.

□

Definición 4.5. A *distancia euclidiana* entre dous vectores dun espazo vectorial euclidiano defínese como a norma da súa diferenza, é dicir, $d(u,v) = \|u - v\|$. Do mesmo xeito, o ángulo que forman dous vectores non nulo dun espazo euclidiano é o ángulo $\theta \in [0, \pi]$ tal que

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}.$$

O ángulo está ben definido porque $\cos\theta \in [-1, 1]$ pola desigualdade de Cauchy–Schwarz. As seguintes propiedades da distancia son inmediatas a partir das correspondentes propiedades para a norma.

Proposición 4.4. Sexa $(E, \langle\cdot,\cdot\rangle)$ un espazo euclidiano. A aplicación $d\colon E \times E \to \mathbb{R}$ definida como $d(u,v) = \|u - v\|$ cumpre as seguintes propiedades.

(a) $d(u,v) \geq 0$, con igualdade se, e soamente se, $u = v$.

(b) $d(u,v) = d(v,u)$.

(c) $d(u,w) \leq d(u,v) + d(v,w)$, para todo $u, v, w \in E$ (desigualdade triangular).

É importante ter presente que as propiedades que establecemos para a norma e para a distancia úsanse a miúdo como definición; iso é así porque é posible traballar con distancias que non proveñen de normas, ou con normas que non proveñen de produtos escalares. Aquí, como estamos especialmente interesados en traballar con produtos escalares, case non consideraremos estas situacións, salvo nalgúns exemplos ao discutir as normas matriciais.
Como é habitual, interésanos considerar a representación matricial dun produto escalar.

Definición 4.6. Sexa E un K-espazo vectorial e sexa $\mathcal{B} = \{v_1, \ldots, v_n\}$ unha base. A *matriz dunha forma bilineal* $\phi\colon E \times E \to K$ nesa base defínese como

$$\mathrm{Mat}(\phi; \mathcal{B}) := (\phi(v_i, v_j))_{1\leq i,j\leq n} = \begin{pmatrix} \phi(v_1,v_1) & \phi(v_1,v_2) & \ldots & \phi(v_1,v_n) \\ \phi(v_2,v_1) & \phi(v_2,v_2) & \ldots & \phi(v_2,v_n) \\ \vdots & \vdots & & \vdots \\ \phi(v_n,v_1) & \phi(v_n,v_2) & \ldots & \phi(v_n,v_n) \end{pmatrix} \in \mathcal{M}_n(K).$$

O feito de que a forma bilineal sexa simétrica é equivalente a dicir que a matriz sexa simétrica en calquera base. Do mesmo xeito que a matriz dunha aplicación lineal permite calcular o valor que toma en función das coordenadas dos vectores, a matriz dunha forma bilineal permite calcular o valor que toma a forma en dous vectores en función das súas coordenadas. Se $u = \sum_{i=1}^n x_i v_i$ e $v = \sum_{i=1}^n y_i v_i$ son vectores de E expresados na base $\mathcal{B}$, o seu produto escalar é

$$\phi(u,v) = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

onde $a_{ij} = \phi(v_i, v_j)$.

Proposición 4.5. A relación entre as matrices dunha forma bilineal en dúas bases $\mathcal{B}_u$ e $\mathcal{B}_v$ do espazo vén dada por $B = P^t A P$, onde $A = \mathrm{Mat}(\phi; B_v)$, $B = \mathrm{Mat}(\phi; \mathcal{B}_u)$ e $P = (\mathbb{I})_{\mathcal{B}_u, \mathcal{B}_v}$ é a matriz de cambio de base.

Demostración. Sexan $A = (a_{ij})$, $B = (b_{ij})$ e $P = (p_{ij})$. Polo tanto, $P^t = (p_{ji})$. Sexan $AP = (\alpha_{ij})$ e $P^t AP = (\beta_{ij})$. Por definición, $a_{ij} = \phi(v_i, v_j)$, $b_{ij} = \phi(u_i, u_j)$ e $u_j = \sum_{i=1}^n p_{ij} v_i$. Entón,

$$b_{rs} = \phi(u_r, u_s) = \phi\Big(\sum_{i=1}^n p_{ir} v_i, \sum_{j=1}^n p_{js} v_j \Big) = \sum_{i=1}^n p_{ir} \sum_{j=1}^n a_{ij} p_{js},$$

e isto é igual a $\sum_{i=1}^n p_{ir} \alpha_{is} = \beta_{rs}$. □

A matriz do cambio de base podémola interpretar tamén en termos da matriz do cambio de base no dual. Deste xeito, $B = P^t AP$ pódese interpretar como segue:

- A matriz P pasa da base u á base v. É dicir, pódese interpretar A como a matriz dunha aplicación lineal

  $$E \longrightarrow E, \quad v \mapsto (\phi(v_1, v), \phi(v_2, v), \dots, \phi(v_n, v));$$

  ao aplicar o cambio de base de u a v estamos convertendo un vector da base u a un vector da base v ao que aplicar esa aplicación lineal.

- A matriz A é a matriz da aplicación lineal descrita no ítem anterior, pero tamén se pode ver como a matriz dunha aplicación lineal entre os espazos duais.

- A matriz P^t pasa da base dual v^* á base dual u^*. Cada unha das filas de A correspóndese coa aplicación $\psi_i \colon E \to K$, dada por $\psi_i(v) = \psi(v_i, v)$. Polo tanto, temos que realizar un cambio de base para pasar de expresar as formas na base v^* á base u^*, e a matriz do cambio é P^t.

Definición 4.7. Dúas matrices cadradas $A, B \in \mathcal{M}_n(K)$ dise que son *congruentes* se existe unha matriz invertible $P \in \mathrm{GL}_n(K)$ tal que $B = P^t AP$.

A relación de congruencia é unha relación de equivalencia. O seguinte resultado dános un criterio para determinar cando unha matriz é definida positiva. É importante incidir no feito de que para podelo aplicar cómpre que a matriz sexa simétrica.

Proposición 4.6 (Criterio de Sylvester)**.** Sexa $A = \mathrm{Mat}(\phi, \mathcal{B}_u) \in \mathcal{M}_n(\mathbb{R})$ a matriz dunha forma bilineal simétrica nun $\mathbb{R}$-espazo vectorial de dimensión finita n. A forma ϕ é definida positiva se, e soamente se, os seus menores principais dominantes son todos estritamente positivos: $\det((a_{ij})_{1 \le i,j \le k}) > 0$ para todo $k = 1, 2, \dots, n$.

Demostración. Para cada $k = 1, 2, \ldots, n$, escribimos A_k para referirnos á submatriz $(a_{ij})_{1 \leq i,j \leq k} \in \mathcal{M}_k(K)$, de maneira que os menores principais dominantes de A son os determinantes destas matrices.
Se $a_{11} \neq 0$ podemos construír unha nova base do espazo, $\mathcal{B}_v$, da maneira seguinte:

$$v_1 = u_1, \quad v_2 = u_2 - \frac{a_{12}}{a_{11}} u_1, \quad \ldots, \quad v_n = u_n - \frac{a_{1n}}{a_{11}} u_1.$$

A matriz de cambio de base é

$$P = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \ldots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Sexa $A = \mathrm{Mat}(\phi; \mathcal{B}_u)$ a matriz do produto escalar nunha base $\mathcal{B}_u$. Entón $a_{11} = \phi(u_1, u_1) > 0$ e pódese definir a base $\mathcal{B}_v$ do xeito que acabamos de describir. Sexa $B = P^t A P$. Para cada índice $j > 1$ tense que

$$\phi(v_1, v_j) = \phi\Big(u_1, u_j - \frac{a_{1j}}{a_{11}} u_1\Big) = a_{1j} - \frac{a_{1j}}{a_{11}} \cdot a_{11} = 0.$$

Polo tanto, a matriz B é da forma $B = \left(\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & B' \end{array}\right)$, onde $B' \in \mathcal{M}_{n-1}(\mathbb{R})$ é a matriz do produto escalar restrinxido ao subespazo $\langle v_2, \ldots, v_n \rangle$.
Para establecer que todo produto escalar cumpre a condición do criterio, procedemos por indución sobre a dimensión n. Se $n = 1$, entón $a_{11} = \phi(u_1, u_1) > 0$. Se o supomos certo para $n-1$, temos que é suficiente agora demostrar que o determinante da matriz dun produto escalar é sempre estritamente positivo, xa que os menores principais dominantes $\det(A_k)$ son os determinantes das matrices do produto escalar restrinxido aos subespazos $\langle u_1, \ldots, u_k \rangle$, e son automaticamente positivos pola hipótese de indución. Empregando de novo a hipótese, temos tamén que $\det(B') > 0$, xa que B' é a matriz da forma bilineal restrinxida a $\langle v_2, \ldots, v_n \rangle$; polo tanto $\det(B) = a_{11} \det(B') > 0$. Como $\det(P) = 1$ deducimos que $\det(A) = \det(B) > 0$.
Para demostrar o recíproco procedemos tamén por indución en n. Supoñamos que a matriz A cumpre a condición do criterio, é dicir, todos os seus menores principais dominantese son estritamente positivos. En particular, $a_{11} > 0$. Polo tanto, podemos definir unha base $\mathcal{B}_v$ igual que antes e matriz asociada ten tamén ceros no resto da primeira fila e da primeira columna. Esta matriz, á que podemos chamar B, cumpre a condición de Sylvester xa que $B_k = P_k^t A_k P_k$ para $k = 1, \ldots, n$; como $\det P_k = 1$, os menores principais dominantes de A e B coinciden. A matriz B' é a da restrición da forma bilineal ϕ ao subespazo $F = \langle v_2, \ldots, v_n \rangle$. A nova matriz tamén cumpre a condición de Sylvester xa que $a_{11} \det(B'_k) = \det(B_{k+1}) > 0$. Polo tanto, $\det(B'_k) > 0$ para todo $k = 1, \ldots, n-1$. Por hipótese de indución a restrición de ϕ ao subespazo F de dimensión $n-1$ é un produto escalar. Como $E = \langle v_1 \rangle \oplus F$, todo vector $v \in E$ é da forma $v = x v_1 + w$, con $x \in \mathbb{R}$ e $w \in F$, e só pode ser cero se $x = 0$ e $w = 0$. Entón, como $\phi(v_1, w) = 0$ temos que

$$\phi(v, v) = x^2 a_{11} + \phi(w, w) > 0$$

se $x \neq 0$ ou se $w \neq 0$, isto é, sempre que $v \neq 0$. $\square$

Exemplo. A matriz

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

é definida positiva porque os seus menores principais dominantes son 2, 3 e 4.

De cara a algunhas aplicacións, especialmente á análise matemática, convén ter en conta a notación de semidefinida positiva, isto é, $\langle v, v\rangle \geq 0$ para todo $v \in E$. Neste caso, non é certo o resulto análogo ao criterio de Sylvester (que diría que unha matriz é semidefinida positiva se os menores principais dominantes son todos maiores ou iguais que cero).
No caso complexo, estes resultados adáptanse de xeito inmediato. A única diferenza que hai que ter en conta é que a relación de congruencia escríbese como $B = P^*AP$, onde P^* é a trasposta da conxugada; a demostración é a mesma que no caso real, simplemente observando que ao aplicar a propiedade de sesquilinealidade temos que considerar o conxugado. Nalgúns textos, e en especial en física, é habitual escribir $P^\dagger$ en lugar de P^*.

4.2. Ortogonalidade e proxección ortogonal

Sexa $(E, \langle\cdot\rangle)$ un espazo vectorial euclidiano. O caso hermítico é totalmente análogo.

Definición 4.8. Dous vectores $u, v \in E$ dise que son *ortogonais* ou *perpendiculares* se $\langle u, v\rangle = 0$. Escribiremos $u \perp v$. Dous subconxuntos $S, T \subset E$ dise que son *ortogonais* se $u \perp v$ para todo $u \in S$ e $v \in T$; poremos $S \perp T$. Unha base do espazo E dise *ortogonal* se cada vector da base é ortogonal a todos os demais, e dise que é *ortonormal* se é ortogonal e todos os seus vectores son unitarios.

Nunha base ortonormal $\{v_1, \ldots, v_n\}$, as coordenadas dun vector $v = \sum_{i=1}^n x_i v_i$ cumpren que $x_i = \langle v, v_i\rangle$ e o produto escalar de dous vectores ou a norma dun vector calcúlanse igual que no caso do espazo euclidiano $\mathbb{R}^n$. É dicir, se $v = \sum_{i=1}^n x_i v_i$ e $w = \sum_{i=1}^n y_i v_i$, entón

$$\langle v, w\rangle = \sum_{i=1}^n x_i y_i, \quad \|v\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Imos describir agora o proceso coñecido como ortonormalización de Gram–Schmidt. É un método algorítmico que permite, a partir dunha base calquera, producir unha base ortonormal (e en particular, amosa que todo espazo euclidiano de dimensión finita ten unha base ortonormal).

Proposición 4.7. Sexa $\{u_1, \ldots, u_n\}$ unha base do espazo vectorial euclidiano $(E, \langle\cdot,\cdot\rangle)$. Os vectores definidos recursivamente pola fórmula

$$w_k = \frac{v_k}{\|v_k\|}, \quad v_k = u_k - \sum_{i=1}^{k-1} \langle u_k, w_i\rangle w_i$$

para $k = 1, \ldots, n$ son unha base ortonormal do espazo.

Demostración. Os vectores w_k son unha base, xa que se obteñen a partir da base u_k realizando transformacións elementais, e teñen norma 1 por construción. É suficiente, polo tanto, ver que son ortogonais, para o cal comprobaremos que cada un é ortogonal a todos os anteriores. Supoñamos que xa o comprobamos ata o vectores w_{k-1}. Para todo índice $j < k$ tense que:

$$\begin{aligned}\langle w_k, w_j\rangle = \frac{1}{\|v_k\|}\langle v_k, w_j\rangle &= \frac{1}{\|v_k\|}\Big\langle u_k - \sum_{i=1}^{k-1}\langle u_k, w_i\rangle w_i, w_j\Big\rangle \\ &= \frac{1}{\|v_k\|}\Big(\langle u_k, w_j\rangle - \sum_{i=1}^{k-1}\langle u_k, w_i\rangle\langle w_i, w_j\rangle\Big).\end{aligned}$$

Por hipótese de indución, cada factor $\langle w_i, w_j\rangle$ da suma é igual a $\delta_{i,j}$. Polo tanto, o valor da suma é $\langle u_k, w_j\rangle$ e concluímos que $\langle w_k, w_j\rangle = 0$. □

Exemplo. Sexa $\{u_1, u_2, u_3\}$ a base de $\mathbb{R}^3$ co produto escalar habitual dada por

$$u_1 = (2, 1, 2),\ u_2 = (0, 1, 1),\ u_3 = (0, 1, -1).$$

Seguindo o proceso de ortonormalización de Gram–Schmidt, temos que o primeiro vector é $w_1 = (2/3, 1/3, 2/3)$. Para o segundo vector, temos

$$v_2 = (0, 1, 1) - 1 \cdot (2/3, 1/3, 2/3) = (-2/3, 2/3, 1/3),$$

e como xa ten norma 1, $w_2 = (-2/3, 2/3, 1/3)$. Finalmente,

$$v_3 = (0, 1, -1) + 1/3 \cdot (2/3, 1/3, 2/3) - 1/3 \cdot (-2/3, 2/3, 1/3) = (4/9, 8/9, -8/9),$$

e normalizando, $w_3 = (1/3, 2/3, -2/3)$.

O seguinte resultado é consecuencia directa da proposición anterior.

Proposición 4.8. Todo espazo euclidiano de dimensión finita ten algunha base ortonormal. Ademais, toda base ortonormal dun subespazo pode ampliarse a unha base ortonormal de todo o espazo.

Demostración. A partir dunha base calquera pode obterse unha base ortonormal aplicando o algoritmo de Gram–Schmidt. Para ver a segunda afirmación, se $w_1, \ldots, w_r$ é unha base ortonormal do subespazo F, completámola polo teorema de Steinitz a unha base de todo o espazo engadindo vectores $w_{r+1}, \ldots, w_n$. Entón, se aplicamos o algoritmo de ortonormalización de Gram–Schmidt á base $\{w_1, \ldots, w_r, w_{r+1}, \ldots, w_n\}$, os primeiros r vectores quedan igual e os demais convértense en novos vectores que dan lugar a unha base ortonormal. □

Para traballar con matrices, é conveniente introducir algunhas notacións sobre matrices.

Definición 4.9. Unha matriz $A \in \mathcal{M}_n(K)$ dise que é *ortogonal* se $A^tA = \mathbb{I}_n$, ou alternativamente, se $A^{-1} = A^t$. As matrices ortogonais forman un grupo co produto, que se chama *grupo ortogonal* e que se denota por $\mathcal{O}_n(K)$.
No caso complexo, diremos que unha matriz $U \in \mathcal{M}_n(\mathbb{C})$ é *unitaria* se $U^*U = \mathbb{I}_n$, onde como é habitual U^* é a matriz complexa conxugada. As matrices unitarias forman un grupo co produto, que se chama *grupo unitario* e que se denota por $U_n(\mathbb{C})$.

O feito de que as matrices ortogonais formen un grupo é unha comprobación inmediata: se $AA^t = \mathbb{I}_n$ e $BB^t = \mathbb{I}_n$, entón $A^{-1} = A^t$ e $B^{-1} = B^t$. Multiplicando ambas igualdades temos que

$$(AB)^{-1} = B^{-1}A^{-1} = B^tA^t = (AB)^t,$$

e entón $(AB)^t(AB) = \mathbb{I}_n$. Para a inversa, está claro que se $A^{-1} = A^t$, tamén se ten que $A = (A^{-1})^t$.

Exemplo. Imos determinar as matrices de $\mathcal{M}_2(\mathbb{R})$ que son ortogonais. Escribindo

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

A condición de ortogonalidade di que

$$a^2 + c^2 = b^2 + d^2 = 1, \quad ab + cd = 0.$$

Podemos pór entón $a = \cos\alpha$, $c = \sin\alpha$, $b = \cos\beta$ e $d = \sin\beta$, onde $\alpha, \beta \in [0, 2\pi)$. De $ac + bd = 0$,

$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta) = 0,$$

polo que $\alpha - \beta = 3\pi/2$ ou $\alpha - \beta = \pi/2$. No primeiro caso,

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix};$$

en termos xeométricos, A correspóndese cun xiro de ángulo α. É dicir, se comezamos cun vector $v = (x, y)$ de norma 1, ao aplicar a matriz obtemos $v' = (x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha)$. Temos que $\|v'\| = 1$ e por outro lado, se φ é o ángulo que forman v e v,

$$\cos\varphi = x^2\cos\alpha - xy\sin\alpha + xy\sin\alpha + y^2\cos\alpha = (x^2 + y^2)\cos\alpha = \cos\alpha.$$

Como o coseno é inxectivo en $[0, \pi)$, temos que $\alpha = \varphi$, o que responde á idea de que a matriz é un xiro de ángulo α.
No segundo caso,

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

que se corresponde coa simetría con respecto a un eixe que pasa pola orixe (ou, alternativamente, a composición dunha simetría en torno ao eixe horizontal e un xiro de ángulo α).
O grupo $\mathcal{O}_2(\mathbb{R})$ está formado por matrices de determinante $+1$ ou -1. As matrices de $\mathcal{O}_2(\mathbb{R})$ con determinante $+1$, que son as do primeiro tipo que amosamos, forman un subgrupo que se chama o subgrupo especial ortogonal, e que se adoita denotar como $\mathrm{SO}_2(\mathbb{R})$.

Para unha matriz, a condición de ser ortogonal é equivalente a dicir que as súas columnas forman unha base ortonormal de $\mathbb{R}^n$. Máis en xeral, tense o seguinte resultado.

Proposición 4.9. Sexa $\mathcal{B}_w = \{w_1, \ldots, w_n\}$ unha base ortonormal do espazo euclidiano E. Entón, outra base $\mathcal{B}_u = \{u_1, \ldots, u_n\}$ é tamén ortonormal se, e soamente se, as matrices dos cambios de base son matrices ortogonais.

Demostración. A condición de que $\mathcal{B}_w$ sexa ortonormal quere dicir que a matriz do produto escalar nesa base sexa a identidade. Se lle chamamos U á matriz do produto escalar na base $\mathcal{B}_u$, entón $U = P^t \mathbb{I}_n P = P^t P$, onde P é a matriz do cambio de base de $\mathcal{B}_u$ a $\mathcal{B}_w$. Entón, $U = \mathbb{I}_n$ se, e soamente se, P e P^{-1} son ortogonais. □

Imos introducir agora a noción de proxección ortogonal. En xeral, en matemáticas, unha proxección é unha aplicación f tal que $f^2 = f$. Desde o punto de vista da diagonalización, tense que $X^2 - X = X(X-1)$ é un polinomio anulador de f, polo que o mínimo será un divisor seu. Automaticamente, iso dinos que f diagonaliza e que os únicos valores propios que pode ter son 0 e 1.

Definición 4.10. Sexa $F \subset E$ un subespazo vectorial ou, máis en xeral, un subconxunto calquera, dun espazo euclidiano. Entón o conxunto

$$F^\perp = \{v \in E \mid v \perp w \text{ para todo } w \in F\}$$

formado polos vectores do espazo que son ortogonais a todos os de F é un subespazo que se chama *ortogonal* do subespazo ou subconxunto F.

Observemos que se $F = \langle w_1, w_2, \ldots, w_r \rangle$, entón $v \in F^\perp$ se e soamente se $v \perp w_i$ para todo $i = 1, \ldots, r$. No espazo euclidiano habitual, se $F \subset \mathbb{R}^n$ é o subespazo xerado por unha familia de vectores $F = \langle x_1, \ldots, x_m \rangle$ con $x_i = (x_{i1}, \ldots, x_{in})$, o subespazo ortogonal $F^\perp$ é o subespazo das solucións do sistema de ecuacións lineais homoxéneo $\{\sum_{j=1}^n x_{ij} X_j = 0 \text{ para } i = 1, \ldots, m\}$. É dicir, o sistema con matriz

$$\begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}.$$

Podemos definir agora a noción de suma ortogonal.

Definición 4.11. A suma $F = F_1 + F_2$ de dous subespazos de E que sexan ortogonais chámase *suma ortogonal* e ás veces escríbese como $F_1 \perp F_2$. O caso de máis de dous espazos que sexan ortogonais cada un cos outros faise de forma análoga. Unha suma ortogonal é sempre unha suma directa.

Para xustificar a última observación, se $F = F_1 \perp \ldots \perp F_n$ e $w_1 + \ldots + w_n = 0$, entón considerando o produto escalar por w_j temos que

$$0 = \langle w_j, 0 \rangle = \langle w_j, \sum_{i=1}^n w_i \rangle = \langle w_j, w_j \rangle,$$

o que quere dicir que $w_j = 0$ para calquera j.

Proposición 4.10. Para todo subespazo $F \subset E$ dun espazo euclidiano de dimensión finita, o espazo E é a suma ortogonal $E = F \perp F^\perp$.

Demostración. A condición de definida positiva do produto escalar implica que a intersección $F \cap F^\perp$ é trivial, xa que se houbera un elemento en común v entón este cumpriría que $\langle v, v \rangle = 0$, e sabemos que $v = 0$ é o único para o cal iso é certo. Polo tanto, é suficiente ver que todo vector $v \in E$ é suma dun de F e doutro de $F^\perp$. Sexa $\{u_1, \ldots, u_r\}$ unha base ortonormal de F e sexa $\{u_{r+1}, \ldots, u_n\}$ unha ampliación a unha base ortonormal de todo E. Polo tanto, todo vector $v \in E$ escríbese como $v = \sum_{i=1}^n x_i u_i$, onde os coeficientes son $x_i = \langle v, u_i \rangle$. Polo tanto, v é a suma de $\sum_{i=1}^r x_i u_i \in F$ e $\sum_{i=r+1}^n x_i u_i \in F^\perp$. □

Definición 4.12. Sexa $F \subset E$ un subespazo vectorial dun espazo euclidiano. A *proxección ortogonal* do espazo E sobre o subespazo F é a aplicación lineal $\pi_F \colon E \to F$ definida como $\pi_F(v) = v_1$ se $v = v_1 + v_2$, con $v_1 \in F$ e $v_2 \in F^\perp$. Do mesmo xeito, a proxección ortogonal en $F^\perp$ é a aplicación $\pi_{F^\perp} \colon E \to F^\perp$ definida por $\pi_{F^\perp} = v_2$.

Proposición 4.11. A proxección ortogonal nun subespazo F dun vector $v \in E$ é o vector do subespazo que está máis próximo a v:

$$d(v, \pi_F(v)) = \text{mín}\{d(v, w) \mid w \in F\}.$$

Demostración. Como $E = F \perp F^\perp$, o vector $v \in E$ pódese escribir de forma única como $v = v_1 + v_2$, con $v_1 \in F$ e $v_2 \in F^\perp$. Entón, para todo vector $w \in F$ temos que

$$d(v,w)^2 = \langle v - w, v - w\rangle = \langle v_1 + v_2 - w, v_1 + v_2 - w\rangle = \langle v_1 - w, v_1 - w\rangle + \langle v_2, v_2\rangle.$$

O valor mínimo desta expresión dáse cando o primeiro sumando vale cero, que corresponde a $w = v_1$; este caso tense que $d(v, w) = \sqrt{\langle v_2, v_2\rangle} = \|\pi_{F^\perp}(v)\|$. □

A distancia $d(v, F)$ dun vector $v \in E$ a un subespazo $F \subset E$ defínese como a distancia mínima $\text{mín}\{d(v, w) \mid w \in F\}$ e a proposición anterior afirma que

$$d(v, F) = \|\pi_{F^\perp}(v)\| = \|v - \pi_F(v)\|.$$

Exemplo. En $\mathbb{R}^3$ consideramos o plano F dado por $z = 0$, que é un subespazo de dimensión 2. O seu ortogonal, $F^\perp$, é a liña xerada por $(0, 0, 1)$. Dado un vector $v = (a, b, c)$, a descomposición da proxección ortogonal é simplemente $v = (a, b, 0)+(0, 0, c)$. O resultado anterior dinos que $(a, b, 0)$ é o vector do plano $z = 0$ máis preto de v, e que polo tanto a distancia de v ao plano F é simplemente $|c|$.

O teorema da proxección ortogonal permítenos caracterizar o ortogonal dun subespazo dado a través de ecuacións. Do mesmo modo que antes discutimos que no caso dun subespazo dado en termos de xeradores o ortogonal se obtiña considerando as solucións do sistema dado polos coeficientes, o recíproco tamén é certo. É dicir, dado un subespazo en forma de ecuacións, os coeficientes das mesmas representan os xeradores do ortogonal.

Proposición 4.12. Consideramos o espazo euclidiano $\mathbb{R}^n$ co produto escalar estándar. Se $F \subset \mathbb{R}^n$ é o subespazo das solucións dun sistema homoxéneo de m ecuacións lineais,

$$F = \Big\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{j=1}^{n} a_{ij}x_j = 0 \text{ para } i = 1, \ldots, m\Big\},$$

entón o seu ortogonal $F^\perp$ é o subespazo xerado polos m vectores

$$v_1 = (a_{11}, a_{12}, \ldots, a_{1n}),\ \ldots,\ v_m = (a_{m1}, a_{m2}, \ldots, a_{mn}).$$

Demostración. Cada v_i é ortogonal a todo vector de F e polo tanto todos os v_i pertencen a $F^\perp$. Sexa r o rango do sistema, que é a dimensión do subespazo xerado polos vectores v_i. Entón, $\dim F = n-r$, e polo teorema da proxección ortogonal, $\dim F^\perp = r$. Como os dous espazos teñen a mesma dimensión e hai unha inclusión, teñen que ser iguais. □

Exemplo. Sexa $ax+by+cz = 0$ un plano en $\mathbb{R}^3$. Entón, o seu ortogonal é o subespazo xerado polo vector (a, b, c). Esta é unha noción moi habitual na análise matemática, onde se adoita falar de vector normal.

Outra consecuencia inmediata do resultado é que $(F^\perp)^\perp = F$. Para velo, sexa $v \in F$. Para $w \in F^\perp$ tense que $v \perp w$, polo que $v \in (F^\perp)^\perp$, co cal $F \subset (F^\perp)^\perp$. Pero ao mesmo tempo temos que $E = F \oplus F^\perp$ e $E = F^\perp \oplus (F^\perp)^\perp$, polo que se ten que $\dim F = n - \dim F^\perp = \dim(F^\perp)^\perp$. Polo tanto, os dous subespazos son iguais.
Imos discutir por último como calcular a proxección ortogonal a un subespazo.

Proposición 4.13. Sexa $\{v_1, \ldots, v_r\}$ unha base de $F \subset E$. Entón, para cada vector $u \in E$, as coordenadas da proxección ortogonal $\pi_F(u) = x_1 v_1 + \ldots + x_r v_r$ son as solucións do sistema lineal $Ax = b$, onde

$$A = \begin{pmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \ldots & \langle v_1, v_r\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \ldots & \langle v_2, v_r\rangle \\ \vdots & \vdots & & \vdots \\ \langle v_r, v_1\rangle & \langle v_r, v_2\rangle & \ldots & \langle v_r, v_r\rangle \end{pmatrix}, \quad b = \begin{pmatrix} \langle v_1, u\rangle \\ \langle v_2, u\rangle \\ \vdots \\ \langle v_r, u\rangle \end{pmatrix}.$$

Demostración. Sexa $\{v_{r+1}, \ldots, v_n\}$ unha base do subespazo ortogonal $F^\perp$. O teorema da proxección ortogonal afirma que $E = F \perp F^\perp$, e polo tanto $v_1, \ldots, v_n$ é unha base de todo o espazo. Podemos entón escribir $u = u_1 + u_2$, onde $u_1 = \sum_{i=1}^r x_i v_i \in F$ e $u_2 = \sum_{i=r+1}^n x_i v_i \in F^\perp$. A proxección ortogonal de u en F é o vector u_1. Para cada índice $j = 1, \ldots, r$ tense que

$$\langle v_j, u\rangle = \sum_{i=1}^n x_i \langle v_j, v_i\rangle = \sum_{i=1}^r \langle v_j, v_i\rangle x_i.$$

Considerando estas expresións para cada índice, recuperamos o sistema de ecuacións do enunciado. □

Exemplo. Consideremos $\mathbb{R}^4$ co produto escalar habitual, e sexa F o subespazo xerado polos vectores $v_1 = (1,1,1,1)$ e $v_2 = (1,1,-1,-1)$. Consideremos o vector $v = (1,0,0,0)$. Entón, a proxección de v en F vén dada por $x_1 v_1 + x_2 v_2$, onde

$$\begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Polo tanto, $x_1 = x_2 = \frac{1}{4}$ e a proxección ortogonal é o vector $(1/2, 1/2, 0, 0)$. En particular, a distancia de v a F é 1.

No caso particular no que $F = \langle v\rangle$, temos que

$$\pi_F(w) = \frac{\langle v, w\rangle}{\langle v, v\rangle} v.$$

Unha das aplicacións máis interesantes da proxección ortogonal é o chamado método dos mínimos cadrados. Sexan $A \in \mathcal{M}_{n\times m}(\mathbb{R})$ e $b \in \mathbb{R}^m$ un vector columna. O sistema lineal $Ax = b$ pode non ter solución; son os chamados sistemas sobredeterminados, que son sistemas que teñen máis ecuacións que incógnitas.

Proposición 4.14. Os vectores $x \in \mathbb{R}^n$ que minimizan o valor de $\|Ax - b\|$ son as solución do sistema lineal $A^t Ax = A^t b$.

Demostración. O vector $y = Ax \in \mathbb{R}^m$ que minimiza a norma $\|y - b\|$ é a proxección ortogonal de $b \in \mathbb{R}^m$ no subespazo $F \subset \mathbb{R}^m$ formado polos vectores Ax, e está xerado polos vectores columna de A.
O resultado anterior sobre o cálculo da proxección ortogonal tamén funciona cando os vectores v_i son unha familia de xeradores de F e non necesariamente unha base. Neste caso, as columnas $v_1, \ldots, v_n$ da matriz A son xeradores do subespazo F. Entón a matriz $(\langle v_i, v_j \rangle)_{1 \leq i,j \leq n}$ é a matriz $A^t A \in \mathcal{M}_n(\mathbb{R})$ e o vector $(\langle v_i, v \rangle) \in \mathbb{R}^n$ é o vector $A^t b$. Polo tanto, o enunciado séguese directamente do resultado anterior. □

Exemplo. Consideremos os puntos $(0, 0)$, $(1, 1)$, $(2, 3)$, $(3, 4)$ e $(4, 4)$. A recta de regresión é a recta da forma $y = ax + b$ que mellor aproxima os puntos dados. Polo tanto, queremos buscar os coeficientes (α, β) que minimicen o erro ao resolver

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

Usando os métodos dos mínimos cadrados, isto correspóndese con resolver o sistema

$$\begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} 35 \\ 12 \end{pmatrix},$$

de onde resulta $\beta = 1{,}1$ e $\alpha = 0{,}2$, polo cal a recta buscada é $y = 1{,}1x + 0{,}2$.

4.3. Endomorfismos autoadxuntos e normais

Imos comezar esta sección introducindo a noción de aplicación lineal adxunta. Informalmente, é o endomorfismo que nos permite *pasar* da primeira á segunda compoñente da forma bilineal. Sexan E e F dous espazos vectoriais euclidianos ou hermíticos; escribiremos $\langle \cdot, \cdot \rangle_E$ e $\langle \cdot, \cdot \rangle_F$ para os correspondentes produtos escalares, aínda que cando a elección do espazo vectorial sexa clara polo contexto eliminaremos o subíndice.

Definición 4.13. Sexa $f \colon E \to F$ unha aplicación lineal. A *aplicación adxunta* de f, ou simplemente *adxunta* de f, é a aplicación $f^* \colon F \to E$ tal que

$$\langle f(v), w \rangle_F = \langle v, f^*(w) \rangle_E$$

para $v \in E$ e $w \in F$. No caso particular no que $F = E$ e $f \in \mathrm{End}(E)$, a adxunta de f é simplemente a aplicación que cumpre que

$$\langle f(v), w \rangle_E = \langle v, f^*(w) \rangle_E$$

para $v, w \in E$.

Proposición 4.15. Sexa $f \colon E \to F$ unha aplicación lineal. A aplicación adxunta $f^* \colon F \to E$ está ben definida.

Demostración. Sexa $w \in F$ e consideremos a aplicación lineal en E que envía $v \in E$ a $\langle f(v), w \rangle_F$; esta aplicación lineal depende de w e de f. Polo teorema de representación de Riesz, existe un único vector $w' \in E$ de maneira que a aplicación se corresponde a facer o produto escalar con w', é dicir, $\langle f(v), w \rangle_F = \langle v, w' \rangle_E$ para todo $v \in E$. Este vector w' é o que chamaremos $f^*(w)$. □

Noutras palabras, $f^*(w)$ é o único vector en E que cumpre $\langle f(v), w\rangle_F = \langle v, f^*(w)\rangle_E$ para todo $v \in E$.

Exemplo. Sexa $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ dada por $f(x, y, z) = (y + 3z, 2x)$. A adxunta de f é

$$f^*(a, b) = (2b, a, 3a).$$

Para demostralo, facemos

$$\begin{aligned}
\langle (x_1, x_2, x_3), f^*(y_1, y_2)\rangle &= \langle f(x_1, x_2, x_3), (y_1, y_2)\rangle \\
&= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\rangle \\
&= x_2y_1 + 3x_3y_1 + 2x_1y_2 \\
&= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\rangle.
\end{aligned}$$

Polo tanto, $f^*(y_1, y_2) = (2y_2, y_1, 3y_1)$.

Se f é unha aplicación lineal, cúmprese que f^* tamén é unha aplicación lineal (comprobación inmediata a partir da definición, vendo que $f^*(w_1 + w_2) = f^*(w_1) + f^*(w_2)$ e $f^*(\lambda w) = \lambda f^*(w)$). Por exemplo, no caso da suma temos que

$$\begin{aligned}
\langle v, f^*(w_1 + w_2)\rangle_E = \langle f(v), w_1 + w_2\rangle_F &= \langle f(v), w_1\rangle_F + \langle f(v), w_2\rangle_F \\
&= \langle v, f^*(w_1)\rangle_E + \langle v, f^*(w_2)\rangle_E \\
&= \langle v, f^*(w_1) + f^*(w_2)\rangle_E.
\end{aligned}$$

O caso do produto por escalares é análogo.
O seguinte resultado afirma que a matriz da aplicación adxunta coincide coa transposta.

Proposición 4.16. Sexa $f\colon E \to F$ unha aplicación lineal entre dous espazos vectoriais reais. Sexa $\{u_1, \dots, u_n\}$ unha base ortonormal de E e $\{v_1, \dots, v_m\}$ unha base ortonormal de F. Sexa A a matriz de f con respecto a esas bases e B a matriz de f^* con respecto ás mesmas bases. Entón, $B = A^t$.
No caso complexo, e coas mesmas notacións, $B = A^*$ xa que a forma bilineal non é simétrica, senón hermítica, e iso introduce a conxugación complexa.

Demostración. A k-ésima columna de A correspóndese a escribir $f(u_k)$ como unha combinación lineal dos vectores $v_1, \dots, v_m$. Como esta é unha base ortonormal temos que

$$f(u_k) = \sum_{i=1}^{m} \langle f(u_k), v_i\rangle_F v_i.$$

Polo tanto, a posición (j, k) da matriz A é $\langle f(u_k), v_j\rangle_F$. Procedendo da mesma maneira, pero intercambiando os papeis da base $\{u_1, \dots, u_n\}$ e $\{v_1, \dots, v_m\}$ vemos que a entrada (j, k) da matriz B é $\langle f^*(v_k), u_j\rangle_E$. Agora ben,

$$\langle f^*(v_k), u_j\rangle_E = \langle v_k, f(u_j)\rangle_F,$$

que se corresponde á entrada (k, j) da matriz A. O caso complexo é totalmente análogo. □

Exemplo. No exemplo anterior no que $f(x, y, z) = (y + 3z, 2x)$, a matriz asociada é

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

Vimos que a adxunta é $f^*(a,b) = (2b, a, 3a)$, que ten por matriz

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

Este resultado permítenos pensar a adxunta dunha aplicación lineal nunha base ortonormal como a súa transposta, e deste xeito podemos establecer algunhas propiedades moi naturais. Por exemplo, $(f^*)^* = f$, ou $(fg)^* = g^*f^*$, para $f, g \in \operatorname{End}(E)$.
O caso de maior interese no contexto desta materia dáse cando unha aplicación lineal coincide coa súa adxunta. Desde o punto de vista da estrutura dos endomorfismos, poderemos asegurar que diagonaliza.

Definición 4.14. Un endomorfismo $f \in \operatorname{End}(E)$ dise que é *autoadxunto* se $f = f^*$. No caso real, as aplicacións lineais autoadxuntas tamén se chaman *simétricas*, e no caso complexo, *hermíticas*.

É importante observar, seguindo o razoamento de resultados anteriores, que a matriz dunha aplicación simétrica nunha base ortonormal é simétrica. Se $\{v_1, \dots, v_n\}$ é a base ortonormal, entón na entrada (i,j) da matriz correspondente ao endomorfismo nesa base temos o elemento $\langle v_i, f(v_j)\rangle$; na entrada (j,i) temos $\langle v_j, f(v_i)\rangle$, e ambos valores coinciden. En cambio, se a base non é ortonormal, iso non ten por que ser certo.
O obxectivo desta sección é dar condicións necesarias para que un endomorfismo diagonalice nunha base ortonormal. Imos comezar estudando o caso real, onde o resultado se basea na seguinte proposición, que afirma que o polinomio característico dunha matriz simétrica descompón completamente sobre os reais. No caso de dimensión 2, iso pode verse directamente, observando que se $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$, entón $\operatorname{Char}(A;X) = X^2 - (a+c)X^2 + ac - b^2$, polo que o discriminante é

$$(a+c)^2 - 4ac + 4b^2 = (a-c)^2 + (2b)^2 \geq 0.$$

Proposición 4.17. Toda matriz simétrica $A \in \mathcal{M}_n(\mathbb{R})$ ten todos os valores propios reais, é dicir, o polinomio característico $\operatorname{Char}(A;X) \in \mathbb{R}[X]$ ten tantas raíces como o seu grao, contando multiplicidades.

Demostración. A matriz pódese interpretar como unha matriz con entradas complexas, de maneira que polo teorema fundamental da álxebra, sabemos que o polinomio característico ten algunha raíz complexa e polo tanto existe algún valor propio complexo λ. Sexa $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ un vector propio de valor propio λ, $Az = \lambda z$. Tomando conxugados temos que $A\bar{z} = \bar{\lambda}\bar{z}$, onde $\bar{z}$ é o vector que se obtén conxugando cada compoñente de z. O número $\sum_{i=1} |z_i|^2 = z^t\bar{z} = \bar{z}^t z$ é un real positivo. Usando agora que A é simétrica, temos que

$$\lambda \sum_{i=1}^{n} |z_i|^2 = \bar{z}^t \lambda z = \bar{z}^t A z = (\bar{z}^t A z)^t = z^t A^t \bar{z} = z^t A \bar{z} = \bar{\lambda} z^t \bar{z} = \bar{\lambda} \sum_{i=1}^{n} |z_i|^2.$$

Polo tanto, $\lambda = \bar{\lambda}$ e λ é un número real. $\square$

O seguinte teorema é un dos resultados fundamentais do tema, e asegura que toda matriz simétrica con entradas reais diagonaliza nunha base ortonormal. A demostración é consecuencia de dous resultados anteriores: o teorema da proxección ortogonal e a proposición que afirma que o polinomio característico dunha matriz simétrica descompón completamente sobre os reais. O nome de *teorema espectral* débese a que en certas áreas das matemáticas é frecuente utilizar o termo *espectro* para refiririse ao conxunto de valores propios dunha aplicación lineal.

Teorema 4.1 (Teorema dos endomorfismos simétricos, tamén chamado teorema espectral)**.** Todo endomorfismo simétrico dun espazo vectorial euclidiano diagonaliza e admite unha base de vectores propios que, ademais, é unha base ortonormal.

Demostración. Procedemos por indución sobre a dimensión n do espazo. Se $n = 1$ é evidente, xa que todo vector non nulo é un vector propio que, ao normalizalo, dá lugar a unha base ortonormal. Supoñamos demostrado o resultado para dimensión $n-1$ e sexa $f \in \mathrm{End}(E)$ un endomorfismo simétrico dun espazo euclidiano de dimensión n. Probamos con anterioridade que o endomorfismo ten valores propios reais. Sexa $\lambda \in \mathbb{R}$ un valor propio e sexa $v \in E$ un vector propio de valor propio λ. Podemos supor, dividindo pola norma, que é un vector unitario. O teorema da proxección ortogonal asegura que $\langle v\rangle \oplus \langle v\rangle^\perp = E$. Afirmamos que o subespazo $\langle v\rangle^\perp$ é invariante polo endomorfismo f; en efecto, como f é simétrico, se $u \in \langle v\rangle^\perp$, entón $\langle u, v\rangle = 0$, polo que

$$\langle f(u), v\rangle = \langle u, f(v)\rangle = \langle u, \lambda v\rangle = \lambda\langle u, v\rangle = 0;$$

iso quere dicir que $f(u) \in \langle v\rangle^\perp$. A restrición do endomorfismo f ao subespazo $\langle v\rangle^\perp$ de dimensión $n-1$ tamén é un endomorfismo simétrico. Pola hipótese de indución, ten unha base ortonormal de vectores propios. Engadindo v a esta base tense unha base ortonormal de vectores propios de todo o espazo. □

O resultado anterior garantiza que unha matriz simétrica sempre vai diagonalizar, e ademais farao nunha base ortonormal.

Exemplo. Consideremos a matriz

$$A = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 5 & -1 \\ -1 & -1 & 5 \end{pmatrix}.$$

O seu polinomio característico é $\mathrm{Char}(A; X) = -(X-6)^2(X-3)$. O teorema espectral asegura que diagonaliza nunha base ortonormal, e en particular, que vectores propios de valores propios diferentes son ortogonais. Polo tanto, o polinomio mínimo será $\mathrm{Min}(A; X) = (X-6)(X-3)$. O subespazo propio para o valor propio 3 é $\langle(1,1,1)\rangle$. O subespazo propio para o valor propio 6 é $\{(x,y,z) \in \mathbb{R}^3 \mid x+y+z = 0\}$. Porén, é importante observar que con basta con coller unha base calquera do subespazo para ter unha base ortogonal: por exemplo, se tomamos os vectores $(1,-1,0)$ e $(1,0,-1)$, claramente non son perpendiculares. Si podemos coller dous vectores arbitrarios e aplicar Gram–Schmidt. Nese caso, o primeiro vector sería $v_1 = (1/\sqrt{2}, -1/\sqrt{2}, 0)$ e o segundo, $v_2 = (1/\sqrt{6}, 1/\sqrt{6}, -2/\sqrt{6})$. Como v_3 temos simplemente o vector propio de valor propio 3 e norma 1, $v_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$.

Un corolario importante do teorema espectral real é o seguinte.

Corolario 4.1. Unha matriz simétrica A con entradas reais ten valores propios non negativos se, e soamente se, existe unha matriz con entradas reais B tal que $A = B^tB$. Ademais, todos os valores propios son positivos se, e soamente se, B é non singular.

Demostración. Se A é simétrica, polo teorema espectral existe unha matriz ortogonal P e unha matriz diagonal D de maneira que $A = PDP^t$. Se todos os valores propios son non negativos, podemos considerar $D^{1/2}$, a matriz diagonal que ten por entradas as raíces cadradas de cada un dos valores propios de D. Entón,

$$A = PDP^t = PD^{1/2}D^{1/2}P^t = (D^{1/2}P^t)^t(D^{1/2}P^t).$$

Ademais, os valores propios son non cero se, e soamente se, as súas raíces cadradas son non cero.
Para o recíproco, supoñamos que $A = B^tB$. Nese caso, todos os valores propios de A son non negativos porque calquera vector propio v de valor propio λ cumpre

$$\lambda = \frac{v^tAv}{v^tv} = \frac{v^tB^tBv}{v^tv} = \frac{\|Bv\|^2}{\|v\|^2} \geq 0.$$

□

En xeral, os endomorfismos simétricos terán unha gran importancia no estudo da xeometría. Por exemplo, no estudo das superficies hai dúas aplicacións simétricas que fan un papel moi importante no estudo da curvatura, e que son a *primeira forma fundamental* e a *segunda forma fundamental*.
No caso complexo, o resultado análogo require introducir outra clase de endomorfismos.

Definición 4.15. Nun espazo euclidiano ou hermítico, un endomorfismo $f \in \mathrm{End}(E)$ dise que é *normal* se conmuta co seu adxunto, isto é $ff^* = f^*f$.

Algúns exemplos de endomorfismos normais son os seguintes:

- As aplicacións lineais unitarias, é dicir, aquelas para as que $f^* = f^{-1}$.
- As aplicacións lineais hermíticas, aquelas para as que $f^* = f$.
- As aplicación lineais antihermíticas, que son as cumpren $f^* = -f$.
- As aplicacións lineais positivas, aquelas para as que $f = gg^*$.

De cara a establecer un análogo do teorema espectral no caso complexo en termos de operadores normais, necesitamos o seguinte resultado. Sobre os reais, o resultado correspondente afirmaría que se E é un $\mathbb{R}$-espazo euclidiano e f é un endomorfismo simétrico tal que $\langle f(v), v\rangle = 0$ para todo $v \in E$, entón $f = 0$. Non imos facer a demostración, porque é análoga á do caso complexo que agora discutimos e que logo empregaremos.

Proposición 4.18. Sexa E un espazo hermítico e $f \in \mathrm{End}(E)$ de maneira que

$$\langle f(v), v\rangle = 0$$

para todo $v \in E$. Entón $f = 0$.

Demostración. Temos que

$$\begin{aligned}\langle f(u), w\rangle = \frac{1}{4}\Big(&\langle f(u+w), u+w\rangle - \langle f(u-w), u-w\rangle \\ &+ i\langle f(u+iw), u+iw\rangle - i\langle f(u-iw), u-iw\rangle.\Big)\end{aligned}$$

Para comprobar esta igualdade, podemos desenvolver termo a termo e vemos que

$$\begin{aligned}
\langle f(u+w), u+w\rangle &= \langle f(u), u\rangle + \langle f(u), w\rangle + \langle f(w), u\rangle + \langle f(w), w\rangle \\
-\langle f(u-w), u-w\rangle &= -\langle f(u), u\rangle + \langle f(u), w\rangle + \langle f(w), u\rangle - \langle f(w), w\rangle \\
i\langle f(u+iw), u+iw\rangle &= i\langle f(u), u\rangle + \langle f(u), w\rangle - \langle f(w), u\rangle + i\langle f(w), w\rangle \\
-i\langle f(u-iw), u-iw\rangle &= -i\langle f(u), u\rangle + \langle f(u), w\rangle - \langle f(w), u\rangle - i\langle f(w), w\rangle.
\end{aligned}$$

Sumando e dividindo por 4, queda precisamente $\langle f(u), w\rangle$.
Entón, como cada termo ao lado dereito da igualdade é da forma $\langle f(v), v)\rangle$, temos que $\langle f(u), w\rangle$ para todo $u, w \in E$. Polo tanto, $f = 0$. □

A idea da demostración é a mesma que a que se emprega no clásico exercicio de calcular a suma $\sum_{j=0}^{n} \binom{4n}{4j}$, no que se considera unha suma sobre as raíces cuartas da unidade que faga que das catro clases de termos correspondentes aos resto módulo 4 só quede unha e as outras se cancelen.
O seguinte resultado vai ser esencial para demostar o teorema espectral complexo. Afirma que unha aplicación lineal e a súa adxunta comparten un conxunto de vectores propios, e que os valores propios son os conxugados.

Proposición 4.19. Unha aplicación lineal é normal se e soamente se, para todo $v \in E$, se cumpre que $\|f(v)\| = \|f^*(v)\|$. Se f é normal e $v \in E$ é un vector propio de f con valor propio λ, entón v tamén é un vector propio de f^* con valor propio $\bar{\lambda}$.

Demostración. Para a primeira parte, observamos que f é normal se, e soamente se, $f^*f - ff^* = 0$. Iso é equivalente a $\langle (f^*f - ff^*)v, v\rangle = 0$ para todo $v \in E$. Polo tanto, unha aplicación lineal é normal se, e soamente se,

$$\langle f^*f(v), v\rangle = \langle ff^*(v), v\rangle$$

para todo $v \in E$. Usando a definición de adxunta, iso é equivalente a $\|f(v)\|^2 = \|f^*(v)\|^2$.
Para a segunda parte é suficiente con observar que $f - \lambda\,\mathrm{Id}$ tamén é normal, xa que

$$\begin{aligned}
(f - \lambda\,\mathrm{Id})(f^* - \bar{\lambda}\,\mathrm{Id}) &= ff^* - \lambda f^* - \bar{\lambda} f + \mathrm{Id} \\
&= f^*f - \lambda f^* - \bar{\lambda} f + \mathrm{Id} \\
&= (f^* - \bar{\lambda}\,\mathrm{Id})(f - \lambda\,\mathrm{Id}).
\end{aligned}$$

Entón, tense que

$$0 = \|(f - \lambda\,\mathrm{Id})v\| = \|(f - \lambda\,\mathrm{Id})^* v\| = \|(f^* - \bar{\lambda}\,\mathrm{Id})v\|,$$

polo que v é un vector propio de f^* con valor propio $\bar{\lambda}$, como se quería. □

Teorema 4.2 (Teorema espectral complexo)**.** Sexa E un espazo hermítico de dimensión finita, e sexa $f \in \mathrm{End}(E)$. Entón, se f é normal, existe unha base ortonormal de E formada por vectores propios de f.

Demostración. Consideremos a base de Jordan de f (que sempre existe pois sobre os complexos o polinomio característico sempre descompón completamente), que en termos matriciais se representa a través dunha matriz triangular superior. Aplicando o proceso de ortogonalización de Gram–Schmidt aos vectores desa base, temos que existe unha base ortonormal $\{v_1, \ldots, v_n\}$ para a cal a matriz $A = (a_{ij})$ é triangular superior.

Isto é consecuencia do feito de que, para calquera $k \in \{1, 2, \ldots, n\}$, v_k só depende dos k primeiros vectores da base, polo que $f(v_k)$ tamén; de aquí, pódese deducir que $f(v_k)$ é unha combinación lineal de $v_1, \ldots, v_k$.
Agora podemos probar o resultado por indución, que é inmediato para $n = 1$.
Imos demostrar que A é de feito diagonal, o que quererá dicir que os vectores que forman a base son vectores propios de f. Como $a_{ij} = 0$ para $i > j$, temos que $f(v_1) = a_{11}v_1$ e $f^*(v_1) = \sum_{k=1}^{n} \bar{a}_{1k}v_k$. Empregando o resultado anterior e o feito de que a base $\{v_1, \ldots, v_n\}$ é ortonormal,

$$|a_{11}|^2 = \|a_{11}v_1\|^2 = \|f(v_1)\|^2 = \|f^*(v_1)\|^2 = \|\sum_{k=1}^{n} \bar{a}_{1k}v_k\|^2 = \sum_{k=1}^{n} |a_{1k}|^2.$$

Polo tanto, $|a_{12}| = \ldots = |a_{1n}| = 0$, e temos entón que o subespazo xerado por $\{v_2, \ldots, v_n\}$ é invariante por f. Entón f é unha aplicación lineal normal nun espazo de dimensión $n - 1$, e pola hipótese de indución a matriz de f é diagonal. □

O recíproco tamén é certo. Se existe unha base ortonormal formada por vectores propios de f, entón f é normal. Para demostralo, sexa $\{e_1, \ldots, e_n\}$ a base, e tense que a matriz correspondente, á que chamaremos A, é diagonal. Ademais, a matriz de f^* é A^* e tamén é diagonal, xa que vimos que os e_i son tamén vectores propios de f^* (sendo os valores propios correspondentes os conxugados). Polo tanto, $ff^* = f^*f$ xa que as matrices correspondentes conmutan.

4.4. Endomorfismos ortogonais e unitarios

Definición 4.16. Sexa E un espazo euclidiano (ou hermítico) e $f \in \mathrm{End}(E)$ un endomorfismo. Dicimos que f é unha *isometría* se conserva o produto escalar (ou hermítico), é dicir, se

$$\langle f(u), f(v)\rangle = \langle u, v\rangle$$

para todo $u, v \in E$.
No caso real, as isometrías adoitan chamarse *aplicación ortogonais*, e no caso complexo, *aplicación unitarias*.

De feito, non é mester esixir que f sexa un endomorfismo. Da condición $\langle f(u), f(v)\rangle = \langle u, v\rangle$ pódese deducir automaticamente que f é lineal. Por simplicidade, imos comezar centrándonos no caso real.

Proposición 4.20. Sexa E un espazo vectorial real e $f\colon E \to E$ unha aplicación que cumpre que $\langle f(u), f(v)\rangle = \langle u, v\rangle$ para todo $u, v \in E$. Entón f é lineal.

Demostración. Hai que demostrar que $f(u+v)-f(u)-f(v) = 0$ e que $f(\lambda v)-\lambda f(v) = 0$, para calquera $u, v \in E$ e $\lambda \in \mathbb{R}$. Imos facer aquí unicamente o caso do produto por escalares. En ambos casos, a demostración consiste en ver que o produto escalar por el mesmo é 0:

$$\begin{aligned}\langle f(\lambda v) - \lambda f(v), f(\lambda v) - \lambda f(v)\rangle &= \langle f(\lambda v), f(\lambda v)\rangle + \lambda^2\langle f(v), f(v)\rangle - 2\lambda\langle f(\lambda v), f(v)\rangle \\ &= \langle \lambda v, \lambda v\rangle + \lambda^2\langle v, v\rangle - 2\lambda\langle \lambda v, v\rangle \\ &= \lambda^2\langle v, v\rangle + \lambda^2\langle v, v\rangle - 2\lambda^2\langle v, v\rangle = 0.\end{aligned}$$

□

Proposición 4.21. Sexa E un espazo vectorial real e sexa $f\colon E \to E$. As dúas seguintes condicións son equivalentes.

(a) f é unha isometría, é dicir, $\langle u, v\rangle = \langle f(u), f(v)\rangle$, para todo $u, v \in E$.

(b) f conserva normas, é dicir, $\|v\| = \|f(v)\|$, para todo $v \in E$.

Demostración. Está claro que se conserva produtos escalares, entón tamén conserva normas, xa que $\|v\|^2 = \langle v, v\rangle$. Para o recíproco, é suficiente observar que

$$2\langle u, v\rangle = \|u + v\|^2 - \|u\|^2 - \|v\|^2.$$

□

A seguinte proposición resume as propiedades máis importantes das isometrías.

Proposición 4.22. Sexa $f\colon E \to E$ unha isometría. Entón:

(a) f conserva ángulos.

(b) Se λ é un valor propio real de f, entón $\lambda = \pm 1$. En particular, f é bixectiva.

(c) Se F é un subespazo f-invariante, entón $F^\perp$ é un subespazo f-invariante.

(d) Se A é a matriz dunha isometría nunha base ortonormal, entón A é unha matriz ortogonal. En particular, $\det(A) = \pm 1$.

Demostración. (a) O ángulo vén dado como o cociente dun produto escalar e o produto de dúas normas, e todas as cantidades se conservan por unha isometría.

(b) Sexa λ un valor propio de f. Se $v \neq 0$ é un vector propio de valor propio λ, entón

$$\langle v, v\rangle = \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda^2 \langle v, v\rangle.$$

Dividindo por $\|v\|^2$ dedúcese que $\lambda^2 = 1$, polo que $\lambda = \pm 1$. A aplicación é bixectiva porque 0 non é un valor propio.

(c) Temos que se $v \in F$, $f(v) \in F$ por ser F invariante. Como f é bixectiva, a restrición $f|F\colon F \to F$ é unha aplicación lineal que conserva os produtos escalares, e $f^{-1}(v) \in F$ para todo $v \in F$. Dado un vector $u \in E$, temos que $u \in F^\perp$ se, e soamente se $\langle u, v\rangle = 0$ para todo $v \in F$. Polo tanto, se $u \in F^\perp$ temos que

$$\langle f(u), v\rangle = \langle f(u), f(f^{-1}(v))\rangle = \langle u, f^{-1}(v)\rangle = 0,$$

para todo $v \in F$. Polo tanto, $u \in F^\perp$ implica que $f(u)$ tamén está en $F^\perp$.

(d) Como f é bixectiva, podemos considerar a súa inversa f^{-1}. Temos que $\langle f(u), v\rangle = \langle u, f^{-1}(v)\rangle$, o que proba que a adxunta de f é igual á súa inversa. En termos matriciais (sempre que a base sexa ortonormal), iso quere dicir que $A^t = A^{-1}$, que é a definición de matriz ortogonal.

□

No caso complexo, se f é unha isometría, tense que existe unha base ortonormal de vectores propios de E de maneira que os correspondentes valores propios teñen módulo 1.

Proposición 4.23. Sexa f unha aplicación unitaria. Entón, f diagonaliza nunha base de vectores propios e os valores propios de f teñen módulo 1.

Demostración. O feito de que f diagonalice é unha consecuencia directa do teorema espectral complexo, xa que unha aplicación lineal unitaria é tamén normal. Para ver que o módulo dun valor propio é 1, procedemos como no segundo epígrafe da proposición anterior e temos que $|\langle v, v\rangle| = |\lambda|^2 \cdot |\langle v, v\rangle|$, polo que o resultado é certo. Podemos comprobar que se u e v son vectores propios correspondentes a vectores propios diferentes λ e μ, temos que

$$\langle u, v\rangle = \langle f(u), f(v)\rangle = \lambda\bar{\mu}\langle u, v\rangle.$$

Se $\langle u, v\rangle \neq 0$, entón $\lambda\bar{\mu} = 1$, o que quere dicir que $\bar{\mu}^{-1} = \lambda$, e de aquí séguese que $\mu = \lambda$, xa que o inverso do conxugado dun complexo de módulo 1 é el mesmo. En calquera caso, esta comprobación da ortogonalidade non tería sido necesaria, xa que é tamén consecuencia do teorema espectral. □

O estudo das isometrías será especialmente relevante en xeometría, ao traballarmos os movementos do espazo euclidiano. No caso real, porén, non é certo que toda isometría diagonalice. Por exemplo, sexa

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

É doado ver que a matriz representa unha isometría xa que conserva normas, é dicir, se $v = (x, y)$, entón

$$Av = (x\cos\alpha + y\sin\alpha, -x\sin\alpha + y\cos\alpha),$$

e cúmprese que $\|Av\| = \sqrt{x^2 + y^2}$. Por outro lado,

$$\operatorname{Char}(A; X) = X^2 - 2\cos\alpha X + 1,$$

que é unha ecuación cuadrática con discriminante $\Delta = 4(\cos^2\alpha - 1) \leq 0$, sendo 0 se, e soamente se, $\cos\alpha = \pm 1$. É dicir, en calquera outro caso o polinomio característico non ten raíces reais e polo tanto a matriz non pode diagonalizar.
O seguinte resultado proba que, en certa maneira, o exemplo anterior permite entender todos os casos nos que as isometrías non diagonalizan.

Teorema 4.3. Sexa E un espazo vectorial euclidiano e $f \in \operatorname{End}(E)$ unha aplicación ortogonal. Existe unha base ortonormal de E na que a matriz correspondente a f é unha matriz diagonal por bloques con elementos iguais a 1 e -1 na diagonal e logo bloques A_i da forma $\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$, con $a_i^2 + b_i^2 = 1$.

Demostración. Antes de comezar a demostración, imos enumerar as propiedades anteriores que se precisan e que se empregarán ao longo da proba.

(i) Os únicos valores propios reais dunha isometría son 1 e -1.

(ii) Unha isometría é unha aplicación normal que, polo tanto, diagonaliza nunha base ortonormal sobre os complexos.

(iii) Os valores propios complexos dunha isometría teñen módulo 1.

(iv) Dous vectores propios dunha isometría con valores propios diferentes son ortogonais.

Sexan E_1 e E_{-1} os subespazos de vectores propios correspondentes aos valores propios 1 e -1, respectivamente, que son subespazos ortogonais para os que, por Gram–Schmidt, podemos coller bases ortonormais. Sexa F un complementario ortogonal de $E_1 \oplus E_{-1}$, que tamén é invariante por f. Temos entón que

$$E = E_1 \oplus E_{-1} \oplus F,$$

e a base buscada é a unión de bases ortonormais destes subespazos. Como F non ten valores propios reais, o polinomio característico da restrición de f a F é o produto de polinomios irredutibles de grao 2, xa que se z é unha raíz, o seu conxugado $\bar{z}$ tamén o será (coa mesma multiplicidade).
Sexa A a matriz da restrición de f a F nunha base ortonormal $\{e_1, \ldots, e_{2r}\}$. Podemos ver agora F como un espazo vectorial sobre os complexos, ao que chamaremos $\hat{F}$ para evitar o abuso de notación, e designar por $\hat{f}$ e $(\hat{e}_1, \ldots, \hat{e}_{2r})$ a aplicación lineal e a base asociada, respectivamente. Neste caso, $\hat{f}$ é unitaria e o polinomio característico coincide co de f, polo que as raíces son complexas e conxugadas dúas a dúas. Tense que $\hat{f}$ é normal, polo que diagonaliza nunha base ortonormal $\{v_1, \ldots, v_{2r}\}$; ademais, os valores propios teñen módulo 1. Porén, os vectores propios estarán definidos sobre os complexos, non sobre os reais.
Se v é un vector propio de valor propio $z = a + bi$ e as súas coordenadas son $(x_1 + iy_1, \ldots, x_n + iy_n)$, temos que $\bar{v}$ é un vector propio de valor propio $\bar{z}$:

$$A\bar{v} = \overline{Av} = \overline{zv} = \bar{z} \cdot \bar{v}.$$

Convén observar que se pode escoller a base ortonormal de maneira que se v forma parte, $\bar{v}$ tamén; isto é así xa que vectores propios de valores propios diferentes sempre son ortogonais e se $\{v_1, \ldots, v_s\}$ é unha base ortonormal para un valor propio z, entón $\{\bar{v}_1, \ldots, \bar{v}_s\}$ éo para $\bar{z}$.
Para cada parella de vectores propios $(v, \bar{v})$, collemos

$$w = \frac{1}{\sqrt{2}}(v + \bar{v}) \quad \text{e} \quad w' = \frac{-i}{\sqrt{2}}(v - \bar{v}).$$

Equivalentemente,

$$v = \frac{w + iw'}{\sqrt{2}} \quad \text{e} \quad \bar{v} = \frac{w - iw'}{\sqrt{2}}.$$

É inmediato ver que w e w' teñen coordenadas reais, norma 1 (xa que $\langle v \pm \bar{v}, v \pm \bar{v} \rangle = 2$) e que son ortogonais entre si porque o seu produto escalar é 0.
Por outra banda,

$$\hat{f}(w) = \frac{\hat{f}(v) + \hat{f}(\bar{v})}{\sqrt{2}} = \frac{zv + \bar{z}\bar{v}}{\sqrt{2}} = w\Big(\frac{z + \bar{z}}{2}\Big) + w'i\Big(\frac{z - \bar{z}}{2}\Big).$$

Por simplicidade, pomos $a = \frac{z+\bar{z}}{2}$ e $b = -i\frac{z-\bar{z}}{2}$, co cal $\hat{f}(w) = aw - bw'$. Repetindo a operación anterior, obtemos que

$$\hat{f}(w') = w\Big(-i\frac{z - \bar{z}}{2}\Big) + w'\Big(\frac{z + \bar{z}}{2}\Big) = bw + aw'.$$

Ademais,
$$a^2 + b^2 = \frac{z^2 + \bar{z}^2 + 2z\bar{z} - z^2 - \bar{z}^2 + 2z\bar{z}}{4} = z\bar{z} = |z|^2 = 1,$$
xa que sabemos que os valores propios dunha aplicación unitaria teñen módulo 1.
En termos matriciais, $Aw = aw - bw'$ e $Aw' = bw + aw'$, e tense a condición do enunciado de que $a^2 + b^2 = 1$. Substituíndo na base $\{v_1, \dots, v_{2r}\}$ cada parella $v, \bar{v}$ polos correspondentes w e w' obtemos unha nova base ortonormal de $\hat{F}$, $\{w_1, w_1', \dots, w_r, w_r'\}$, formada por vectores de coordenadas reais. Sexa u_i o vector de F que na base $e_1, \dots, e_{2r}$ ten as mesmas coordenadas que w_i na base $\hat{e}_1, \dots, \hat{e}_{2r}$. A base $\{u_1, \dots, u_{2r}\}$ é ortonormal e está formada por parellas u, u' con $Au = au - bu'$ e $Au' = bu + au'$, onde $a, b \in \mathbb{R}$. En consecuencia, matriz de f nesta base é da forma descrita no enunciado do teorema. □

Exemplo. En $\mathbb{R}^3$, se impomos a condición de que o determinante sexa $+1$, as matrices que cumpren a condición son da forma
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$
Se a base da matriz é $\{v_1, v_2, v_3\}$, xeometricamente correspóndese cunha rotación de eixe v_1 e ángulo α. O coseno do ángulo pode lerse directamente en termos da traza da matriz, $\cos\alpha = \frac{1}{2}(\mathrm{Tr}(f) - 1)$ (o seno depende do que se chama a orientación da base, que é un concepto de carácter xeométrico).

4.5. Descomposición en valores singulares

Imos comezar a sección introducindo a noción de norma matricial. Este concepto é de especial relevancia non só no estudo da álxebra, senón que aparece frecuentemente en matemática aplicada. Nesta sección, imos supoñer que $K = \mathbb{R}$, aínda que os resultados se adaptan sen ningún problema para $K = \mathbb{C}$.

Definición 4.17. Sexa $A \in \mathcal{M}_n(\mathbb{R})$. A *norma matricial* asociada á norma vectorial $\|\cdot\|$ está dada por
$$\|A\| := \max_{\|v\|=1} \|Av\|.$$
Observamos que o máximo sempre existe: por estarmos en $\mathbb{R}^n$, temos que o conxunto de vectores de norma 1 é un compacto xa que é limitado e pechado; como a norma é unha función continua, temos que polo teorema de Weierstrass unha función continua nun compacto sempre alcanza un máximo.
A definición non require que a norma proveña dun produto escalar. De feito, en $\mathbb{R}^n$, dúas das normas máis habituais, e que non proveñen de produtos escalares, son $\|\cdot\|_1$ e $\|\cdot\|_\infty$: se $x = (x_1, \dots, x_n)$, entón
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_\infty = \max_{1\le i\le\infty} |x_i|.$$

Proposición 4.24. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Téñense as seguintes igualdades.

(a) $\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|$.

(b) $\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$.

Demostración. Imos comezar coa primeira parte. Sexa x un vector con $\|x\|_1 = 1$. Entón,

$$\begin{aligned}\|Ax\|_1 &= \sum_{i=1}^{m} \Big| \sum_{j=1}^{n} a_{ij} x_j \Big| \leq \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| \cdot |x_j| \\ &= \sum_{j=1}^{n} |x_j| \sum_{i=1}^{m} |a_{ij}| \leq \sum_{j=1}^{n} |x_j| \Big(\max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \Big) \\ &= \Big(\max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \Big) \sum_{j=1}^{n} |x_j| = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.\end{aligned}$$

Cómpre que exista un x para o que se cumpre a igualdade; para iso, collemos un vector que sexa 1 na compoñente j para a cal $\sum_{i=1}^{m} |a_{ij}|$ toma o valor máximo, e que sexa 0 nas demais.
Para a segunda parte, collemos x con $\|x\|$ con $\|x\|_\infty = 1$. Imos ver que cada compoñente de Ax é menor ou igual que $\max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$:

$$\Big| \sum_{j=1}^{n} a_{ij} x_j \Big| \leq \sum_{j=1}^{n} |a_{ij}| \cdot |x_j| \leq \sum_{j=1}^{n} |a_{ij}| \leq \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

Para ver que se pode cumprir a igualdade, consideramos o índice i tal que $\sum_{j=1}^{n} |a_{ij}|$ é máximo. Collemos entón o vector que na posición j-ésima ten un 1 se $a_{ij} \geq 0$ e un -1 se $a_{ij} < 0$. □

De agora en adiante, sexa $(E, \langle \cdot, \cdot \rangle)$ un espazo euclidiano e sexa $\| \cdot \|$ a norma asociada. Outra das normas que consideraremos é a inducida polo produto escalar habitual en $\mathbb{R}^n$; poremos $\| \cdot \|_2$. A seguinte definición introduce a chamada norma de Frobenius.

Definición 4.18. En $\mathcal{M}_n(\mathbb{R})$, sexa $\langle \cdot, \cdot \rangle$ o produto escalar dado por $\langle A, B \rangle := \operatorname{Tr}(A^t B)$. A norma inducida chámase *norma de Frobenius* e escríbese $\|A\|_F$.

Se $A = (a_{ij})$, temos entón que

$$\|A\|_F^2 = \sum_{i,j=1}^{n} |a_{ij}|^2.$$

A descomposición en valores singulares estende a idea de diagonalización a matrices non cadradas nun espazo euclidiano. Antes de comezar, enfatizamos que este proceso ten múltiples aplicacións importantes en matemática aplicada. En particular, en moitas situacións interésanos considerar aproximacións de matrices por outras de rango baixo. En aprendizaxe automático (*machine learning*) pode reducir o tamaño de conxuntos de datos de grandes dimensións, o que fai que os datos se poidan almacenar e procesar de xeito máis doado. Ademais, tamén contribúe a reducir o *ruído* e eliminar información irrelevante, aumentando a precisión da análise de datos. No lado negativo, supón unha perda de información, de aí que sexa importante escoller un valor do rango que permita reducir o tamaño, pero ao mesmo tempo non perder demasiado no proceso. Algúns ámbitos nos que se usa a descomposición en valores singulares son a compresión de imaxes, os sistemas de recomendación ou o tratamento de linguaxes naturais.

Definición 4.19. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Unha *descomposición en valores singulares* (SVD, polas siglas en inglés) de A é unha factorización da forma $A = U\Sigma V^t$, onde

$U \in \mathcal{M}_m(\mathbb{R})$ é unha matriz ortogonal, $\Sigma \in \mathcal{M}_{m\times n}(\mathbb{R})$ é unha matriz rectangular diagonal con entradas non negativas e $V \in \mathcal{M}_n(\mathbb{R})$ é outra matriz ortogonal.
No caso complexo, a definición é a mesma, cambiando a condición de ortogonal por unitaria e a de trasposta por conxugada trasposta.

O resultado principal que imos probar nese sentido é o seguinte. En termos matriciais, afirma que toda matriz admite unha descomposición en valores singulares.

Proposición 4.25. Sexan E e F dous espazos euclidianos de dimensións n e m, respectivamente e sexa $f\colon E \to F$ unha aplicación lineal. Entón, f admite unha descomposición en valores singulares, é dicir, existen bases ortonormais de E e de F, $\mathcal{B}_v = \{v_1, \dots, v_n\}$ e $\mathcal{B}_u = \{u_1, \dots, u_m\}$, de maneira que

$$f(w) = \sum_{i=1}^{\text{mín}\{m,n\}} \sigma_i \langle w, v_i \rangle u_i,$$

onde os $\sigma_i \geq 0$ son os chamados *valores singulares* de f.

Demostración. Sexa $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ a matriz asociada a f. Sexa $C = A^t A \in \mathcal{M}_n$, que é unha matriz cadrada, simétrica e semidefinida positiva. Polo teorema espectral sabemos que $C = V\Lambda V^t$, onde $V \in \mathcal{M}_n(\mathbb{R})$ é unha matriz ortogonal e $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ é unha matriz diagonal con $\lambda_1 \geq \ldots \geq \lambda_r > 0 = \lambda_{r+1} = \ldots = \lambda_n$, onde r é o rango de A. Sexan $\{v_1, \dots, v_n\}$ os vectores columna de V. Se $i > r$, entón $v_i^t C v_i = (Av_i)^t (Av_i) = 0$, polo que en particular $Av_i = 0$.
Sexa $\sigma_i = \sqrt{\lambda_i}$ e definimos $\Sigma \in \mathcal{M}_{m\times n}$ como a matriz que no bloque $r \times r$ situado arriba á esquerda se corresponde á matriz diagonal cos valores σ_i, e que ten 0 no resto das posicións. Definimos tamén $u_i = \frac{1}{\sigma_i} A v_i \in \mathbb{R}^m$, para cada $1 \leq i \leq m$. Observamos entón que os vectores $u_1, \dots, u_r$ son vectores ortonormais:

$$\begin{aligned} u_i^t u_j &= \left(\frac{1}{\sigma_i} A v_i\right)^t \left(\frac{1}{\sigma_j} A v_j\right) = \frac{1}{\sigma_i \sigma_j} v_i^t A^t A v_j \\ &= \frac{\lambda_j}{\sigma_i \sigma_j} v_i^t v_j = \frac{\sigma_j}{\sigma_i} v_i^t v_j = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j, \end{cases} \end{aligned}$$

onde usamos que $A^t A v_j = \lambda_j v_j$. Podemos agora completar a base $\{u_1, \dots, u_r\}$ con vectores $u_{r+1}, \dots, u_m$ de maneira que teñamos unha base ortonormal de $\mathbb{R}^m$. Para concluír, temos que ver que $A = U\Sigma V^t$, ou equivalentemente, que $AV = U\Sigma$. Iso é o mesmo que demostrar que se $1 \leq i \leq n$, entón $Av_i = \sigma_i u_i$. Cando $1 \leq i \leq r$, temos que $Av_i = \sigma_i u_i$ por definición; se $i > r$, entón $Av_i = 0$ e $\sigma_i = 0$, polo que o resultado é certo. □

É unha convención habitual colocar os valores singulares en orde, de xeito que $\sigma_1 \geq \sigma_2 \geq \ldots$

Exemplo. Imos achar a descomposición en valores singulares da matriz

$$A = \begin{pmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{pmatrix}.$$

Para iso, consideramos o produto $A^t A$:

$$\begin{pmatrix} 81 & -27 \\ -27 & 9 \end{pmatrix}.$$

Os valores propios son $\lambda = 90, 0$, e os vectores propios normalizados son

$$v_1 = \Big(\frac{-3}{\sqrt{10}}, \frac{1}{\sqrt{10}}\Big), \quad v_2 = \Big(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}}\Big).$$

Os valores singulares son $\sigma_1 = 3\sqrt{10}$ e $\sigma_2 = 0$. Para a matriz U, collemos

$$u_1 = \frac{1}{\sigma_1} A v_1 = \begin{pmatrix} 1/3 \\ -2/3 \\ -2/3 \end{pmatrix}.$$

Para completar a base, collemos dous vectores ortogonais a u_1 e aplicando Gram–Schmidt, obtemos unha base ortonormal. Por exemplo, $u_2 = (2/3, -1/3, 2/3)$ e $u_3 = (2/3, 2/3, -1/3)$. Polo tanto,

$$A = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ -2/3 & 2/3 & -1/3 \end{pmatrix} \begin{pmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}^t.$$

Proposición 4.26. Sexa $A \in \mathcal{M}_n(\mathbb{R})$ unha matriz cadrada e sexan $\sigma_1 \geq \sigma_2 \geq \ldots$ os seus valores singulares. Entón, $\|A\|_2 = \sigma_1$ e $\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}$.

Demostración. Pomos $B = A^t A$, que é unha matriz simétrica e que polo tanto admite unha base ortonormal. Poñamos $v_1, \ldots, v_n$ para os vectores propios e $\lambda_1, \ldots, \lambda_n$ para os valores propios correspondentes, con $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$. Se $x = \sum_{i=1}^n a_i v_i$, entón $Bx = \sum_{i=1}^n \lambda_i a_i v_i$. Polo tanto,

$$\begin{aligned} \|Ax\|_2 &= \sqrt{\langle Ax, Ax \rangle} = \sqrt{\langle x, Bx \rangle} \\ &= \sqrt{\langle \sum_{i=1}^n a_i v_i, \sum_{i=1}^n \lambda_i a_i v_i \rangle} = \sqrt{\sum_{i=1}^n \lambda_i a_i^2} \leq \sqrt{\lambda_1} \cdot \|x\|_2 \end{aligned}$$

e a igualdade alcánzase considerando o vector $x = v_1$.
Para a segunda igualdade simplemente observamos que

$$\|A\|_F = \sqrt{\mathrm{Tr}(A^t A)} = \sqrt{\mathrm{Tr}(V \Sigma U^t U \Sigma V^t)} = \sqrt{\mathrm{Tr}(V \Sigma^2 V^t)} = \sqrt{\sum_{i=1}^n \sigma_i^2}.$$

□

Proposición 4.27 (Teorema de Eckart–Young–Mirsky)**.** Sexa $A \in \mathcal{M}_{m\times n}$ unha matriz cunha SVD da forma $A = U \Sigma V^t$. Entón, $A_k = \sum_{i=1}^k \sigma_i u_i v_i^t$ é unha matriz de rango como moito k que minimiza $\|A - A_k\|_2$.

Demostración. Comezamos observando que $\|A - A_k\|_2 = \| \sum_{i=k+1}^n \sigma_i u_i v_i^t \|_2 = \sigma_{k+1}$. Cómpre probar entón que se $B_k = XY^t$, onde X e Y teñen k columnas, entón $\sigma_{k+1} \leq \|A - B_k\|_2$. A primeira observación é consecuencia directa do feito que unha matriz M de rango r pódese escribir como o produto de NP, onde N ten por columnas r columnas de M linealmente independentes e P ten por columnas as combinacións lineais de cada unha das columnas de N usadas para dar a descomposición de M.
Como Y ten k columnas, existe unha combinación lineal w das primeiras $k+1$ columnas de V que cumpre $Y^t w$, é dicir, se $v_1, \ldots, v_n$ son as columnas de v, entón para unha

elección dos γ_i, $w = \sum_{i=1}^{k+1} \gamma_i v_i$ cumpre que $Y^t w = 0$. Sen perder xeralidade, podemos supor que $\|w\|_2 = 1$, ou o que é o mesmo, que $\gamma_1^2 + \ldots + \gamma_{k+1}^2 = 1$. Entón,

$$\|A - B_k\|_2^2 \geq \|(A - B_k)w\|_2^2 = \|Aw\|_2^2 = \sum_{i=1}^{k+1} \gamma_i^2 \sigma_i^2 \geq \sigma_{k+1}^2.$$

A conclusión séguese tomando raíces cadradas na igualdade anterior. □

A matriz A_k non ten por que ser a única que resolve o problema de minimización (aínda que si é o caso cando todos os valores singulares son distintos). O mesmo resultado é certo para a norma de Frobenius. Porén, para demostralo necesitamos un resultado previo sobre os rangos de matrices.

Lema 4.1. Sexa K un corpo e sexan $A, B \in \mathcal{M}_{m\times n}(K)$ dúas matrices $m \times n$ con entradas nese corpo. Cúmprese que

$$\operatorname{rango}(A + B) \leq \operatorname{rango}(A) + \operatorname{rango}(B).$$

Demostración. Para unha matriz $X \in \mathcal{M}_{m\times n}(K)$, pomos $C(X)$ para o subespazo de K^m xerado polas columnas de X. Temos entón que $\operatorname{rango}(X) = \dim C(X)$. O resultado quedará probado se vemos que $C(A + B) \subset C(A) + C(B)$. Se $v \in C(A + B)$, iso quere dicir que existe un vector $w \in K^n$ de maneira que $(A + B)w = v$. Nese caso, temos que $v = Aw + Bw$, con $Aw \in C(A)$ e $Bw \in C(B)$, polo que temos o resultado buscado. □

Proposición 4.28. Sexa $A \in \mathcal{M}_{m\times n}$ unha matriz con SVD $A = U\Sigma V^t$. Entón, $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^t$ é unha matriz de rango como moito k que minimiza $\|A - A_k\|_F$.

Demostración. Temos que $\|A - A_k\|_F^2 = \sum_{i=k+1}^{n} \sigma_i^2$. Como no resultado anterior, cómpre ver que se $B_k = XY^t$, con X e Y matrices con k columnas, entón $\|A - A_k\|_F^2 \leq \|A - B_k\|_F^2$.
Imos considerar unha descomposición calquera da forma $A = A' + A''$, que logo aplicaremos a unhas matrices concretas. Pola desigualdade triangular para a norma $\|\cdot\|_2$, $\sigma_1(A) \leq \sigma_1(A') + \sigma_1(A'')$; para demostrar isto, sexa v un vector de norma 1de xeito que $\|A\|_2 = \|Av\|_2$. Entón,

$$\|A\|_2 = \|Av\|_2 \leq \|A'v\|_2 + \|A''v\|_2 \leq \|A'\| + \|A''\|.$$

Supoñamos que A'_k e A''_k son as aproximacións de rango k de A' e A'' dadas pola SVD. Entón, para calquera par $i, j \geq 1$,

$$\begin{aligned}
\sigma_i(A') + \sigma_j(A'') &= \sigma_1(A' - A'_{i-1}) + \sigma_1(A'' - A''_{j-1}) \\
&\geq \sigma_1(A - A'_{i-1} - A''_{j-1}) \\
&\geq \sigma_1(A - A_{i+j-2}) \\
&= \sigma_{i+j-1}(A),
\end{aligned}$$

onde usamos que o rango de $A'_{i-1} + A''_{j-1}$ é menor ou igual que o de A_{i+j-2}, xa que en xeral o rango dunha suma de matrices é sempre menor ou igual que a suma dos rangos. Como $\sigma_{k+1}(B_k) = 0$, se $A' = A - B_k$ e $A'' = B_k$ concluímos que para $i \geq 1$ e $j = k + 1$, $\sigma_i(A - B_k) \geq \sigma_{k+i}(A)$. Polo tanto,

$$\|A - B_k\|_F^2 = \sum_{i=1}^{n} \sigma_i(A - B_k)^2 \geq \sum_{i=k+1}^{n} \sigma_i(A)^2 = \|A - A_k\|_F^2.$$

□

Exemplo. Consideremos a descomposición en valores singulares

$$A = \begin{pmatrix} 4 & 0 \\ 3 & -5 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2\sqrt{10} & 0 \\ 0 & \sqrt{10} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

A matriz de rango 1 que mellor aproxima a matriz A é

$$A_1 = \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2\sqrt{10} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 4 & -4 \end{pmatrix}.$$

Observamos que

$$A - A_1 = \begin{pmatrix} 2 & 2 \\ -1 & -1 \end{pmatrix};$$

o cadrado da norma de Frobenius é $\|A - A_1\|_F^2 = 10$, isto é, a norma é $\sqrt{10}$, que coincide co segundo valor singular. O mesmo sucede para $\|A - A_1\|_2^2$.

4.6. Formas cuadráticas

Nesta sección, imos considerar que K é un corpo arbitrario, aínda que cara ao final, volveremos centrarnos nos casos nos que $K = \mathbb{R}$ ou $K = \mathbb{C}$.

Definición 4.20. Unha *forma cuadrática* é unha aplicación $q\colon E \to K$ tal que existe unha forma bilineal simétrica $\phi\colon E \times E \to K$ de maneira que $q(x) = \phi(x, x)$. Aquí, q é a forma cuadrática asociada a ϕ e escribimos $q = q_\phi$. Por definición, a matriz de $q = q_\phi$ na base $\mathcal{B}_e$ é a matriz de ϕ na base e, $\operatorname{Mat}(\phi, \mathcal{B}_e)$.

Alternativamente, podemos escribir $q_\phi(x) = x^t \operatorname{Mat}(\phi, \mathcal{B}_e) x$, se x se expresa na base e. Observemos que en ningún momento nos estamos restrinxindo a formas bilineais definidas positivas, senón que unicamente esiximos a condición de simetría.

Exemplo. Nun K-espazo vectorial E, consideramos $q(x, y) = 2x^2 - 4xy + 5y^2$. Trátase dunha forma cuadrática que ten por forma bilineal asociada $\phi\colon E \times E \to K$

$$\phi((x_1, x_2), (y_1, y_2)) = 2x_1y_1 - 2x_2y_1 - 2x_1y_2 + 5x_2y_2,$$

e a súa matriz na base canónica é

$$\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

Podemos comprobar que

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2x^2 - 4xy + 5y^2.$$

Un primeiro resultado que podemos probar para formas bilineais e que nos será útil é o seguinte. Convén recordar que a característica dun corpo é o menor enteiro positivo n tal que $n \cdot 1 = 0$; se non existe ningún enteiro con esa propiedade, dise que a característica é 0. Por exemplo, $\mathbb{Z}/p\mathbb{Z}$ ten característica p, mentres que $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$ teñen característica 0.

Proposición 4.29. Sexa K un corpo de característica diferente de 2 (é dicir, $2 \neq 0$) e ϕ unha forma bilineal simétrica diferente de 0. Entón, existe un vector $v \in E$ de maneira que $\phi(v, v) \neq 0$.

Demostración. Como ϕ non é identicamente nula, existen dous vectores $v_1, v_2 \in E$ de maneira que $\phi(v_1, v_2) \neq 0$. Se $\phi(v_1, v_1) \neq 0$ ou $\phi(v_2, v_2) \neq 0$, xa acabamos. Senón, observamos que a súa suma é

$$\begin{aligned} q(v_1 + v_2) &= \phi(v_1 + v_2, v_1 + v_2) \\ &= \phi(v_1, v_1) + \phi(v_2, v_2) + 2\phi(v_1, v_2). \\ &= 2\phi(v_1, v_2) \neq 0 \end{aligned}$$

□

A condición de que a característica sexa diferente de 2 é necesaria. Consideremos por exemplo o caso de $K = (\mathbb{Z}/2\mathbb{Z})$ e $E = K^2$. Se $\phi\colon E \times E \to K$ está dada por

$$\phi((x_1, x_2), (y_1, y_2)) = x_1 y_2 + x_2 y_1,$$

temos que a matriz da forma bilineal na base canónica é

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

e en particular é non cero. En cambio, temos que

$$q((0,0)) = q((1,0)) = q((0,1)) = q((1,1)) = 0,$$

é dicir, a avaliación da forma cuadrática en calquera elemento de E dá 0.
O resultado anterior pódese estender e, máis en xeral, temos que nun corpo K de característica diferente de 2, podemos achar unha forma bilineal asociada a unha forma cuadrática $q\colon E \to K$.

Proposición 4.30. Nun corpo de característica diferente de 2, hai unha bixección entre as formas bilineais simétricas e as formas cuadráticas. Estes conxuntos, á súa vez, están en bixección coas matrices $A \in \mathcal{M}_n(K)$ tales que $A^t = A$.

Demostración. Dada unha forma cuadrática $q\colon E \to K$, definimos a forma bilineal asociada como

$$\psi_q(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)),$$

onde necesitamos que a característica sexa diferente de 2 para poder dividir. É unha comprobación rutineira ver que a forma cuadrática asociada á forma bilineal ψ_q é precisamente q, e que a forma bilineal asociada á forma cuadrática é a forma bilineal de partida. □

É posible dar unha definición equivalente de forma cuadrática cando a característica do corpo non é 2, pedindo que cumpra as seguintes propiedades. O resultado vén a dicir que se a forma cuadrática é 0, entón a correspondente forma bilineal tamén.

Proposición 4.31. Sexa K un corpo de característica diferente de dous. A definición de forma cuadrática é equivalente a dar unha aplicación $q\colon E \to K$ que cumpra as seguintes condicións.

(i) $q(x + y + z) = q(x + y) + q(x + z) + q(y + z) - q(x) - q(y) - q(z)$.

(ii) $q(\lambda x + y) = q(\lambda x) + q(y) + \lambda(q(x + y) - q(x) - q(y))$.

(iii) $q(\lambda x) = \lambda^2 q(x)$.

Demostración. Se $q = q_\phi$, onde ϕ é unha forma bilineal, é unha comprobación inmediata ver que cumpren as tres condicións do enunciado. Para ver o recíproco, supoñamos que q cumpre esas propiedades e definamos

$$\psi(x,y) = q(x+y) - q(x) - q(y).$$

Hai que comprobar que ψ é unha forma bilineal. En concreto, hai que ver que respecta a suma, o produto por escalares e que é simétrica. Da definición, é evidente que $\psi(x,y) = \psi(y,x)$. Para ver que $\psi(x+z,y) = \psi(x,y) + \psi(z,y)$ (e o mesmo na segunda variable) é suficiente con escribir as definicións correspondentes e usar a primeira propiedade. Para ver que $\psi(\lambda x, y) = \lambda\psi(x,y)$ (e o mesmo na segunda variable), usamos a segunda condición:

$$\begin{aligned}\psi(\lambda x, y) &= q(\lambda x + y) - q(\lambda x) - q(y)\\ &= q(\lambda x) + q(y) + \lambda q(x+y) - \lambda q(x) - \lambda q(y) - q(\lambda x) - q(y)\\ &= \lambda q(x+y) - \lambda q(x) - \lambda q(y)\\ &= \lambda\psi(x,y).\end{aligned}$$

Finalmente, para ver que q está asociada á forma bilineal $\frac{1}{2}\psi$, empregamos a terceira propiedade:

$$\frac{1}{2}\psi(x,x) = \frac{1}{2}q(2x) - q(x) = 2q(x) - q(x) = q(x).$$

□

Definición 4.21. O *rango* de $\phi\colon E \times E \to K$ é o rango da matriz asociada á forma bilineal $A = \mathrm{Mat}(\phi, \mathcal{B}_e)$.

A definición é correcta, xa que se $B = \mathrm{Mat}(\phi, \mathcal{B}_u)$ é a matriz de ϕ noutra base u de E, entón $B = S^t A S$ e o rango de B coincide co de A.

Definición 4.22. O *radical* de $\phi\colon E \times E \to K$ é o conxunto

$$\mathrm{rad}(\phi) = \{x \in E \mid \phi(x,y) = 0 \text{ para todo } y \in E\}.$$

De xeito trivial tense que o radical é un subespazo vectorial de E. Dise que ϕ é non dexenerada cando o radical de ϕ é $\{0\}$. A modo de notación, e como xa fixemos en seccións anteriores, escribimos $\phi_x\colon E \to K$ para referirnos á aplicación lineal na que fixamos que a primeira compoñente da forma bilineal sexa x, é dicir, $\phi_x(y) = \phi(x,y)$. A modo de notación, diremos que unha base $\mathcal{B}_u = \{u_1, \ldots, u_n\}$ é ϕ-ortogonal se $\phi(u_i, u_j) = 0$ para todo $i \neq j$. Nese caso, a matriz $\mathrm{Mat}(\phi, \mathcal{B}_u)$ é diagonal e falamos da forma reducida de ϕ.

Proposición 4.32. Sexa $u = u_1, \ldots, u_n$ unha base de E de maneira que $\mathrm{Mat}(\phi, \mathcal{B}_u)$ sexa a matriz diagonal $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Supoñamos que os vectores están ordenados de maneira que $\lambda_i = 0$ para $i = 1, \ldots, s$ e $\lambda_i \neq 0$ para $i = s+1, \ldots, n$. Entón, $\mathrm{rad}(\phi) = \langle u_1, \ldots, u_s\rangle$. En particular, a suma da dimensión do radical e do rango coincide coa dimensión de E.

Demostración. Comezamos observando que $\langle u_1, \dots, u_s \rangle \subset \mathrm{rad}(\phi)$, xa que se escribimos $y = \sum_{i=1}^n y_i u_i \in E$, entón temos que

$$\phi(u_j, y) = \sum_{i=1}^n y_i \phi(u_i, u_j) = y_j \phi(u_j, u_j) = 0.$$

Polo tanto, $u_j \in \mathrm{rad}(\phi)$. Do mesmo xeito, se $x \in \mathrm{rad}(\phi)$ se escribe como $x = \sum_{i=1}^n x_i u_i$ tense que

$$0 = \phi(x, u_j) = \sum_{i=1}^n x_i \phi(u_i, u_j) = x_j \phi(u_j, u_j).$$

Polo tanto, se algún u_j, con $j > s$, é tal que $x_j \neq 0$, teremos que $\phi(u_j, u_j) = 0$, o que é unha contradición. $\square$

Unha última noción xeral para o estudo das formas cuadráticas é a de vector isótropo.

Definición 4.23. Sexa $q\colon E \to K$ unha forma cuadrática. Dise que un vector v é *isótropo* se $q(v) = 0$. O conxunto de todos os vectores isótropos denomínase *cono de isotropía.*

En xeral, o cono de isotropía non será un subespazo vectorial. Por exemplo, se $q(x, y) = 2xy$ temos que os vectores isótropos son os da forma $(a, 0)$ e $(0, b)$, con $a, b \in K$.
O seguinte obxectivo que temos ao estudar as formas cuadráticas é encontrar unha base $\mathcal{B}_u = \{u_1, \dots, u_n\}$ de E de maneira que a matriz sexa diagonal. É dicir, queremos encontrar S de maneira que $S^t AS = D$. Á base $\mathcal{B}_u$ chámaselle *base adaptada* e á matriz D chámaselle *forma reducida* de ϕ. Ás veces tamén falaremos dunha *base ortogonal* de ϕ. O seguinte resultado pódese interpretar en termos do teorema espectral cando o corpo K é igual a $\mathbb{R}$, xa que nos está a dicir que unha matriz simétrica diagonaliza nunha base ortogonal (aquí non temos o concepto de norma, polo que non falamos de ortonormalidade).

Proposición 4.33. Sexa $\phi\colon E \times E \to K$ unha forma bilineal simétrica. Entón existe unha base ortogonal de ϕ.

Demostración. Faremos a demostración por indución na dimensión do espazo. Se a dimensión é 1, o resultado é evidente. Supoñamos que a dimensión de E é $n > 1$ e que o resultado é certo para espazos de dimensión $n - 1$. Se $\phi = 0$, entón calquera base é ortogonal. Se $\phi \neq 0$, en particular $q_\phi \neq 0$. Polo tanto, existe $u_1 \in E$, con $u_1 \neq 0$ de maneira que $q(u_1) = \phi(u_1, u_1) \neq 0$. Polo tanto, $\phi_{u_1} : E \to K$ é unha aplicación non nula e o seu núcleo, F, ten dimensión $n - 1$. Como $\langle u_1 \rangle \cap F = 0$, temos que $E = \langle u_1 \rangle \oplus F$. Sexa ψ a restrición de ϕ a F. Temos que ψ tamén é unha forma bilineal simétrica, polo que grazas á hipótese de indución existe unha base $u_2, \dots, u_n$ para a cal a forma bilineal é ortogonal. Polo tanto, $\{u_1, \dots, u_n\}$ é unha base ortogonal para ϕ. $\square$

Para levar isto á práctica imos introducir o chamado método de Gauss (tamén chamado, nalgúns textos, congruencia-pivote). Recordamos que facer unha transformación elemental por filas nunha matriz A é o mesmo que multiplicar por unha matriz elemental F á esquerda de A. Analogamente, facer unha transformación elemental por columnas en A é o mesmo que multiplicar por unha matriz elemental C á dereita de A. O método de Gauss consiste en facer transformacións elementais por filas e despois repetir as mesmas operacións por columnas ata conseguir diagonalizar a matriz (alternativamente, primeiro por columnas e logo por filas). Podemos ir alternando as transformacións e cada vez que facemos unha, inmediatamente despois repetímola, de xeito que se primeiro era por filas logo é por columnas, e viceversa. Deste xeito, podemos facer

$$\begin{aligned}(A|\mathbb{I}) &\sim (F_1A|F_1) \sim (F_1AC_1|F_1) \sim (F_2F_1AC_1|F_2F_1) \sim \\ &\sim (F_2F_1AC_1C_2|F_2F_1) \sim \ldots \sim (F_r \ldots F_1AC_1 \ldots C_r|F_r \ldots F_1).\end{aligned}$$

Se a matriz final é diagonal, escribimos $D := F_r \ldots F_1AC_1 \ldots C_r$ e $S^t := F_r \ldots F_1$. Polo tanto,

$$S = (F_r \ldots F_1)^t = F_1^t \ldots F_r^t = C_1 \ldots C_r$$

e $D = S^tAS$. Collendo a base $u = u_1, \ldots, u_n$ definida polas columnas de S, temos que a matriz da forma bilineal nesa base é diagonal.

Exemplo. Sexa $q\colon \mathbb{R}^3 \to \mathbb{R}$ a forma cuadrática

$$q(x, y, z) = x^2 + 4xy + 2xz + 3y^2 + 10yz + 4z^2.$$

A matriz de q na base canónica é

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 5 \\ 1 & 5 & 4 \end{pmatrix}.$$

Realizamos as transformacións $f_2' = f_2 - 2f_1$ e $f_3' = f_3 - f_1$, e logo por columnas $c_2' = c_2 - 2c_1$ e $c_3' = c_3 - c_1$; logo facemos $f_3' = f_3 + 3f_2$ e $c_3' = c_3 + 3c_2$. Deste xeito chegamos a

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 12 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Temos que o resultado de avaliar $q(x, y, z)$ nos vectores columna de S dános os valores da diagonal:

$$q(1, 0, 0) = 1,\ q(-2, 1, 0) = -1,\ q(-7, 3, 1) = 12.$$

Observamos ademais que se ten que

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -7 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 5 \\ 1 & 5 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & -7 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 12 \end{pmatrix}.$$

Porén, poderiamos ter multiplicado por $\frac{1}{\sqrt{12}}$ a terceira fila e logo a terceira columna. Nese caso,

$$D' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S' = \begin{pmatrix} 1 & -2 & -\frac{7}{\sqrt{12}} \\ 0 & 1 & \frac{3}{\sqrt{12}} \\ 0 & 0 & \frac{1}{\sqrt{12}} \end{pmatrix}.$$

É dicir, a forma diagonal reducida non é única.

No caso de característica 2 o método de congruencia-pivote pode non funcionar. Podemos considerar, a modo de exemplo, o da forma bilineal $\phi((x_1, x_2), (y_1, y_2)) = x_1y_2 + x_2y_1$.

Para ilustrar o comentario final do exemplo, imos considerar o caso máis elemental posible, o das matrices 1×1. Neste caso, un cambio de base consiste en multiplicar por un escalar non nulo tanto a fila como a columna, é dicir, en pasar de ter un elemento

(α) a ter $(d^2\alpha)$, onde $d \in K^\times$. Polo tanto, dúas matrices (α) e (β) representan a mesma forma cuadrática se, e soamente se, $\alpha = d^2\beta$, con $d \neq 0$; trátase polo tanto dunha relación de equivalencia. Como no caso das formas bilineais, diremos que as formas cuadráticas que están na mesma clase de equivalencia son congruentes. Imos explorar algúns casos.

- Se $K = \mathbb{C}$ e $\alpha, \beta \neq 0$, entón sempre se ten que α/β ten unha raíz cadrada. Polo tanto, no caso complexo, só hai dúas clases de congruencia, a de (0) e a de (1).
- Se $K = \mathbb{R}$ e $\alpha, \beta \neq 0$, tense que α/β ten raíz cadrada se, e soamente se, α e β teñen o mesmo signo. Polo tanto, hai tres clases de congruencia, a de (0), a de (1) e a de (-1).
- Se $p > 2$ e $K = \mathbb{Z}/p\mathbb{Z}$, podemos ver ao igual que sucede nos números reais, hai dúas clases de congruencia entre os elementos non cero: os cadrados e os non cadrados. Por exemplo, se $K = \mathbb{Z}/7\mathbb{Z}$, os cadrados son $\{1, 2, 4\}$ e os non cadrados $\{3, 5, 6\}$. Dado dous non cadrados, tense que o seu cociente sempre é un cadrado, por exemplo, $5/3 = 4$. Temos tamén 3 clases de congruencia, a do cero, a dos cadrados non cero e a dos non cadrados.
- Se $K = \mathbb{Q}$ a situación é moito máis rica e hai infinitas clases de congruencia.

O obxectivo principal desta sección é estudar cando dúas formas cuadráticas representan a mesma en diferentes bases, é dicir, en termos matriciais, cando se pode escoller unha matriz P de maneira que $B = PAP^t$. A noción é diferente que no caso das aplicación lineais, cando se puña P^{-1} en vez de P^t.
Con estes conceptos, podemos formular o problema de clasificación das formas cuadráticas. En primeiro lugar, observamos que hai dúas maneiras naturais de proceder: son as chamadas clasificacións afín e proxectiva.

Proposición 4.34. Sexan E e F dous K-espazos vectoriais da mesma dimensión n e sexa $\phi\colon E \times E \to K$ unha forma bilineal simétrica. Sexa $\alpha\colon E \to F$ un isomorfismo e $\xi\colon F \times F \to K$ definida por $\xi = \phi \circ (\alpha^{-1} \times \alpha^{-1})$. Entón, temos o seguinte:

(a) A aplicación ξ é unha forma bilineal simétrica.

(b) Se $\mathcal{B}_v$ é unha base de E, entón $\mathcal{B}_u = \alpha(\mathcal{B}_v)$ é unha base de F e $(\xi)_{\mathcal{B}_u} = (\phi)_{\mathcal{B}_v}$.

(c) Se $q = q_\phi$ é a forma cuadrática asociada a ϕ, entón a forma cuadrática asociada a ξ é $q_\xi = q_\phi \circ \alpha^{-1}$.

Demostración. (a) Temos que

$$\begin{aligned}\xi(x_1 + x_2, y) &= \phi(\alpha^{-1}(x_1 + x_2), \alpha^{-1}(y))\\ &= \phi(\alpha^{-1}(x_1) + \alpha^{-1}(x_2), \alpha^{-1}(y))\\ &= \phi(\alpha^{-1}(x_1), \alpha^{-1}(y)) + \phi(\alpha^{-1}(x_2), \alpha^{-1}(y))\\ &= \xi(x_1, y) + \xi(x_2, y).\end{aligned}$$

A comprobación da suma na segunda variable, do produto por escalares e da simetría é igualmente inmediata.

(b) Dados dous índices i, j, consideramos os vectores v_i, v_j e $u_i = \alpha(v_i)$, $u_j = \alpha(v_j)$. Entón,
$$\xi(u_i, u_j) = \phi(\alpha^{-1}(\alpha(v_i)), \alpha^{-1}(\alpha(v_j))) = \phi(v_i, v_j),$$
o que demostra que a matriz de ξ na base $\mathcal{B}_u$ coincide coa matriz de ϕ na base $\mathcal{B}_v$.

(c) Temos que $q_\xi = \phi(\alpha^{-1}(x), \alpha^{-1}(x)) = (q_\phi \circ \alpha^{-1})(x)$, polo que a afirmación é certa. $\square$

Coas notacións anteriores, diremos que ϕ e equivalente a ψ (ou congruente con ψ) cando existe un isomorfismo $\alpha \colon E \to F$ tal que $\psi = \phi \circ (\alpha^{-1} \times \alpha^{-1})$. Temos por tanto unha relación de equivalencia no espazo de formas cuadráticas.
De cara ao estudo da xeometría, interésanos dar tamén outra noción de equivalencia.

Definición 4.24. Diremos que ϕ e ψ son *proxectivamente equivalentes* se existe un isomorfismo $\alpha \colon E \to F$ e un escalar $\lambda \neq 0$ de maneira que $\psi = \lambda\phi \circ (\alpha^{-1} \times \alpha^{-1})$.

Como sucedía no caso anterior, é unha comprobación sinxela ver que esta relación é de equivalencia. Se dúas formas son equivalentes no contexto afín, tamén o serán no proxectivo, tomando $\lambda = 1$. En cambio, pódeno ser no sentido proxectivo, pero non no afín. Por exemplo,
$$q_1(x, y) = x^2 + y^2, \quad q_2(x, y) = -x^2 - y^2$$
son proxectivamente equivalentes, pero non equivalentes no sentido afín.
Imos pasar a estudar agora o caso real e o caso complexo, que son aqueles nos que podemos dar respostas máis simples. Durante o estudo dos produtos escalares, fixemos énfase na noción de que unha forma bilineal fose definida positiva, e demos un criterio (de Sylvester) para saber cando iso sucede. De cara ao estudo das formas cuadráticas, convén lembrar a seguinte notación.

Definición 4.25. Dicimos que unha forma bilineal ϕ é *definida positiva* (resp. *definida negativa*) se $\phi(x, x) > 0$ (resp. $\phi(x, x) < 0$) para todo $x \in E$ con $x \neq 0$. Do mesmo xeito, dicimos que é *semidefinida positiva* (resp. *semidefinida negativa*) se $\phi(x, x) \geq 0$ (resp. $\phi(x, x) \leq 0$) para todo $x \in E$. Finalmente, dicimos que é *indefinida* se ϕ non é semidefinida.

Unha matriz simétrica é definida positiva (resp. definida negativa) se todos os valores propios son maiores que 0; é semidefinida positiva (resp. semidefinida negativa) se todos os valores propios son maiores ou iguais que 0. É indefinida cando hai tanto valores propios positivos como negativos. De cara ao cálculo de extremos no cálculo diferencial, é importante ter en conta que as matrices definidas positivas, definidas negativas ou indefinidas permiten obter directamente información sobre o carácter dos puntos críticos. En cambio, as semidefinidas positivas ou semidefinidas negativas requiren dun estudo máis sutil. Convén ter en conta que o criterio de Sylvester pódese adaptar facilmente para matrices definidas negativas ou indefinidas.

Proposición 4.35 (Criterio de Sylvester para matrices definidas negativas)**.** Sexa $A = \mathrm{Mat}(\phi, \mathcal{B}_u) \in \mathcal{M}_n(\mathbb{R})$ a matriz dunha forma bilineal simétrica nun $\mathbb{R}$-espazo vectorial de dimensión finita n. A forma ϕ é definida negativa se, e soamente se, os determinantes $(-1)^k \det((a_{ij})_{1 \leq i,j \leq k}) > 0$ para todo $k = 1, 2, \ldots, n$.

Demostración. Temos que A é definida negativa se, e soamente se, $-A$ é definida positiva. Polo criterio de Sylvester, tense que $-A$ é definida positiva se, e soamente se,

$$\det((-a_{ij})_{1\leq i,j\leq k}) = (-1)^k \det((a_{ij})_{1\leq i,j\leq k}) > 0$$

para todo $k = 1, 2, \ldots, n$. □

Definición 4.26. Sexa $\phi\colon E \times E \to \mathbb{R}$ unha forma bilineal simétrica. Consideramos $D = \mathrm{Mat}(\phi, \mathcal{B}_u) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ unha forma reducida de ϕ.

(i) O *índice de inercia positivo* (resp. *negativo*) defínese como o número de $\lambda_i > 0$ (resp. $\lambda_i < 0$) e denótase como $i_+(\phi)$ (resp. $i_-(\phi)$).

(ii) O *índice de inercia de Sylvester* de ϕ é o mínimo de i_+ e i_- e o rango de ϕ correspóndese co número de λ_i que non son 0. Escribimos $i(\phi)$.

(iii) Escribimos $i_0(\phi)$ para o número de ceros que ten unha forma reducida.

Observamos que se cumpre sempre que $i_+(\phi)+i_-(\phi)+i_0(\phi) = n$, polo que se para dúas formas cuadráticas en espazos da mesma dimensión dúas das cantidades coinciden, a terceira tamén o fará.

Proposición 4.36. Sexa $\phi\colon E \times E \to \mathbb{R}$ unha forma bilineal simétrica. Entón, $i_+(\phi)$ é a máxima dimensión dun subespazo F de maneira que $\phi|F$ é definida positiva. O resultado análogo para $i_-(\phi)$ tamén é certo.

Demostración. Sexa u unha base ortogonal para a forma bilineal. Como a restrición ao subespazo xerado polos vectores correspondentes aos elementos positivos da diagonal é definida positiva, temos que é suficiente probar que non pode haber subespazos de dimensión maior con esa propiedade.
Sexa F de maneira que $\phi|F$ é definida positiva. Supoñamos que a dimensión de F é maior que $i_+(\phi)$ e chegaremos a unha contradición. Denotamos por $F_-(u)$ e $F_0(u)$ o subespazo xerado polos vectores da base u correspondentes aos elementos negativos e nulos da diagonal, respectivamente. Afirmamos agora que $F \cap (F_-(u) \oplus F_0(u)) = 0$, xa que se $x \in F$ e $x = y + z$, con $y \in F_-(u)$ e $z \in F_0(u)$ teriamos que

$$0 \leq \phi(x, x) = \phi(y, y) + 2\phi(y, z) + \phi(z, z).$$

Como $y \in F_-(u)$, entón $\phi(y, y) \leq 0$. Como z está no radical, entón $\phi(y, z) = \phi(z, z) = 0$, Polo tanto, $0 \leq \phi(x, x) = \phi(y, y) \leq 0$, de onde $\phi(x, x) = 0$. Como a restrición de ϕ a F é definida positiva, entón $x = 0$. Polo tanto, empregando a fórmula de Grassmann temos que

$$\dim(F + (F_-(u) + F_0(u))) = \dim(F) + \dim(F_-(u) + F_0(u)) \geq n + 1,$$

que é unha contradición.
Esta caracterización non depende da base escollida, xa que a imaxe dun subespazo F por un cambio de base mantén a propiedade de ser definida positiva. □

A seguinte proposición é o que se coñece habitualmente como *a lei de inercia de Sylvester.*

Proposición 4.37. (a) Se $K = \mathbb{R}$, entón dúas forma cuadráticas ϕ e ψ son equivalentes se, e soamente se, coinciden os rangos e $i_+(\phi) = i_+(\psi)$.

(b) Se $K = \mathbb{C}$, entón dúas formas cuadráticas son equivalentes se, e soamente se, coinciden os seus rangos.

Demostración. No caso real, cun cambio de base, podemos conseguir que todos os elementos positivos (resp. negativos) nunha forma reducida pasen a ser 1 (resp. -1). Pola proposición anterior, temos que tanto o rango como a cantidade $i_+(\phi)$ (resp. $i_-(\phi)$ son invariantes dunha forma cuadrática que se manteñen por cambios de base, polo que a cantidade de elementos positivos e negativos sempre se ten que manter constante. Polo tanto, a condición do enunciado é suficiente, pero tamén necesaria.
No caso complexo, cun cambio de base podemos conseguir que todos os elementos non nulos nunha forma reducida pasen a ser 1. Como o rango é constante por cambios de base, que estes coincidan é condición necesaria e suficiente para que dúas formas cuadráticas sexan equivalentes. □

Proposición 4.38. (a) Se $K = \mathbb{R}$, entón dúas forma cuadráticas ϕ e ψ son proxectivamente equivalentes se coinciden os rangos e $i(\phi) = i(\psi)$.

(b) Se $K = \mathbb{C}$, entón dúas formas cuadráticas son proxectivamente equivalentes se coinciden os seus rangos.

Demostración. A demostración é igual que no caso afín, tendo en conta que agora podemos cambiar de signo e conseguir que onde unha forma cuadrática era definida positiva agora sexa definida negativa (e viceversa). □

Pódese comprobar que nun espazo vectorial real de dimensión 3 hai 10 opcións para o par (r, i_+), polo que hai 10 clases de equivalencia afín. En cambio, só hai 6 clases de equivalencia proxectiva. As ecuacións reducidas das 10 formas cuadráticas son as seguintes:

- $q(x, y, z) = 0$.
- $q(x, y, z) = x^2$.
- $q(x, y, z) = -x^2$.
- $q(x, y, z) = x^2 + y^2$.
- $q(x, y, z) = x^2 - y^2$.
- $q(x, y, z) = -x^2 - y^2$.
- $q(x, y, z) = x^2 + y^2 + z^2$.
- $q(x, y, z) = x^2 + y^2 - z^2$.
- $q(x, y, z) = x^2 - y^2 - z^2$.
- $q(x, y, z) = -x^2 - y^2 - z^2$.

No caso proxectivo, as 6 formas cuadráticas son:

- $q(x, y, z) = 0$.
- $q(x, y, z) = x^2$.
- $q(x, y, z) = x^2 + y^2$.

- $q(x, y, z) = x^2 - y^2$.
- $q(x, y, z) = x^2 + y^2 + z^2$.
- $q(x, y, z) = x^2 + y^2 - z^2$.

No caso complexo, en cambio, o único que importa é o rango (tanto desde o punto de vista afín como desde o proxectivo), polo que hai catro clases de equivalencia correspondentes a cada un dos catro posibles rangos.
Algunhas das ideas que introducimos durante a clasificación de formas cuadráticas nos reais ou nos complexos pódense estender a corpos arbitrarios. Neses casos, un dos resultados máis importante é o chamado *teorema de cancelación de Witt*. Para dar unha idea aproximada do seu contido, diremos que un *espazo cuadrático* é un par (E, q), onde E é un K-espazo vectorial e $q\colon E \to K$ é unha forma cuadrática en E. Unha descomposición $(E, q) = (E_1, q_1) \oplus (E_2, q_2)$ indica que $E = E_1 \oplus E_2$ e que se ϕ_q é a forma bilineal asociada a q, entón $\phi_q(x, y) = 0$ para todo $x \in E_1$ e $y \in E_2$. O teorema de cancelación de Witt afirma que se (E, q), (E_1, q_1) e (E_2, q_2) son tres espazos cuadráticos sobre un corpo K e

$$(E_1, q_1) \oplus (E, q) \simeq (E_2, q_2) \oplus (E, q),$$

entón $(E_1, q_1) \simeq (E_2, q_2)$. Este resultado é a base para establecer resultados de clasificación sobre corpors arbitrarios, que nunha primeira aproximación veñen a dicir que a condición necesaria e suficiente para que dúas formas cuadráticas sexan equivalentes é que os elementos non nulos da diagonal coincidan módulo cadrados.

4.7. Formas simplécticas

A sección anterior centrouse no estudo das aplicacións bilineais simétricas. É natural preguntarse se existe unha teoría análoga para as aplicacións bilineais antisimétricas, que xa apareceron no estudo dos determinantes. A seguinte definición ten gran importancia no estudo da xeometría.

Definición 4.27. Sexa E un K-espazo vectorial. Unha *forma simpléctica* en E é unha aplicación bilineal $\phi\colon E \times E \to K$ que ten as seguintes dúas propiedades:

(i) $\phi(v, v) = 0$ para todo $v \in E$.

(ii) Se $\phi(u, v) = 0$ para todo $v \in E$, entón $u = 0$.

Un espazo vectorial simpléctico é un par (E, ϕ)

Se a característica é diferente de dous, entón a primeira propiedade é equivalente a esixir que $\phi(u, v) = -\phi(v, u)$ para todo $u, v \in E$. A seguinte proposición resume as propiedades principais das formas simplécticas.

Proposición 4.39. Sexa E un K-espazo vectorial e $\phi\colon E \times E \to K$ unha forma simpléctica.

(a) Para calquera elección dunha base de E, a matriz de ϕ é antisimétrica e non singular.

(b) A dimensión de E é par.

(c) Se a dimensión de E é $2n$, existe unha base de E na cal a matriz de ϕ ten a forma

$$\left(\begin{array}{c|c} 0 & \mathbb{I}_n \\ \hline -\mathbb{I}_n & 0 \end{array}\right).$$

Unha base así chámase *simpléctica.*

(d) Se $F \subset E$ é un subespazo para o cal $\phi(u, v) = 0$ se $u, v \in F$, entón F ten dimensión como moito n. Se F ten dimensión igual a n, chámase *subespazo lagranxiano.*

Demostración. A antisimetría é clara. Se a matriz fose singular, existirían $\lambda_1, \ldots, \lambda_n$ de maneira que a forma $\sum_{i=1}^n \lambda_i \phi(\cdot, u_i)$ é identicamente cero. Polo tanto, iso querería dicir, definindo $u = \sum_{i=1}^n \lambda_i u_i$, que $\phi(\cdot, u)$ é a aplicación lineal nula, e pola hipótese de non dexeneración, iso quere dicir que $u = 0$.
Para demostrar os apartados (b) e (c) imos proceder por indución na dimensión de E. Cando esta é 0 ou 1 o resultado é evidente. Supoñamos agora que $\dim E = n$ e que o resultado se cumpre para calquera espazo vectorial de dimensión $n - 2$. Sexa $v \in E$. Como ϕ é non dexenerada, existe $w \in E$ de maneira que $\phi(v, w) \neq 0$. Cambiando w por un múltiplo seu, podemos supor que $\phi(v, w) = 1$, e temos entón que $\phi(w, v) = -1$. Consideramos o subespazo

$$F = \{u \in E \mid \phi(u, v) = 0,\ \phi(u, w) = 0\}.$$

Observamos que $F \cap \langle v, w\rangle = \{0\}$; se non fose así, existiría un elemento $u \in F \cap \langle v, w\rangle$. Polo tanto, $u = \alpha v + \beta w$. Como $\phi(u, v) = \beta\phi(w, v) = 0$ temos que $\beta = 0$, e de xeito similar $\alpha = 0$. Polo tanto, $u = 0$. De xeito similar, se consideramos a aplicación

$$E \to K^2, \quad u \mapsto (\phi(u, v), \phi(u, w))$$

esta é sobrexectiva porque ϕ é non dexenerada. Polo tanto, o seu núcleo, que se corresponde con F, ten dimensión $n - 2$. Isto proba que $F \oplus \langle v, w\rangle = E$.
De cara a aplicar a hipótese de indución temos que comprobar que a restrición de Ω a F é unha forma simpléctica. A condición de $\phi(u, u) = 0$ é inmediata, polo que é suficiente ver que é non dexenerada. Para iso, collemos un elemento $u \in F$ arbitrario, e sabemos que existe $u' = r + s$, con $r \in F$ e $s \in \langle v, w\rangle$ de maneira que $\phi(u, u') \neq 0$. Por definición, $\phi(u, u') = \phi(u, r) + \phi(u, s) = \phi(u, r) \neq 0$, polo que collendo $r \in F$ temos que se cumpre a condición. Agora, pola hipótese de indución temos que $\dim F$ é par e admite unha base simpléctica $\{v_1, \ldots, v_{m-1}, w_1, \ldots, w_{m-1}\}$, onde $2m = n - 2$. Polas definicións, temos que $\{v, v_1, \ldots, v_{m-1}, w, w_1, \ldots, w_{m-1}\}$ é unha base simpléctica.
Para probar a última parte, sexa F un subespazo para o cal $\phi(u, v) = 0$, e sexa $\{v_1, \ldots, v_m\}$ unha base de F. Consideramos a aplicación

$$\psi \colon E \to K^m, \quad v \mapsto (\phi(v, v_1), \ldots, \phi(v, v_m)).$$

Polo feito de que ϕ sexa non dexenerada, a dimensión da imaxe é m. Polo tanto, a dimensión do núcleo é como moito $2n - m$. Como $F \subset \ker\psi$ por hipótese, tense que $m \leq 2n - m$, e polo tanto $m \leq n$, como queriamos demostrar. □

Imos describir agora un procedemento para obter unha base simpléctica; o algoritmo pode verse como unha variación do que explicamos para Gram–Schmidt. Sexa $\phi \colon E \times E \to K$ unha forma bilineal simpléctica; supoñamos que E ten dimensión $n = 2k$.

1. Fixamos unha base $\{u_1, \ldots, u_n\}$ de E.

2. Collemos un vector calquera v_1 e buscamos v_2 de maneira que $\phi(v_1, v_2) = 1$ (como existe un para o que $\phi(v_1, v_2) \neq 0$, sempre se pode tomar de maneira que dea 1).

3. Para $i = 3, 4, \dots, n$, cambiamos u_i por
$$u_i' = u_i - \phi(u_i, v_2)v_1 + \phi(u_i, v_1)v_2.$$
É unha comprobación inmediata ver que $\phi(v_1, u_i') = \phi(v_2, u_i') = 0$, polo que tanto v_1 como v_2 son ortogonais ao resto de vectores da base; por exemplo, no primeiro caso,
$$\begin{aligned}\phi(v_1, u_i') &= \phi(v_1, u_i) - \phi(u_i, v_2)\phi(v_1, v_1) + \phi(u_i, v_1)\phi(v_1, v_2)\\ &= \phi(v_1, u_i) + \phi(u_i, v_1) = 0.\end{aligned}$$

4. Repetimos o proceso con $(u_3', u_4', \dots, u_n')$ para obter vectores v_3 e v_4, de xeito que $\phi(v_3, v_4) = 1$ e v_3 e v_4 sexan ortogonais ao resto de vectores. Iterando o algoritmo, acabamos cunha colección $(v_1, v_2, \dots, v_n)$ e temos que, pola propia construción, $\{v_1, v_3, \dots, v_k, v_2, v_4, \dots, v_{2k}\}$ forma unha base simpléctica.

Exemplo. Sexa $\{e_1, e_2, e_3, e_4\}$ a base canónica de $\mathbb{R}^4$ e sexa ϕ a forma bilineal representada pola matriz
$$A = \begin{pmatrix} 0 & -1 & 2 & 0\\ 1 & 0 & 0 & 1\\ -2 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\end{pmatrix}.$$
Tense que ϕ é simpléctica porque a matriz é antisimétrica e non singular. Como $\phi(e_1, e_2) = -1$, collemos $v_1 = e_1$ e $v_2 = -e_2$. A continuación, consideramos
$$\begin{aligned}e_3' &= e_3 - \phi(e_3, v_2)v_1 + \phi(e_3, v_1)v_2 = (0, 2, 1, 0),\\ e_4' &= e_4 - \phi(e_4, v_2)v_1 + \phi(e_4, v_1)v_2 = (-1, 0, 0, 1).\end{aligned}$$
Como $\phi(e_3', e_4') = 2$, collemos $v_3 = e_3'$ e $v_4 = e_4'/2$. Polo tanto, $\{v_1, v_3, v_2, v_4\}$ é unha base simpléctica na que a matriz é da forma
$$\begin{pmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\end{pmatrix}.$$
O caso de catro elementos é simple porque chega con aplicar o proceso de ortogonalización unha vez. De termos seis vectores, teriamos obtido v_1 e v_2 como no exemplo, e logo quedaríannos catro vectores $\{u_3', u_4', u_5', u_6'\}$ sobre os que habería proceder do mesmo xeito, collendo dous deles, v_3 e v_4, que cumprisen $\phi(v_3, v_4) = 1$ e cambiando os vectores u_5' e u_6' por vectores u_5'' e u_6'' ortogonais aos outros dous.

4.8. Problemas

O espazo euclidiano.

Problema 4.1. Sexan U, V dous $\mathbb{R}$-espazos vectoriais e $f\colon U \to V$ unha aplicación lineal. Dado un produto escalar en V $\langle\cdot,\cdot\rangle\colon V \times V \longrightarrow V$, definimos en U o produto escalar
$$\langle\cdot,\cdot\rangle_f\colon U \times U \longrightarrow \mathbb{R}, \quad \langle u, v\rangle_f = \langle f(u), f(v)\rangle,$$

onde $u, v \in U$. Demostrar que $\langle \cdot, \cdot \rangle_f$ é un produto escalar en U se, e soamente se, f é inxectiva.

Solución. A bilinealidade e a simetría de $\langle \cdot, \cdot \rangle_f$ séguense inmediatamente a partir das propiedades correspondentes de $\langle \cdot, \cdot \rangle$. Usando que $\langle \cdot, \cdot \rangle$ é definida positiva, temos que $\langle \cdot, \cdot \rangle_f$ é definida positiva se, e soamente se, f é inxectiva: se $v \neq 0$, entón $\langle v, v \rangle_f = \langle f(v), f(v) \rangle$ ten que ser maior que 0 e iso require $f(v) \neq 0$.

Problema 4.2. Dicir se as seguintes expresións definen ou non un produto escalar en $\mathbb{R}^2$:

(a) $\langle (x_1, x_2), (y_1, y_2) \rangle = 2x_1y_1 + x_1y_2 + x_2y_1 + 4x_2y_2$.

(b) $\langle (x_1, x_2), (y_1, y_2) \rangle = x_1y_1 - x_2y_2$.

Solución. Ambas son formas bilineais simétricas. A primeira é definida positiva xa que

$$\langle (x, y), (x, y) \rangle = 2x^2 + 4y^2 + 2xy = (x + y)^2 + x^2 + 3y^2 > 0$$

se $(x, y) \neq (0, 0)$. A segunda non é definida positiva xa que $\langle (0, 1), (0, 1) \rangle = -1$.

Problema 4.3. Comprobar que $F((x_1, x_2), (y_1, y_2)) = 2x_1\bar{y}_1 + (1+i)x_1\bar{y}_2 + (1-i)x_2\bar{y}_1$ é unha forma hermítica.

Solución. É inmediato ver que respecta a suma en ambas variables e o produto por escalares na primeira. O feito de que sexa sesquilineal na segunda séguese de que as correspondentes compoñentes ($\bar{y}_1$ e $\bar{y}_2$) sempre aparecen conxugadas.

Problema 4.4. Demostrar que as seguintes aplicacións son produtos escalares sobre os espazos correspondentes:

(a) $\langle A, B \rangle = \mathrm{Tr}(A \cdot B^t)$ no espazo $\mathcal{M}_n(\mathbb{R})$.

(b) $\langle p(x), q(x) \rangle = \int_{-1}^{1} w(x)p(x)q(x)\,dx$ no espazo $\mathbb{R}_2[x]$ dos polinomios de grao ≤ 2, onde w é unha función continua con valores reais positivos no intervalo $[-1, 1]$, isto é, $w(x) > 0$ para todo $x \in [-1, 1]$.

(c) $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$ no espazo $\mathcal{C}([0, 1], \mathbb{R})$ de funcións continuas no intervalo $[0, 1]$.

Escribir o enunciado da desigualdade de Cauchy–Schwarz en cada caso.

Solución. (a) A bilinealidade séguese da propiedade distributiva do produto de matrices e do feito que a transposición tamén é lineal. Do mesmo xeito, é simétrica porque a traza dunha matriz coincida coa da súa transposta. Hai que demostrar agora que é definida positiva. Sexa $A = (a_{ij})$ e $B = (b_{ij})$. Entón, a traza da matriz produto é

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}b_{ij},$$

que no caso no que $A = B$ queda $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$, que é 0 se, e soamente se, $A = 0$. A desigualdade de Cauchy–Scharz afirma que

$$|\,\mathrm{Tr}(AB^t)| \leq \sqrt{\mathrm{Tr}(AA^t)}\sqrt{\mathrm{Tr}(BB^t)}.$$

(b) A bilinealidade e a simetría son consecuencia da propiedade distributiva do produto de funcións e da integración, así como da conmutatividade do produto de funcións. Para ver que é definida positiva, chega con observar que a integral dunha función continua e non negativa nun intervalo $[a, b]$ é 0 se, e soamente se, é igual a 0: se nun punto do intervalo se ten que $f(x) = \epsilon > 0$, hai un entorno de x no que é maior que $\epsilon/2$, polo que a integral é estritamente positiva. A desigualdade de Cauchy–Schwarz neste caso asegura que

$$\Big| \int_{-1}^{1} w(x)p(x)q(x)\,dx \Big| \leq \sqrt{\int_{-1}^{1} w(x)p(x)^2\,dx} \sqrt{\int_{-1}^{1} w(x)q(x)^2\,dx}.$$

(c) Faise igual que o apartado anterior, pondo $w(x) = 1$ e cambiando o intervalo $[-1, 1]$ polo $[0, 1]$.

Problema 4.5. Sexan $a, b, c, d > 0$ números reais positivos. Demostrar que

$$(a + b + c + d)\Big(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\Big) \geq 16$$

e determinar cando se cumpre a igualdade.

Solución. En $\mathbb{R}^4$ co produto escalar habitual consideramos os vectores

$$u = (\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}), \quad v = (\sqrt{1/a}, \sqrt{1/b}, \sqrt{1/c}, \sqrt{1/d}).$$

Entón, pola desigualdade de Cauchy–Schwarz temos directamente a desigualdade do enunciado.
A igualdade cúmpre se, e soamente se, os vectores u e v son proporcionais, isto é,

$$\frac{\sqrt{a}}{\frac{1}{\sqrt{a}}} = \frac{\sqrt{b}}{\frac{1}{\sqrt{b}}} = \frac{\sqrt{c}}{\frac{1}{\sqrt{c}}} = \frac{\sqrt{d}}{\frac{1}{\sqrt{d}}}.$$

En particular, a igualdade dáse se, e soamente se, $a = b = c = d$.

Problema 4.6. Considérase o produto escalar no espazo $\mathbb{R}_2[X]$ definido por

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\,dx.$$

Determinar o ángulo que forman os polinomios $p(X) = 1 + X + X^2$ e $q(X) = 1 + X$.

Solución. Para determinar o ángulo, hai que calcular $\langle p(x), q(x) \rangle$ e tamén as normas de cada vector.

- Para o produto escalar,

$$\langle p(X), q(X) \rangle = \int_{-1}^{1} (X^3 + 2X^2 + 2X + 1)\,dx = \frac{10}{3}.$$

- A norma de $p(X)$ é a raíz cadrada de

$$\langle p(X), p(X) \rangle = \int_{-1}^{1} (X^4 + 2X^3 + 3X^2 + 2X + 1)\,dx = \frac{22}{5}.$$

- A norma de $p(X)$ é a raíz cadrada de

$$\langle q(X), q(X)\rangle = \int_{-1}^{1} (X^2 + 2X + 1)\, dx = \frac{8}{3}.$$

Polo tanto, se α é o ángulo que forman, tense que

$$\cos\alpha = \frac{10/3}{\sqrt{(22/5)}\sqrt{(8/3)}} = \frac{5\sqrt{15}}{6\sqrt{11}}.$$

Problema 4.7. Considérase a forma bilineal $\phi\colon \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$\phi((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_2y_2 - x_1y_2 + 2x_2y_1.$$

(a) Encontrar a súa matriz na base canónica.

(b) Encontrar a súa matriz na base $\{u_1, u_2\}$, con $u_1 = (1, 1)$ e $u_2 = (1, 2)$.

(c) Existe algunha base de $\mathbb{R}^2$ na cal a matriz de ϕ sexa simétrica?

Solución. A matriz na base canónica é

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Para achar a matriz na base $\{u_1, u_2\}$, imos dar dúas alternativas. A primeira é directamente coa definición:

$$(\phi(u_i, u_j)) = \begin{pmatrix} \phi(u_1, u_1) & \phi(u_1, u_2) \\ \phi(u_2, u_1) & \phi(u_2, u_2) \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 7 \end{pmatrix}.$$

Alternativamente, podemos considerar a matriz P que pasa da base $\{u_1, u_2\}$ á base canónica. Temos que $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, polo que a matriz buscada é

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 6 & 7 \end{pmatrix}.$$

Se a matriz nunha certa base, C, é unha matriz simétrica, entón noutra base vén dada por Q^tCQ, onde Q é a matriz do cambio. Entón, se C é simétrica, a matriz noutra base tamén o será, xa que $(Q^tCQ)^t = Q^tC^tQ = Q^tCQ$. Como a matriz non é simétrica na base canónica, non o pode ser en ningunha.

Problema 4.8. Encontrar os valores dos parámetros $a, b \in \mathbb{R}$ para os cales as seguintes matrices definen un produto escalar.

(a) $\begin{pmatrix} 1 & -1 & 1 \\ -1 & 3 & 0 \\ 1 & a & b \end{pmatrix}.$

(b) $\begin{pmatrix} 1 & a \\ a^2 & 2 \end{pmatrix}.$

Solución. (a) Para que sexa simétrica, a matriz ten que cumprir que $a = 0$, e para que sexa definida positiva, aplicando o criterio de Sylvester, temos que

$$1 > 0, \quad 2 > 0, \quad 3b - b - 3 > 0,$$

o que quere dicir que $b > 3/2$.

(b) Para que sexa simétrica tense que cumprir que $a = a^2$, polo que as únicas posibilidades son 0 e 1. A condición de definida positiva é $2 - a^2 > 0$, que se cumpre para ambos valores.

Ortogonalidade e proxección ortogonal.

Problema 4.9. Considérase o produto escalar no espazo de funcións continuas no intervalo $[-t, t]$ definido por $\langle f, g\rangle = \int_{-t}^{t} f(x)g(x)\,dx$. Demostrar que as funcións $\sin(ax)$ e $\cos(bx)$ son ortogonais para $a, b \in \mathbb{R}$.

Solución. É suficiente con demostrar que

$$\int_{-t}^{t} \sin(ax)\cos(bx)\,dx = 0.$$

Para iso, chega con observar que $\sin(ax)\cos(bx)$ é unha función impar, polo que a súa integral nun intervalo simétrico con respecto ao 0 é 0. Alternativamente, recordamos que

$$\sin(ax)\cos(bx) = \frac{\sin((a+b)x) + \sin((a-b)x)}{2}.$$

Polo tanto,

$$\int_{-t}^{t} \sin(ax)\cos(bx)\,dx = \frac{1}{2}\int_{-t}^{t} \sin((a+b)x)\,dx + \frac{1}{2}\int_{-t}^{t} \sin((a-b)x)\,dx = 0.$$

Problema 4.10. Aplicar o proceso de ortogonalización de Gram–Schmidt ás seguintes bases.

(a) Base $\{(1, 1, -1), (1, -1, 1), (-1, 1, 1)\}$ en $\mathbb{R}^3$ co produto estándard e co produto escalar que na base canónica ten matriz

$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 3 \end{pmatrix}.$$

(b) Base $\{(1, 1, 1, 0), (1, 1, 0, 1), (1, 0, 1, 1), (0, 1, 1, 1)\}$ en $\mathbb{R}^4$ co produto estándard e co produto escalar definido por

$$\langle x, y\rangle = x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2 + x_3y_3 + x_3y_4 + x_4y_3 + 2x_4y_4.$$

(c) Bases $\{1, X, X^2\}$ e $\{X^2, X, 1\}$ no espazo $\mathbb{R}_2[X]$ co produto escalar

$$\langle p(X), q(X)\rangle = \int_0^t p(x)q(x)\,dx,$$

onde $t > 0$ é un número real.

Solución. (a) Co produto estándard, a norma do primeiro vector é 3, polo que $w_1 = \frac{1}{\sqrt{3}}(1,1,-1)$. Para o segundo vector, calculamos

$$v_2 = (1,-1,1) + (1/3,1/3,-1/3) = (4/3,-2/3,2/3);$$

ao normalizar, vemos que $w_2 = \frac{1}{\sqrt{6}}(2,-1,1)$. Por último,

$$v_3 = (-1,1,1) + (1/3,1/3,-1/3) + (2/3,-1/3,1/3) = (0,1,1),$$

de onde vemos que $w_3 = \frac{1}{\sqrt{2}}(0,1,1)$.

Para o outro produto escalar, o proceso é similar. A norma do primeiro vector é 4, polo que $w_1 = \frac{1}{2}(1,1,-1)$. Os outros dous vectores son $w_2 = \frac{1}{\sqrt{2}}(1,0,0)$ e $w_3 = \frac{1}{2}(1,3,1)$.

(b) Co produto escalar estándard obtemos a base

$$\Big\{\frac{1}{\sqrt{3}}(1,1,1,0), \frac{1}{\sqrt{15}}(1,1,-2,3), \frac{1}{\sqrt{35}}(1,-4,3,3), \frac{1}{\sqrt{7}}(-2,1,1,1)\Big\}.$$

Co outro produto escalar temos

$$\Big\{\frac{1}{\sqrt{2}}(1,1,1,0), (0,0,-1,1), \frac{1}{\sqrt{3}}(0,-1,1,0), \frac{1}{\sqrt{6}}(-3,-1,1,0)\Big\}.$$

(c) Vamos comezar co caso de $\{1, X, X^2\}$. A norma do primeiro vector é $\sqrt{t}$, polo que $w_1 = \frac{1}{\sqrt{t}}$. Temos agora que

$$v_2 = X - \frac{1}{t}\int_0^t X = X - \frac{t}{2}.$$

O cadrado da norma é $\frac{t^3}{3} - t\frac{t^2}{2} + \frac{t^2}{4}\cdot t = \frac{t^3}{12}$. Polo tanto, obtense que

$$w_2 = \frac{\sqrt{3}}{\sqrt{t}}\Big(\frac{2}{t}X - 1\Big).$$

De xeito análogo obtense que

$$w_3 = \sqrt{\frac{5}{t}}\Big(\frac{6}{t^2}X^2 - \frac{6}{t}X + 1\Big).$$

Pasamos agora ao caso de $\{X^2, X, 1\}$. Para o primeiro vector temos que $w_1 = \sqrt{\frac{5}{t^5}}X^2$. Para o segundo,

$$v_2 = X - \frac{5X^2}{t^5}\int_0^t X^3 = X - \frac{5}{4t}X^2.$$

A norma ao cadrado de v_2 é

$$\langle v_2, v_2\rangle = \frac{t^3}{3} - \frac{5}{2t}\cdot\frac{t^4}{4} + \frac{25}{16t^2}\cdot\frac{t^5}{5} = \frac{t^3}{48}.$$

Polo tanto,

$$w_2 = \sqrt{\frac{3}{t}}\Big(\frac{-5}{t^2}X^2 + \frac{4}{t}X\Big).$$

No caso do terceiro vector,

$$v_3 = 1 - \frac{4}{t}X + \frac{10}{3t^2}X^2,$$

e por conseguinte, como a norma é $\sqrt{t}/3$,

$$w_3 = \frac{1}{\sqrt{t}}\Big(\frac{10}{t^2}X^2 - \frac{12}{t}X + 3\Big).$$

A base é entón

$$\Big\{\sqrt{\frac{5}{t^5}}X^2, \sqrt{\frac{3}{t}}\Big(\frac{-5}{t^2}X^2 + \frac{4}{t}X\Big), \frac{1}{\sqrt{t}}\Big(\frac{10}{t^2}X^2 - \frac{12}{t}X + 3\Big)\Big\}.$$

Problema 4.11. No espazo euclidiano $\mathbb{R}^3$ consideramos os subespazos

$$F = \{(x,y,z) \in \mathbb{R}^3 \mid x + ay + z = 0\}, \quad G = \{(x,y,z) \in \mathbb{R}^3 \mid x - y = y - bz = 0\},$$

con $a, b \in \mathbb{R}$. Para que valores dos parámetros a e b os subespazos F e G son complementarios? Para que valores son ortogonais?

Solución. Como F ten dimensión 2 e G ten dimensión 1, a condición de que sexan complementarios quere dicir que as tres ecuacións que definen a intersección dean lugar a un sistema compatible determinado. O sistema

$$x - y = y - bz = x + ay + z = 0$$

ten unicamente a solución trivial cando a matriz correspondente ten determinante non nulo. O determinante é $1 + b + ab$, polo que a condición de que sexan complementarios é

$$1 + b + ab \neq 0.$$

Para que sexan ortogonais, tense que cumprir que o subespazo ortogonal a F, que é o xerado polo vector $(1, a, 1)$, coincida co subespazo xerado por G, que é o $(b, b, 1)$. Iso só sucede se $a = b = 1$.

Problema 4.12. O apartado (a) deste problema é o que se coñece como teorema de representación de Riesz; o apartado (b) mostra que en espazos de dimensión infinita, o resultado non sempre é certo.

(a) Sexa E un $\mathbb{R}$-espazo euclidiano de dimensión finita e $\varphi \in E^*$. Demostrar que existe un único vector $u \in E$ tal que

$$\varphi(v) = \langle v, u\rangle$$

para todo $v \in E$.

(b) Sexa $\mathcal{C}([-1,1],\mathbb{R})$ o espazo vectorial das funcións continuas en $[-1,1]$ co produto escalar habitual dado por

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx.$$

Sexa φ o elemento do dual dado por $\varphi(f) = f(0)$. Demostrar que non existe $g \in \mathcal{C}([-1,1],\mathbb{R})$ de maneira que $\varphi(f) = \langle f, g\rangle$ para calquera $f \in \mathcal{C}([-1,1],\mathbb{R})$.

Solución. (a) Trátase dunha cuestión vista en teoría. Se $u \in E$, podemos considerar o elemento $\varphi_u \in E^*$ dado por

$$\varphi_u \colon v \mapsto \langle v, u \rangle.$$

A asignación $u \mapsto \varphi_v$ define unha aplicación lineal inxectiva entre un espazo e o seu dual, que no caso de dimensión finita, é tamén bixectiva. Polo tanto, se $\varphi \in E^*$, existe $u \in E$ de maneira que $\varphi = \varphi_u$.

(b) A estratexia da solución pasará por construír unha sucesión de funcións cuxa integral tenda a 0, pero que en $x = 0$ teñan un valor constante e non nulo. Para iso, supoñamos que existe unha aplicación g como a do enunciado, e sexa $M = \max_{x \in [-1,1]} |g(x)|$. Consideramos as funcións $f_n(x) = (1-|x|)^n$, que cumpren $f_n(0) = 1$ para todo n. Entón

$$\begin{aligned}
\Big| \int_{-1}^{1} f_n(x) g(x)\, dx \Big| &\leq \int_{-1}^{1} |f_n(x) g(x)|\, dx \\
&\leq M \int_{-1}^{1} f_n(x)\, dx \\
&= 2M \int_{0}^{1} (1-x)^n\, dx \\
&= \frac{2M}{n+1},
\end{aligned}$$

e o límite de $\frac{2M}{n+1}$ cando n tende a infinito é 0. Polo tanto, non hai ningunha $g \in \mathcal{C}([-1,1], \mathbb{R})$ que represente φ a través dun produto escalar.

Problema 4.13. Encontrar o punto do plano $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ máis próximo ao punto $(1, 2, 3)$.

Solución. O punto buscado é a proxección ortogonal do punto $(1, 2, 3)$ en W. A recta ortogonal ao plano é $W^\perp = \langle (1, 1, 1) \rangle$. Tense polo tanto que

$$\pi_{W^\perp}((1,2,3)) = \frac{\langle (1,1,1), (1,2,3) \rangle}{\langle (1,1,1), (1,1,1) \rangle} (1,1,1) = (2,2,2).$$

A proxección ortogonal no plano é a diferenza

$$\pi_W((1,2,3)) = (1,2,3) - \pi_{W^\perp}((1,2,3)) = (-1,0,1).$$

Alternativamente, podemos considerar a base do plano $\{(1, 0, -1), (0, 1, -1)\}$ e resolver o sistema lineal

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix},$$

de onde se ten que $(x, y) = (-1, 0)$. Polo tanto, o vector buscado é o $(-1, 0, 1)$.

Problema 4.14. Consideremos o plano $W = \{(x, y, z) \in \mathbb{R}^3 \mid z = x + y\}$. Para cada $v = (a, b, c) \in \mathbb{R}^3$, encontrar o vector do plano que está máis preto de v, e dicir cal é a distancia $d(v, W)$.

Solución. O plano é o ortogonal da súa recta ortogonal, $W^\perp = \langle w \rangle$, onde $w = (1, 1, -1)$. O vector de W máis próximo a v é

$$\begin{aligned}\pi_W(v) &= v - \pi_{W^\perp}(v) = v - \frac{\langle v, w\rangle}{\langle w, w\rangle} w \\ &= (a, b, c) - \frac{a+b-c}{3}(1, 1, -1) \\ &= \frac{1}{3}(2a - b + c, -a + 2b + c, a + b + 2c).\end{aligned}$$

A distancia polo tanto é

$$d(v, W) = \|\pi_{W^\perp}(v)\| = \frac{|a+b-c|}{\sqrt{3}}.$$

Problema 4.15. Sexa $E = \mathcal{M}_2(\mathbb{R})$ e sexa $F \subset E$ un subespazo vectorial tal que, respecto a certo produto escalar en E, a proxección ortogonal sobre F cumpre

$$\pi_F \begin{pmatrix} 0 & 7 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad \pi_F \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Atopar

$$\pi_F \begin{pmatrix} -1 & 0 \\ 2 & 2 \end{pmatrix}.$$

Solución. Un elemento está na imaxe de π_F se, e soamente se, é un punto fixo de π_F. Polo tanto,

$$\pi_F \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad \pi_F \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Polo tanto, sabemos as imaxes de 4 elementos linealmente independentes, e iso é suficiente para achar a de calquera outro. Resolvendo o sistema de ecuacións, temos que

$$\begin{pmatrix} -1 & 0 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 7 \\ 2 & 0 \end{pmatrix} + (-3) \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} + (-2) \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} + 5 \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

Por conseguinte,

$$\begin{aligned}\pi_F \begin{pmatrix} -1 & 0 \\ 2 & 2 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} + (-3) \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} + (-2) \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} + 5 \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix}.\end{aligned}$$

Problema 4.16. (a) En $\mathbb{R}_n[X]$ considérase o produto escalar no cal a base $\{1, X, \dots, X^n\}$ é unha base ortonormal. Sexa $W = \{P(X) \in \mathbb{R}_n[X] \text{ con } P(1) = 0\}$. Dar unha base do complementario ortogonal de W.

(b) Coas notacións do apartado anterior, demostrar que a distancia dun polinomio $P(X) \in \mathbb{R}_n[X]$ ao subespazo W é $\frac{|P(1)|}{\sqrt{n+1}}$.

Solución. (a) O produto escalar neste espazo é o que dá

$$\left\langle \sum_{i=0}^{n} a_i X^i, \sum_{j=0}^{n} b_j X^j \right\rangle = \sum_{i=0}^{n} a_i b_i.$$

O subespazo W está formado polos polinomios $P(X) = \sum_{i=0}^{n} a_i X^i$ tales que $\sum_{i=0}^{n} a_i = 0$. Ten dimensión n e unha base está dada polos polinomios $X - 1, X^2 - 1, \ldots, X^n - 1$. O seu complementario ortogonal está xerado polo polinomio $Q(X) = 1 + X + X^2 + \ldots + X^n$.

(b) Dado un polinomio xenérico $P(X) = \sum_{i=0}^{n} a_i X^i$ temos que

$$\langle P(X), Q(X) \rangle = \sum_{i=0}^{n} a_i = P(1).$$

Polo tanto,

$$\begin{aligned} \pi_W(P(X)) &= P(X) - \pi_{W^\perp}(P(X)) \\ &= P(X) - \frac{\langle Q(X), P(X) \rangle}{\langle Q(X), Q(X) \rangle} Q(X) \\ &= P(X) - \frac{P(1)}{n+1} Q(X). \end{aligned}$$

A distancia de $P(X)$ a esta proxección é a norma de $\pi_{W^\perp}(P(X))$, que é

$$\frac{|P(1)|}{n+1} \|Q(X)\| = \frac{|P(1)|}{\sqrt{n+1}}.$$

Problema 4.17. Consideramos o espazo vectorial $E = \mathbb{R}_2[X]$ co produto escalar definido por

$$\langle p(X), q(X) \rangle = \int_{-1}^{1} p(X) q(X).$$

É unha comprobación inmediata ver que define unha forma bilineal, simétrica e definida (non é preciso facelo!).

(a) Atopar a matriz do produto escalar na base canónica $\{1, X, X^2\}$.

(b) Dar unha base ortonormal de E.

(c) Atopar un elemento non nulo de E que sexa ortogonal a $p_1(X) = 1$ e a $p_2(X) = X$.

(d) Atopar o complemento ortogonal de $F = \langle X^2 \rangle$ en E.

Solución. (a) Facendo os produtos escalares correspondentes, temos que a matriz buscada é

$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix}.$$

(b) Imos dar unha base ortonormal $\{w_1, w_2, w_3\}$ aplicando o algoritmo de Gram–Schmidt a $\{1, X, X^2\}$. Temos que a norma do polinomio 1 é 2, polo que $w_1 = \frac{1}{\sqrt{2}}$. Como $\langle 1, X \rangle = 0$, temos que

$$w_2 = \frac{X}{\|X\|} = \frac{X}{\sqrt{2/3}}.$$

Finalmente, temos que

$$X^2 - \langle X^2, w_1 \rangle w_1 - \langle X^2, w_2 \rangle w_2 = X^2 - \frac{\sqrt{2}}{3} w_1 = X^2 - \frac{1}{3},$$

polo que

$$w_3 = \frac{X^2 - 1/3}{\|X^2 - 1/3\|} = \frac{X^2 - 1/3}{\sqrt{8/45}} = \frac{3X^2 - 1}{2\sqrt{2/5}}.$$

(c) Escribimos ese elemento como $p(X) = \alpha + \beta X + \gamma X^2$. Impondo que $\langle p(X), 1\rangle = 0$, quédanos que

$$2\alpha + \frac{2}{3}\gamma = 0.$$

De xeito similar, pondo $\langle p(X), X\rangle = 0$, temos

$$\frac{2}{3}\beta = 0.$$

Polo tanto, calquera elemento que sexa ortogonal a 1 e a X será un múltiplo de $-1 + 3X^2$. Alternativamente, empregando o apartado anterior, poderiamos ter chegado á mesma conclusión.

(d) Sexa $q(X) = \alpha + \beta X + \gamma X^2$ un elemento no ortogonal de X^2. Impondo que $\langle q(X), X^2\rangle = 0$, temos que

$$\frac{2\alpha}{3} + \frac{2\gamma}{5}.$$

Iso permíteno pór $F = \langle X, -3 + 5X^2\rangle$.

Problema 4.18. En $\mathbb{R}^3$ consideramos o produto escalar que na base canónica ten por matriz

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & a \\ 0 & a & b \end{pmatrix}.$$

Sabemos que a proxección ortogonal do vector $u = (1, 1, 1)$ sobre o subespazo $F\colon x + 2y + 3z = 0$ é o vector $u' = (2, -1, 0)$.

(a) Atopar os valores de a e de b.

(b) Calcular o ángulo entre u e u'.

(c) Dar unha base ortonormal de F e unha de $F^{\perp}$.

(d) Dado un vector $w = (\alpha, \beta, \gamma)$, sexa $w' = \pi_F(w)$. Calcular as coordenadas de w' na base canónica e determinar a matriz da proxección ortogonal sobre F na base canónica.

Solución. (a) A diferenza $u'' = u - u' = (-1, 2, 1)$ é un vector perpendicular a F. Se pomos $F = \langle v_1, v_2\rangle$, con $v_1 = (2, -1, 0)$ e $v_2 = (3, 0, -1)$, temos que se cumpre

$$0 = \langle u'', v_1\rangle = -1 - a, \quad 0 = \langle u'', v_2\rangle = 3 - 2a - b.$$

Polo tanto, $a = -1$ e $b = 5$.

Imos dar unha solución alternativa empregando o teorema da proxección ortogonal. Consideramos os vectores $v_1 = (2, -1, 0)$ e $v_2 = (3, 0, -1)$, que forman unha base de F. A proxección ortogonal de $u = (1, 1, 1)$ en F é un vector que se pode escribir como $u' = xv_1 + yv_2$, onde (x, y) é a solución do sistema

$$\begin{pmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \langle u, v_1\rangle \\ \langle u, v_2\rangle \end{pmatrix}.$$

Calculando os produtos escalares coa matriz A, temos o sistema

$$\begin{pmatrix} 2 & a+3 \\ a+3 & b+9 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -a+1 \\ -a-b+6 \end{pmatrix}.$$

Como sabemos que a solución é $(1,0)$, temos que

$$\begin{cases} 2 = -a+1 \\ a+3 = -a-b+6. \end{cases}$$

Resolvendo o sistema temos que $a=-1$ e $b=5$.

(b) Tense que

$$\cos\alpha = \frac{\langle u, u'\rangle}{\|u\|\|u'\|} = \frac{2}{\sqrt{8}\sqrt{2}} = \frac{1}{2},$$

polo que $\alpha = \pi/3$.

(c) Aplicando o algoritmo de Gram–Schmidt no subespazo F, temos $w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(2,-1,0)$ e $w_2 = \frac{1}{\sqrt{12}}(1,1,-1)$. Alternativamente, podemos impor que $w_2 \in F$ e $\langle w_1, w_2\rangle = 0$. Para a base de $F^\perp$ é suficiente coller $w_2 = \frac{u''}{\|u''\|} = \frac{1}{\sqrt{6}}(-1,2,1)$.

(d) Empregando a base ortonormal que atopamos no apartado anterior, temos que

$$\begin{aligned} w' &= \langle w, w_1\rangle w_1 + \langle w, w_2\rangle w_2 \\ &= \frac{1}{2}(\alpha+\gamma)(2,-1,0) + \frac{1}{12}(2\alpha+4\beta-6\gamma)(1,1,-1) \\ &= \left(\frac{7}{6}\alpha + \frac{1}{3}\beta + \frac{1}{2}\gamma, -\frac{1}{3}\alpha + \frac{1}{3}\beta - \gamma, -\frac{1}{6}\alpha - \frac{1}{3}\beta + \frac{1}{2}\gamma\right). \end{aligned}.$$

Polo tanto, a matriz que proxecta vectores de $\mathbb{R}^3$ sobre F é, na base canónica,

$$\begin{pmatrix} 7/6 & 1/3 & 1/2 \\ -1/3 & 1/3 & -1 \\ -1/6 & -1/3 & 1/2 \end{pmatrix}.$$

Alternativamente poderiamos ter empregado o segundo método desenvolvido no apartado (a).

Problema 4.19. Encontrar o polinomio $p(X) \in \mathbb{R}_3[X]$ tal que $p(0) = p'(0) = 0$ que faga que

$$\int_0^1 (2+3x-p(x))^2\,dx$$

sexa mínimo.

Solución. Un polinomio $p(X) = a + bx + cX^2 + dX^3$ cumpre que $p(0) = p'(0) = 0$ se, e soamente se, $a = b = 0$. Trátase polo tanto de calcular a proxección ortogonal de $2+3X$ no subespazo xerado polos polinomios $p_1(X) = X^2$ e $p_2(X) = X^3$, onde o produto escalar está dado por

$$\langle p(X), q(X)\rangle = \int_0^1 p(X)q(X)\,dX.$$

Temos que
$$\langle X^2, X^2\rangle = \frac{1}{5},\ \langle X^2, X^3\rangle = \frac{1}{6},\ \langle X^3, X^3\rangle = \frac{1}{7}.$$
Por outro lado,
$$\langle X^2, 2+3X\rangle = \frac{17}{12},\ \langle X^3, 2+3X\rangle = \frac{11}{10}.$$
Temos que resolver o sistema
$$\begin{pmatrix} 1/5 & 1/6 \\ 1/6 & 1/7 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 17/12 \\ 11/10 \end{pmatrix}.$$
A solución é $a = 24$ e $b = -20{,}3$, polo que o polinomio é $p(X) = 24X^2 - 20{,}3X^3$.

Problema 4.20. No espazo vectorial $\mathcal{C}^0([-1,1],\mathbb{R})$, consideramos o vector $v = X^3$. Encontrar o polinomio $p(X) \in \mathbb{R}_2[X] \subset \mathcal{C}^0([-1,1],\mathbb{R})$ que minimiza o valor de
$$\int_{-1}^{1} (X^3 - p(X))^2\, dx$$
e determinar cal é dito mínimo.

Solución. En $\mathcal{C}^0([-1,1],\mathbb{R})$, consideramos o produto escalar $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$. Aplicando os resultados sobre o cálculo da proxección ortogonal, resolvemos o sistema
$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 2/5 \\ 0 \end{pmatrix},$$
polo que queda $a = c = 0$ e $b = 3/5$. Polo tanto, o polinomio buscado é $\frac{3}{5}X$. Tense ademais que
$$\int_{-1}^{1} \left(X^3 - \frac{3}{5}X\right)^2 dx = \left(\frac{X^7}{7} - \frac{6X^5}{25} + \frac{3X^3}{25}\right)\Big|_{-1}^{1} = \frac{2}{7} - \frac{12}{25} + \frac{6}{25} = \frac{8}{175}.$$

Problema 4.21. Resolver os seguintes sistemas sobredeterminados comprobando primeiro que as matrices teñen os rangos axeitados. É algún dos sistemas compatible determinado?

(a) $\begin{pmatrix} 1 & 3 \\ -1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}.$

(b) $\begin{pmatrix} 1 & -4 \\ 2 & -1 \\ 3 & 2 \\ 4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ 4 \\ 3 \end{pmatrix}.$

Solución. (a) Se A é a matriz de coeficientes e b é o vector, sabemos que a solución do sistema determinado vén dada por
$$A^tA = \begin{pmatrix} 6 & 2 \\ 2 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} = A^tb,$$

que ten por solución $x = 3/7$ e $y = -2/7$. Tense que

$$\begin{pmatrix} 1 & 3 \\ -1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 3/7 \\ -2/7 \end{pmatrix} - \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4/7 \\ -12/7 \\ -8/7 \end{pmatrix};$$

a norma deste último vector, que é $\frac{4\sqrt{2}}{\sqrt{7}}$, é o que se adoita coñecer como *residuo* do sistema sobredeterminado.

(b) Procedendo como no apartado anterior, $x = 2$ e $y = -1$; ao calcular o residuo vemos que dá 0. Isto quere dicir que o sistema é compatible.

Problema 4.22. Encontrar a recta de regresión que definen os puntos seguintes:

$$(1, -1),\ (3, 0),\ (5, 0),\ (7, 1),\ (9, 2),\ (12, 4).$$

Solución. A recta $m+nx$ é unha combinación lineal das funcións 1 e x. PAra encontrar os coeficientes, temos que resolver o sistema sobredeterminado

$$\begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \\ 1 & 9 \\ 1 & 12 \end{pmatrix} \begin{pmatrix} m \\ n \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 2 \\ 4 \end{pmatrix}.$$

Chamando A á matriz e b ó vector, temos que a solución se obtén ao resolver $A^tAx = A^tb$, e en particular, $(m, n) = \left(\frac{-162}{97}, \frac{42}{97}\right)$. Polo tanto, a recta buscada é

$$y = \frac{-162 + 42x}{97}.$$

Problema 4.23. Encontrade a parábola do tipo $y = a + bx + cx^2$ que aproxime os catro puntos $(0, 1)$, $(1, 3)$, $(2, 4)$ e $(3, 5)$ en $\mathbb{R}^2$ cun erro mínimo.

Solución. Obtemos de novo un sistema sobredeterminado da forma

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} \begin{pmatrix} m \\ n \\ p \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \end{pmatrix}.$$

A solución é $(m, n, p) = (1{,}1, 2{,}05, -0{,}25)$. Polo tanto, a parábola buscada é $y = -0{,}25x^2 + 2{,}05x + 1{,}05$.

Endomorfismos simétricos.

Problema 4.24. Encontrar o endomorfismo simétrico $f \in \mathrm{End}(\mathbb{R}^3)$ que cumpra as seguintes tres condicións.

(a) $f(1, 1, 1) = (2, 2, 2)$.

(b) $\mathrm{Tr}(f) = 5$.

(c) $(1,-1,0) \in \ker(f)$.

Solución. Todo endomorfismo simétrico ten todos os valores propios reais e diagonaliza nunha base ortonormal. Polo tanto, vectores propios de valores propios diferentes son ortogonais.
Da primeira condición temos que $(1,1,1)$ é un vector propio de valor propio 2; e $(1,-1,0)$ é un vector propio de valor propio 0. O terceiro valor propio é 3 xa que a traza é 5, e polo dito no parágrafo anterior, o vector propio correspondente é ortogonal aos outros dous, polo que ten que ser calquera múltiplo de $(1,1,-2)$. Polo tanto, a matriz na base canónica é

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}^{-1} = \frac{1}{6} \begin{pmatrix} 7 & 7 & -2 \\ 7 & 7 & -2 \\ -2 & -2 & 16 \end{pmatrix}.$$

Problema 4.25. Consideremos o espazo $\mathbb{R}^3$ co produto escalar habitual. Encontrar a matriz en base canónica do endomorfismo g que cumpre simultaneamente as seguintes condicións:

(i) g é simétrico.

(ii) O plano $\pi_1 : x + y = 0$ é g-invariante.

(iii) O plano $\pi_2 : x - y = 0$ é g-invariante.

(iv) $g(2,0,1) = (1,1,-1)$.

Solución. A intersección $\pi_1 \cap \pi_2$ é invariante, polo que $(0,0,1)$ é un vector propio. Como g é simétrico, diagonaliza nunha base ortonormal; o mesmo é certo para a restrición de g a calquera subespazo invariante. Se consideramos $g|\pi_1$, temos que $(0,0,1)$ é un vector propio, e polo tanto o outro vector propio ten que ser un que sexa ortogonal a el, isto é, $(1,-1,0)$. Do mesmo modo, $\langle (0,0,1) \rangle^{\perp} \cap \pi_2 = (1,1,0)$ tamén é un vector propio. Poñamos

$$g(0,0,1) = a(0,0,1),\ g(1,-1,0) = b(1,1,0),\ g(1,-1,0) = c(1,-1,0).$$

Sumando as tres ecuacións, temos que

$$g(2,0,1) = (b+c, b-c, a),$$

e da última condición do enunciado é inmediato ver que $a = -1$, $b = 1$ e $c = 0$. Calculando a imaxe dos vectores da base canónica, a matriz buscada é

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Problema 4.26. Encontrar unha matriz ortogonal P_i tal que $P_i^{-1} A_i P_i$ sexa diagonal para as matrices

$$A_1 = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Solución. O teorema espectral real asegura que as tres matrices diagonalizan e que existen matrices ortogonais P_i que cumpren a condición do enunciado e para as que $P_i^{-1} = P_i^t$. Para obtelas, é suficiente con diagonalizalas asegurándose de que a base de vectores propios é ortonormal.
No primeiro caso, $\mathrm{Char}(A_1; X) = (X-6)(X-1)$, e tense que $\ker(A_1 - 6\,\mathrm{Id}) = \langle(-1, 2)\rangle$ e $\ker(A_1 - \mathrm{Id}) = \langle(2, 1)\rangle$. Dividindo pola norma, que é $\sqrt{5}$, obtemos

$$P_1 = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}, \quad P_1^t A_1 P_1 = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}.$$

O segundo exemplo é totalmente análogo e ten como resultado

$$P_2 = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \quad P_2^t A_2 P_2 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

No último caso, no que $\mathrm{Char}(A_3; X) = -(X-1)^2(X-4)$ hai que sei máis coidadosos, xa que $\ker(A_3 - \mathrm{Id})$ ten dimensión 2 e é preciso coller unha base ortonormal (aplicando, por exemplo, Gram–Schmidt, aínda que os cálculos son o suficientemente doados como para facelo a simple vista). Máis concretamente, podemos pór $\ker(A_3 - \mathrm{Id}) = \langle(1, -1, 0), (1, 1, -2)\rangle$ e os dous vectores son ortogonais entre si. Por outra banda, $\ker(A_3 - 4\,\mathrm{Id}) = \langle(1, 1, 1)\rangle$. Dividindo todos os vectores pola norma, quédanos finalmente que

$$P_3 = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}, \quad P_3^t A_3 P_3 = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Problema 4.27. Sexa $f \in \mathrm{End}(E)$ un endomorfismo simétrico dun espazo vectorial real. Demostrar que $\ker(f) = \ker(f^m)$ para todo enteiro $m \geq 1$.

Solución. Imos dar dúas solucións deste problema. Para a primeira, observamos que como o endomorfismo é simétrico, para calquera valor propio λ tense que o subespazo de vectores propios xeralizados coincide co subespazo de vectores propios, polo que $\ker((f-\lambda)) = \ker((f-\lambda)^m)$ se $m \geq 1$. É dicir, o menor m para o cal $\ker(f^m) = \ker(f^{m+1})$ é o tamaño do maior bloque de Jordan correspondente ao valor propio 0, polo que $m \leq 1$ ($m = 1$ se 0 é un valor propio e 0 en caso contrario).
A segunda solución baséase en considerar que calquera endomorfismo f cumpre

$$\ker f \subset \ker f^2 \subset \ldots \subset \ker f^m \subset \ldots$$

Sexa $f \in \ker f^2$. Entón, $\langle f(v), f(v)\rangle = \langle v, f^2(v)\rangle = 0$, e como o produto escalar é definido positivo tense que $f(v) = 0$ e $\ker f^2 = \ker f$. Máis en xeral, se $n \geq 1$ e $v \in \ker f^{n+1}$,

$$\langle f^n(v), f^n(v)\rangle = \langle f^{n-1}(v), f^{n+1}(v)\rangle = 0,$$

de onde se deduce que $v \in \ker f^n$.

Problema 4.28. En $\mathbb{R}_2[X]$ consideramos o produto escalar definido por

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx.$$

Sexa T o endomorfismo de $\mathbb{R}_2[X]$ dado por $T(a_0 + a_1 X + a_2 X^2) = a_1 X$.

(a) Probar que T non é simétrico.

(b) Achar a matriz de T con respecto á base $\{1, X, X^2\}$ e comprobar que é simétrica.

(c) Explicar por que as conclusións dos dous apartados anteriores non son unha contradición entre si.

Solución. (a) Se T fose simétrico teriamos que $\langle T(1), X\rangle = 0$ e $\langle T(X), 1\rangle = \langle X, 1\rangle = \frac{1}{2}$, que obviamente non é certo.

(b) A matriz é

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

que claramente é simétrica.

(c) Que un endomorfismo sexa simétrico quere dicir que é simétrico con respecto a calquera base ortonormal. Se é simétrico con respecto a unha base ortonormal éo con respecto a calquera. En cambio, aquí o que temos é que $\{1, X, X^2\}$ non é unha base ortonormal; por exemplo, $\langle 1, X\rangle = \frac{1}{2} \neq 0$.

Problema 4.29. Sexa V un espazo vectorial hermítico. Probar que calquera endomorfismo normal T en V ten unha raíz cadrada, é dicir, existe un endomorfismo S de maneira que $S^2 = T$. Se V é un espazo euclidiano (en particular, real), é certo que calquera operador autoadxunto en V ten unha raíz cúbica?

Solución. Polo teorema espectral complexo temos que T diagonaliza nunha base ortonormal, é dicir, $T = PDP^*$, onde $P^* = P^{-1}$. Collendo $S = PD^{1/2}P^*$ temos que

$$S^2 = PD^{1/2}P^*PD^{1/2}P^* = PDP^* = T.$$

Observamos que a raíz cadrada non ten por que ser única, xa que para cada elemento da diagonal podemos escoller unha raíz cadrada, co que $D^{1/2}$ non está univocamente determinada.
No caso da raíz cúbica procedemos da mesma forma usando o teorema espectral real e pondo $T = PDP^t$, con $P^t = P^{-1}$. Cada elemento real ten unha única raíz cúbica, polo que se pode coller $PD^{1/3}P^t$.

Problema 4.30. Sexa $(E, \langle\cdot,\cdot\rangle)$ un espazo vectorial de dimensión n sobre $\mathbb{R}$ cun produto escalar e sexa $T : E \to E$ unha aplicación lineal. O obxectivo dos apartados (a), (c) e (d) deste problema é dar unha demostración alternativa de que $\mathrm{Char}(T; X)$ descompón completamente (polo que para eses apartados non se pode empregar o teorema espectral).

(a) Sexan $b, c \in \mathbb{R}$ dous números reais de maneira que $b^2 < 4c$. Demostrar que se T é simétrica, entón a aplicación lineal $T^2 + bT + c\,\mathrm{Id}$ ten inversa.

(b) É certa a conclusión do apartado anterior se non pomos a condición de que T sexa simétrica?

(c) Supoñamos que T é simétrica. Sexa $v \in E$ con $v \neq 0$, e consideremos números reais $a_0, \ldots, a_n$, non todos cero, de maneira que

$$0 = a_0 v + a_1 T v + \ldots + a_n T^n v.$$

Sexa $p(X) = a_0 + a_1 X + \ldots + a_n X^n \in \mathbb{R}[X]$. Aplicando o apartado (a) ao polinomio $P(T)$, demostrar que T ten un valor propio real.

(d) Empregar os apartados anteriores para dar unha demostración alternativa de que o polinomio característico dun endomorfismo simétrico descompón completamente no corpo dos números reais.

(e) Dar unha demostración alternativa do resultado do apartado (a) empregando o teorema espectral.

Solución. (a) Sexa $v \in E$ un vector non nulo, e imos establecer que $(T^2 + bT + c\,\mathrm{Id})v \neq 0$, vendo así que a aplicación é inxectiva. Para iso, observamos que

$$\begin{aligned}\langle (T^2 + bT + c\,\mathrm{Id})(v), v\rangle &= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\ &\geq \|Tv\|^2 - |b|\|Tv\|\|v\| + c\|v\|^2 \\ &= \left(\|Tv\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\ &> 0,\end{aligned}$$

onde na primeira desigualdade empregamos Cauchy–Schwarz para afirmar que

$$|b|\|Tv\|\|v\| \geq |b|\langle Tv, v\rangle \geq -b\langle Tv, v\rangle.$$

(b) Non é certa. Chega con coller o endomorfismo que, na base canónica, ten matriz

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

Tomamos agora $b = 0$ e $c = 1$, que cumpren que $0^2 < 4 \cdot 1$. Cúmprese que $T^2 + \mathrm{Id} = 0$, polo que a aplicación lineal non ten inversa. Máis en xeral, é posible tomar calquera matriz 2×2 que teña un polinomio característico con dúas raíces complexas conxugadas, e tomar b e c para que $\mathrm{Min}(T; X) = X^2 + bX + c$.

(c) Imos proceder por redución ao absurdo, supondo que $p(T)$ non ten ningún valor propio real. Polo tanto, $n = 2k$ e tense que

$$p(X) = a_n \prod_{i=1}^{k} (X^2 + b_i X + c_i),$$

onde se cumpre que $b_i^2 < 4c_i$ para todo i; isto é certo dado que calquera polinomio non constante en $\mathbb{R}[X]$ factoriza como o produto de factores irredutibles de grao 1 ou de grao 2. Ao considerarmos $p(T)$, temos que

$$p(T) = a_n \prod_{i=1}^{k} (T^2 + b_i T + c_i\,\mathrm{Id}),$$

polo que $p(T)$ é un produto de endomorfismos invertibles e, polo tanto, invertible. Iso é unha contradición co feito de que $v \in \ker(p(T))$. Isto permítenos concluír que o polinomio característico de T ten algún valor propio real, xa que $p(X)$ é un divisor do polinomio mínimo e, polo tanto, tamén do polinomio característico.

(d) Imos demostrar o resultado por indución na dimensión do espazo que o polinomio característico dun endomorfismo simétrico descompón completamente sobre os reais. Se a dimensión é 1, o resultado é evidente. Se o resultado é certo para dimensión $n-1$, supoñamos que $\dim E = n$. Polo apartado anterior, sexa λ un valor propio real e v un vector propio asociado a λ. Polo teorema da proxección ortogonal, $E = \langle v\rangle \oplus \langle v\rangle^{\perp}$. Tense, ademais, que $\langle v\rangle^{\perp}$ é un subespazo invariante e, polo tanto,

$$\operatorname{Char}(T;X) = (\lambda - X)\operatorname{Char}(T|\langle v\rangle^{\perp};X).$$

Por hipótese de indución, $\operatorname{Char}(T|\langle v\rangle^{\perp};X)$ descompón completamente, polo que $\operatorname{Char}(T;X)$ tamén.

(e) Empregando o teorema espectral real, podemos tomar unha base ortonormal de vectores propios para T, $\{v_1,\dots,v_n\}$, que existe xa que T é un endomorfismo simétrico. Sexan $\lambda_1,\dots,\lambda_n$ os valores propios correspondentes. Entón,

$$(T^2 + bT + c\,\mathrm{Id})(v_i) = (\lambda_i^2 + b\lambda_i + c)(v_i),$$

polo que $\{v_1,\dots,v_n\}$ tamén é unha base de vectores propios para $T^2 + bT + c\,\mathrm{Id}$. É suficiente agora observar que ningún dos valores propios pode ser 0, xa que a ecuación $X^2 + bX + c = 0$ non ten solucións reais, debido a que o discriminante $b^2 - 4c$ é menor que 0 polas hipóteses do enunciado.

Problema 4.31. Sexa $(E,\langle\cdot,\cdot\rangle)$ un espazo vectorial sobre $\mathbb{C}$ cun produto hermítico e sexa $T : E \to E$ unha aplicación lineal normal.

(a) Demostrar que se T só ten valores propios reais, entón T é autoadxunta.

(b) Supoñamos que $T^9 = T^8$. Demostrar que T é autoadxunta e que $T^2 = T$.

Supoñamos agora que T non é necesariamente normal.

(c) Dar un exemplo dunha aplicación lineal T nun espazo vectorial complexo tal que $T^9 = T^8$, pero $T^2 \neq T$.

Solución. (a) Como o endomorfismo T é normal, diagonaliza nunha base ortonormal, que escribimos como $\{v_1,\dots,v_n\}$; sexan $\lambda_1,\dots,\lambda_n$ os valores propios correspondentes. Entón, tense que

$$\|(T - \lambda_i\,\mathrm{Id})v_i\| = \|(T^* - \bar{\lambda}\,\mathrm{Id})v_i\|.$$

Iso quere dicir que T^* diagonaliza na mesma base e os valores propios son os conxugados dos de T; pero como T só ten valores propios reais temos que T^* ten os mesmos valores propios e os mesmos vectores propios, polo que son iguais.

Alternativamente, podemos empregar o teorema espectral complexo, que garante que existe unha base ortonormal $\{v_1,\dots,v_n\}$ de vectores propios para o endomorfismo T. Nesa base, a matriz de T é $\operatorname{diag}(\lambda_1,\dots,\lambda_n)$, e como a base é ortonormal, sabemos que a matriz de T^* nesa mesma base é $\operatorname{diag}(\lambda_1^*,\dots,\lambda_n^*) = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$. Como as matrices de T e T^* coinciden nesa base, os endomorfismos son iguais.

(b) Imos demostrar en primeiro lugar que T é autoadxunto. Como $T^9 - T^8 = 0$, os valores propios de T están entre os ceros do polinomio mínimo, que son 0 e 1. Polo apartado anterior, se T é unha aplicación normal que só ten valores propios reais podemos asegurar que é autoadxunta.

Para demostrar que $T^2 = T$ observamos que o endomorfismo T diagonaliza por ser autoadxunto. Entón, o polinomio mínimo, que é un divisor de $T^8(T-1)$ só ten factores de grao 1, polo que ten que ser un divisor de $T(T-1)$, que sempre vai ser un polinomio anulador; polo tanto, $T^2 = T$.

(c) É suficiente con considerar o endomorfismo dado pola matriz

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

ou calquera outro para o cal o polinomio mínimo sexa un divisor de $T^8(T-1)$, pero non de $T(T-1)$; é dicir, que haxa un subespazo invariante para o cal T^8 é un polinomio anulador, pero non T.

Isometrías.

Problema 4.32. Sexa $f \in \mathrm{End}(E)$ tal que para todo $u, v \in E$ se cumpre

$$\langle f(u), f(v) \rangle = \lambda \langle u, v \rangle,$$

onde $\lambda > 0$ está fixo.

(a) Demostrar que $\|f(v)\| = \mu \|v\|$ con $\mu = +\sqrt{\lambda}$, para todo $v \in E$.

(b) Demostrar que f é bixectiva.

(c) Demostrar que os únicos valores propios de f son $\pm\mu$.

Solución. (a) Tense que

$$\|f(u)\|^2 = \langle f(u), f(u) \rangle = \lambda \langle u, u \rangle = \lambda \|u\|^2.$$

Tomando raíces cadradas temos o resultado do enunciado.

(b) Supoñamos que $f(u) = 0$. Entón,

$$0 = \langle f(u), f(u) \rangle = \lambda \langle u, u \rangle,$$

polo que $u = 0$. Isto proba que f é inxectiva, e como f é un endomorfismo dun espazo vectorial de dimensión finita, tamén é bixectiva.

(c) Se $f(v) = av$, tomando normas temos que $\|f(v)\| = |a|\|v\|$. Comparando co resultado do primeiro apartado vemos que $\mu = |a|$, polo que $a = \pm\mu$.

Problema 4.33. Sexa $A \in \mathcal{M}_n(\mathbb{R})$, con $A = (a_{ij})$. Demostrar que $C = \sum_{i,j=1}^n a_{ij}^2$ é invariante por cambios de base ortogonais (é dicir, cambios de base nos que a matriz de cambio de base é ortogonal).

Solución. Comezamos notando que $C = \mathrm{Tr}(AA^t)$. Se aplicamos un cambio de base ortogonal, teremos unha nova matriz $B = QAQ^t$; esta cumprirá

$$\begin{aligned} \mathrm{Tr}(BB^t) &= \mathrm{Tr}(QAQ^tQA^tQ^t) \\ &= \mathrm{Tr}(QAA^tQ^t) \\ &= \mathrm{Tr}(QAA^tQ^{-1}) \\ &= \mathrm{Tr}(AA^t). \end{aligned}$$

A segunda igualdade é certa porque $QQ^t = 1$; a terceira porque $Q^t = Q^{-1}$; e a última, porque a traza é invariante por cambios de base.

Imos dar unha solución alternativa. Coas notacións anteriores, pomos $A = (a_{ij})$ e $Q = (q_{ij})$. Temos que $QQ^t = \mathbb{I}$, polo que $\sum_{k=1}^{n} q_{ik}q_{jk} = \delta_{i,j}$ e $\sum_{k=1}^{n} q_{ki}q_{kj} = \delta_{i,j}$. O problema é equivalente a demostrar que

$$\sum_{i,j=1}^{n} \Big(\sum_{k,\ell=1}^{n} q_{ik}a_{k\ell}q_{j\ell} \Big)^2 = \sum_{k,\ell=1}^{n} a_{k\ell}^2.$$

Temos entón que

$$\begin{aligned}
\sum_{i,j=1}^{n} \Big(\sum_{k,\ell=1}^{n} q_{ik}a_{k\ell}q_{j\ell} \Big)^2 &= \sum_{i,j=1}^{n} \Big(\sum_{k_1,\ell_1=1}^{n} q_{ik_1}a_{k_1\ell_1}q_{j\ell_1} \Big)\Big(\sum_{k_2,\ell_2=1}^{n} q_{ik_2}a_{k_2\ell_2}q_{j\ell_2} \Big) \\
&= \sum_{i,j=1}^{n} \sum_{k_1,k_2=1}^{n} \sum_{\ell_1,\ell_2=1}^{n} q_{ik_1}q_{ik_2}q_{j\ell_1}q_{j\ell_2}a_{k_1\ell_1}a_{k_2\ell_2} \\
&= \sum_{k_1,k_2=1}^{n} \sum_{\ell_1,\ell_2=1}^{n} \delta_{k_1,k_2}\delta_{\ell_1,\ell_2}a_{k_1\ell_1}a_{k_2\ell_2} \\
&= \sum_{k,\ell=1}^{n} a_{k\ell}^2.
\end{aligned}$$

Problema 4.34. Sexa A a matriz dada por

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

É A a matriz dunha isometría? Atopar os valores propios complexos de A e comprobar que teñen módulo 1, e achar unha descomposición diagonal ou en bloques 2×2 da forma $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ con $a^2 + b^2 = 1$.

Solución. Para ver que unha matriz A é unha isometría, é suficiente con ver que $A^t = A^{-1}$. Iso é así porque unha isometría é un endomorfismo que preserva normas. Como $\langle Av, Av \rangle = \langle v, A^tAv \rangle$, cómpre que $A^tAv = v$ para todo v, o que quere dicir que $A^tA = \mathbb{I}$. Neste caso, é inmediato ver que A^tA é a matriz identidade.
Por outro lado, a matriz ten determinante 1, e tratándose dunha isometría 3×3 iso que quere dicir que 1 é un valor propio. Podemos ver que $(-1, 1, 1)$ é un vector propio de valor propio 1. A forma reducida será polo tanto da forma

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix}.$$

A traza é $1 + 2\cos\alpha = 2$, polo que $\cos\alpha = 1/2$; porén, iso deixa dúas opcións para $\sin\alpha$, polo que é precisa ou ben algunha interpretación xeométrica ou proceder do xeito habitual.
Os outros valores propios son $\lambda_2 = \frac{1+i\sqrt{3}}{2}$ e $\lambda_3 = \frac{1-i\sqrt{3}}{2}$, que teñen módulo 1. Os vectores propios son (v_2, v_3), con

$$v_2 = \Big(\frac{1-i\sqrt{3}}{2}, \frac{-1-i\sqrt{3}}{2}, 1\Big), \quad v_3 = \Big(\frac{1+i\sqrt{3}}{2}, \frac{-1+i\sqrt{3}}{2}, 1\Big).$$

Definindo $w_2 = \frac{v_2+v_3}{2}$ e $w_3 = \frac{-i(v_2-v_3)}{2}$ temos a base (v_1, w_2, w_3), con

$$v_1 = (-1, 1, 1), \quad w_2 = (1/2, -1/2, 1), \quad w_3 = (-1/2, -1/2, 0),$$

que é ortogonal (pero non ortonormal, para iso habería que dividir cada vector pola súa norma). En calquera caso, a forma reducida é

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & \sqrt{3}/2 \\ 0 & -\sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Convén observar que normalizar w_2 e w_3 dividindo por 2 ou por $\sqrt{2}$, como se fixo ao traballar o teorema de estrutura das isometrías, non afecta ao resultado.

Problema 4.35. Sexa A a matriz dada por

$$A = \begin{pmatrix} 1{,}4 & -1{,}2 & -5{,}2 & 8{,}32 \\ 0{,}6 & 0{,}2 & -2{,}6 & 3{,}64 \\ 0 & 0 & -7{,}4 & 12{,}48 \\ 0 & 0 & -4{,}8 & 7{,}96 \end{pmatrix}.$$

Atopar os valores propios complexos de A e comprobar que teñen módulo 1, e achar unha descomposición en bloques 2×2 da forma $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ con $a^2+b^2 = 1$. É A a matriz dunha isometría? É iso unha contradición con algún resultado teórico?

Solución. Os valores propios de A son $\lambda_1 = \frac{7+24i}{25}$, $\lambda_2 = \frac{7-24i}{25}$, $\lambda_3 = \frac{4+3i}{5}$ e $\lambda_4 = \frac{4-3i}{5}$. Os vectores propios correspondentes son

$$\begin{aligned} v_1 &= \Big(\frac{3-i}{5}, \frac{2+i}{5}, \frac{8-i}{5}, 1\Big) \\ v_2 &= \Big(\frac{3+i}{5}, \frac{2-i}{5}, \frac{8+i}{5}, 1\Big) \\ v_3 &= (1+i, 1, 0, 0) \\ v_4 &= (1-i, 1, 0, 0). \end{aligned}$$

Definimos os vectores $w_1 = \frac{v_1+v_2}{2}$, $w_2 = \frac{-i(v_1-v_2)}{2}$, $w_3 = \frac{v_3+v_4}{2}$ e $w_4 = \frac{-i(v_3-v_4)}{2}$. Entón, temos á súa vez que $v_1 = w_1 + iw_2$, $v_2 = w_1 - iw_2$, $v_3 = w_3 + iw_4$ e $v_4 = w_3 - iw_4$. Para calcular $f(w_1)$, observamos que

$$f(w_1) = f\Big(\frac{v_1+v_2}{2}\Big) = \frac{\lambda_1 v_1 + \lambda_2 v_2}{2} = \frac{7}{25}w_1 - \frac{24}{25}w_2.$$

De xeito análogo chegamos a

$$\begin{aligned} f(w_2) &= \frac{24}{25}w_1 + \frac{7}{25}w_2 \\ f(w_3) &= \frac{4}{5}w_3 - \frac{3}{5}w_4 \\ f(w_4) &= \frac{3}{5}w_3 + \frac{4}{5}w_4. \end{aligned}$$

Observamos que nesta descomposición os vectores da base non son ortornormais, pero a descomposición en matrices de senos e cosenos é a mesma. Máis en concreto, temos que

$$A = \begin{pmatrix} 0{,}6 & -0{,}2 & 1 & 1 \\ 0{,}4 & 0{,}2 & 1 & 0 \\ 1{,}6 & -0{,}2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0{,}28 & 0{,}96 & 0 & 0 \\ -0{,}96 & 0{,}28 & 0 & 0 \\ 0 & 0 & 0{,}8 & 0{,}6 \\ 0 & 0 & -0{,}6 & 0{,}8 \end{pmatrix} \begin{pmatrix} 0{,}6 & -0{,}2 & 1 & 1 \\ 0{,}4 & 0{,}2 & 1 & 0 \\ 1{,}6 & -0{,}2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1}.$$

Para ver que unha matriz A é unha isometría, é suficiente con ver que $A^t = A^{-1}$. Iso é así porque unha isometría é un endomorfismo que preserva normas. Como $\langle Av, Av \rangle = \langle v, A^t Av \rangle$, cómpre que $A^t Av = v$ para todo v, o que quere dicir que $A^t A = \mathbb{I}$. Neste caso, é inmediato ver que $A^t A$ non é a matriz identidade. O feito de que os valores propios teñan módulo 1 e que polo tanto se poida construír unha matriz con bloques de senos e cosenos como a descrita no enunciado non implica que a matriz sexa a dunha isometría: para iso, sería preciso ademais que a base de vectores propios na que diagonaliza fose ortonormal, e ese non é o caso.

Problema 4.36. Consideramos a matriz B que na forma canónica se expresa como

$$\begin{pmatrix} 1+i & 1 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 1 \\ 0 & 0 & 0 & 1-i \end{pmatrix}.$$

Que sucede ao aplicar o procedemento para expresala en bloques 2×2 de senos e cosenos?

Solución. Chamamos (v_1, v_2, v_3, v_4) á base na que a matriz admite forma de Jordan. Tomando a base (w_1, w_2, w_3, w_4), con

$$w_1 = \frac{v_1 + v_3}{2}, \quad w_2 = \frac{-i(v_1 - v_3)}{2}, \quad w_3 = \frac{v_2 + v_4}{2}, \quad w_4 = \frac{-i(v_2 - v_4)}{2}.$$

Nesa base, a matriz escríbese como

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

Isto é o que se coñece como *forma de Jordan real*: se a matriz non diagonaliza en $\mathbb{C}$, podemos escribila en forma de bloques 2×2 e en vez de ter uns na diagonal superior o que teremos é bloques 2×2 correspondentes a $\mathbb{I}_2$ por riba dos bloques da forma $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

Descomposición en valores singulares.

Problema 4.37. Atopar a descomposición en valores singulares de

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

e atopar a matriz de rango 1 máis próxima a ela coa norma euclidiana.

Solución. Comezamos considerando A^tA:

$$A^tA = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Os valores propios correspondentes son 3, 1 e 0, e a matriz ortogonal V que ten por columnas os vectores propios correspondentes de norma 1 é

$$V = \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Por outro lado,

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Finalmente, para achar os vectores u_1 e u_2 que forman as columnas de U facemos $u_1 = \frac{1}{\sqrt{3}}Av_1$ e $u_2 = Av_2$. Entón

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Polo tanto, a descomposición buscada é da forma

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}^t.$$

A matriz de rango 1 máis próxima a ela, polo teorema de Eckart–Young–Mirsky, é

$$A_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}^t = \begin{pmatrix} 1 & -1/2 & 1/2 \\ -1 & 1/2 & -1/2 \end{pmatrix}.$$

Problema 4.38. Consideramos o endomorfismo de $\mathbb{R}^3$ dado por $f(x,y,z) = (\alpha x + y, x + \alpha y, \alpha z)$, con $\alpha \geq 1$, e sexa A a súa matriz na base canónica.

(a) Xustificar que f diagonaliza en $\mathbb{R}$ e encontrar os seus valores propios.

(b) Dar unha descomposición en valores singulares de A.

(c) Encontrar α para que 3 sexa o valor máximo de $\|f(x,y,z)\|$ entre os vectores (x,y,z) de norma 1.

Solución. (a) A matriz na base canónica é

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

O polinomio característico é $\mathrm{Char}(A; X) = -(X - (\alpha - 1))(X - (\alpha + 1))(X - \alpha)$, polo que as raíces son α, $\alpha + 1$ e $\alpha - 1$. Un vector propio para α é o $(0,0,1)$; para $\alpha + 1$ é o $(1/\sqrt{2}, 1/\sqrt{2}, 0)$ e para $\alpha - 1$ é o $(1/\sqrt{2}, -1/\sqrt{2}, 0)$.

(b) Como a matriz é simétrica e diagonaliza nunha base ortonormal, unha descomposición da forma $A = PDP^{-1}$, con $P^{-1} = P^t$ é tamén unha descomposición en valores singulares. Nese caso,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha+1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha-1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}^t.$$

(c) O valor máximo correspóndese co maior valor singular, é dicir, con $\alpha + 1$. Polo tanto, $\alpha = 2$.

Problema 4.39. Consideremos a matriz

$$A = \begin{pmatrix} a & b & c & c \\ b & a & c & c \\ c & c & a & b \\ c & c & b & a \end{pmatrix} \in \mathcal{M}_4(\mathbb{R}),$$

con $a > b > c > 0$.

(a) Sabendo que $w_1 = (1, 1, 1, 1)$, $w_2 = (1, 1, -1, -1)$ e $w_3 = (1, -1, 1, -1)$ son vectores propios, dar os valores propios de A (coas súas multiplicidades) e unha base ortonormal de vectores propios.

(b) Calcular unha descomposición en valores singulares de A. Como cambiaría a resposta se $b > a > c > 0$?

(c) No caso particular $a = 4$, $b = 2$ e $c = 1$, dar a matriz de rango 2 máis próxima a A.

(d) Repetir o apartado (b) no caso no que $b > a > c > 0$.

Solución. (a) Como A é unha matriz simétrica, o teorema espectral afirma que diagonaliza nunha base ortonormal de vectores propios. Para os vectores propios do enunciado, temos que

$$Aw_1 = (a + b + 2c)w_1, \quad Aw_2 = (a + b - 2c)w_2, \quad Aw_3 = (a - b)w_3.$$

Polo tanto, $a+b+2c$, $a+b-2c$ e $a-b$ son valores propios de A (con $a+b+2c > a+b-2c > a-b$). Para encontrar o outro valor propio, usamos que a súa suma é igual á traza, que é $4a$. Polo tanto, o outro valor propio é $a - b$. Temos entón que $a-b$ é un valor propio con multiplicidade xeométrica 2. Considerando o subespazo propio asociado, temos que unha base ortogonal de vectores propios está dada por $\{(1, -1, 1, -1), (1, -1, -1, 1)\}$.

Alternativamente, como $\langle w_1, w_2, w_3 \rangle$ é invariante, o ortogonal tamén o será. Isto quere dicir que

$$\langle w_1, w_2, w_3 \rangle^\perp = \langle (1, -1, -1, 1) \rangle$$

é un subespazo propio de valor propio $a - b$. Dividindo pola norma, podemos coller a base ortonormal

$$v_1 = \frac{1}{2}w_1, \quad v_2 = \frac{1}{2}w_2, \quad v_3 = \frac{1}{2}w_3, \quad v_4 = \frac{1}{2}(1, -1, -1, 1).$$

(b) Como A é simétrica, $S := A^tA = A^2$ e os valores propios de S son λ_i^2, onde λ_i son os valores propios de A. A matriz V da descomposición en valores singulares está formada polos vectores v_1, v_2, v_3, v_4 en columna. Os valores singulares de A son por tanto $\sqrt{\lambda_i^2} = |\lambda_i|$. Iso quere dicir que os valores singulares son

$$\sigma_1 = a + b + 2c, \quad \sigma_2 = a + b - 2c, \quad \sigma_3 = \sigma_4 = a - b,$$

xa que polas desigualdades do enunciado todos eles son positivos. Entón,

$$D = \begin{pmatrix} a+b+2c & 0 & 0 & 0 \\ 0 & a+b-2c & 0 & 0 \\ 0 & 0 & a-b & 0 \\ 0 & 0 & 0 & a-b \end{pmatrix}.$$

Como D coincide coa matriz de valores propios, a decomposición $A = VDV^t$ que diagonaliza A (notemos que $V^t = V^{-1}$ xa que V é unha matriz ortogonal por ser un cambio de base entre bases ortogonais) xa nos dá a SVD con $U = V$. Se feito, se calculamos $u_i = \frac{1}{\sigma_i}Av_i$ obtemos $u_i = \frac{1}{\lambda_i}\lambda_i v_i = v_i$ para todo $i = 1, 2, 3, 4$.

(c) A matriz de rango 2 máis próxima a A obtense considerando os dous primeiros valores singulares na SVD, $\sigma_1 = a + b + 2c = 8$ e $\sigma_2 = a + b - 2c = 4$. Entón, obtemos

$$V \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} V^t = \begin{pmatrix} 3 & 3 & 1 & 1 \\ 3 & 3 & 1 & 1 \\ 1 & 1 & 3 & 3 \\ 1 & 1 & 3 & 3 \end{pmatrix}.$$

(d) Se $b > a > c > 0$, os valores singulares son

$$\sigma_1 = a + b + 2c, \quad \sigma_2 = a + b - 2c, \quad \sigma_3 = \sigma_4 = b - a,$$

xa que $a + b - 2c \geq b - a$. A matriz V é a mesma (e polo tanto V^t tamén), pero a matriz U ten agora por columnas os vectores $(v_1, v_2, -v_3, -v_4)$, xa que $u_3 = \frac{1}{\sigma_3}Av_3 = -v_3$ e o mesmo ocorre para u_4.

Problema 4.40. Consideramos o espazo vectorial $E = \mathcal{M}_n(\mathbb{R})$ co produto escalar $\langle A, B\rangle = \operatorname{Tr}(A^tB)$ e a norma correspondente $\|A\|_F = \sqrt{\operatorname{Tr}(A^tA)}$. Sexan F_s e F_a os subespazos formados polas matrices simétricas e antisimétricas, respectivamente.

(a) Demostrar que $F_a = F_s^\perp$ e determinar a matriz simétrica máis próxima a unha matriz dada $A \in E$ con respecto á norma que consideramos.

(b) Expresar $\|A\|_F$ en termos dos valores singulares da matriz A e caracterizar para que matrices A se cumpre $\|A\|_F = \|A\|_2$.

Solución. (a) Imos comezar comprobando que $F_a \subset F_s^\perp$, é dicir, que se $A \in F_a$ e $B \in F_s$, entón $\langle A, B\rangle = 0$. Comezamos observando que

$$\langle A, B\rangle = \operatorname{Tr}(A^tB) = \operatorname{Tr}(-AB) = -\operatorname{Tr}(AB),$$

onde usamos que a traza é lineal. Por outro lado,

$$\langle B, A\rangle = \operatorname{Tr}(B^tA) = \operatorname{Tr}(BA) = \operatorname{Tr}(AB),$$

onde agora empregamos que $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. Como $\langle A, B\rangle = \langle B, A\rangle$, concluísmo que $\mathrm{Tr}(AB) = -\mathrm{Tr}(AB)$, polo que $\mathrm{Tr}(AB) = 0$ e, polo tanto, $\langle A, B\rangle = 0$.

Unha vez demostramos que $F_a \subset F_s^\perp$, para que ver son iguais é suficiente comprobar que as dimensións de ambos coinciden. Tense que $\dim(F_a) = \frac{n^2-n}{2}$, xa que se podemos escoller arbitrariamente o valor de todos os coeficiente a_{ij} con $i > j$; por outro lado, $\dim(F_s) = \frac{n^2+n}{2}$, xa que aquí podemos escoller o valor dos a_{ij} con $i \geq j$. Entón, $\dim(F_s^\perp) = n^2 - \dim(F_s) = \frac{n^2-n}{2}$.

Polo teorema da proxección ortogonal, todo elemento A se escribe de forma única como suma un de F_s e un de F_a; máis en concreto, podemos pór

$$A = \Big(\frac{A + A^t}{2}\Big) + \Big(\frac{A - A^t}{2}\Big).$$

Polo tanto, a matriz máis próxima a $A \in E$ será $A = \frac{A+A^t}{2}$, e a distancia virá dada pola norma de $\frac{A-A^t}{2}$.

(b) Polo visto en teoría, se os valores singulares de A son $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$, entón temos que $\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}$ e $\|A\|_2 = \sigma_1$. Polo tanto, cómpre que $\sigma_2 = \ldots = \sigma_r = 0$. En particular, isto quere dicir que o rango de A^tA é como moito 1.

Problema 4.41. Sexa T un endomorfismo dun espazo euclidiano V. Sexa $\hat{s}$ o menor valor singular de T e sexa s o maior valor singular de T.

(a) Probar que $\hat{s}\|v\| \leq \|Tv\| \leq s\|v\|$ para cada $v \in V$.

(b) Sexa λ un valor propio de T. Probar que $\hat{s} \leq |\lambda| \leq s$.

Solución. (a) Observamos que se U é unha matriz ortogonal, entón $\|Uv\| = \|v\|$ xa que

$$\|Uv\|^2 = \langle Uv, Uv\rangle = \langle v, U^tUv\rangle = \|v\|^2.$$

Por outra banda, se D é unha matriz diagonal da forma $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, con $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$, temos que $\lambda_n\|v\| \leq \|Dv\| \leq \lambda_1\|v\|$. Polo tanto, se consideramos a descomposición en valores singulares de T, $T = U\Sigma V^t$, temos que $\|Tv\| = \|U(\Sigma V^t v)\| = \|\Sigma V^t v\|$. Entón,

$$\hat{s}\|v\| = \hat{s}\|V^t v\| \leq \|\Sigma V^t v\| \leq s\|V^t v\| = s\|v\|.$$

(b) Sexa v un vector propio de T con valor propio λ. Entón, $Tv = \lambda v$. Nese caso, $\|Tv\| = |\lambda|\|v\|$. Polo tanto, usando o apartado anterior temos que

$$\hat{s} \leq |\lambda| \leq s,$$

como queriamos.

Problema 4.42. Achar a descomposición en valores singulares $U\Sigma V^t$ da matriz

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

Pódese usar que $\mathrm{Char}(A^tA; X) = -X(X-9)(X-25)$.

Solución. Comezamos observando que

$$A^tA = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix},$$

e tense entón que $\mathrm{Char}(A^tA) = -X(X-9)(X-25)$. Os vectores propios normalizados (v_1, v_2, v_3) asociados aos valores propios $(25, 9, 0)$, respectivamente, veñen dados por

$$v_1 = \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\Big),\ v_2 = \Big(\frac{1}{3\sqrt{2}}, \frac{-1}{3\sqrt{2}}, \frac{4}{3\sqrt{2}}\Big),\ v_3 = \Big(\frac{-2}{3}, \frac{2}{3}, \frac{1}{3}\Big).$$

Do mesmo xeito, temos que os valores singulares son $\sigma_1 = \sqrt{25} = 5$ e $\sigma_2 = \sqrt{9} = 3$. Finalmente, os vectores columna da matriz U son

$$u_1 = \frac{1}{5}Av_1 = \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\Big),\ u_2 = \frac{1}{3}Av_2 = \Big(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\Big).$$

Polo tanto, a descomposición buscada é

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & \frac{1}{3} \end{pmatrix}^t.$$

Formas cuadráticas.

Problema 4.43. Clasificar as seguintes formas cuadráticas encontrando unha forma reducida, o rango e o índice. Indicade se son definidas positivas, definidas negativas ou non definidas.

(a) $q(x, y, z) = 2x^2 + 5y^2 + 2z^2 - 4xy - 2xz + 4yz$ en $\mathbb{R}^3$.

(b) $q(x, y, z, t) = 2xt + 6yz$ en $\mathbb{R}^4$.

Solución. No primeiro caso, a matriz da forma cuadrática é

$$\begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}.$$

Aplicando congruencia pivote chegamos a que a forma reducida é $2\bar{x}^2 + 3\bar{y}^2 + \frac{7}{6}\bar{z}^2$ e que unha matriz de cambio de base é

$$S = \begin{pmatrix} 1 & 1 & 1/6 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix}.$$

O rango é 3, o índice de inercia positivo é 3, o negativo é 0 e a forma cuadrática é definida positiva.
No segundo caso, a matriz da forma cuadrática é

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Aplicando congruencia pivote chegamos a que a forma reducida é $2\bar{x}^2+6\bar{y}^2-\frac{3}{2}\bar{z}^2-\frac{1}{2}\bar{t}^2$ e que unha matriz de cambio de base é

$$S=\begin{pmatrix}1 & 0 & 0 & -1/2\\ 0 & 1 & -1/2 & 0\\ 0 & 1 & 1/2 & 0\\ 1 & 0 & 0 & 1/2\end{pmatrix}.$$

O rango é 4, o índice de inercia positivo é 2, o negativo é 2 e a forma cuadrática é non definida.

Problema 4.44. Dar a clasificación afín e proxectiva, encontrando unha forma reducida, o rango e o índice, das formas cuadráticas que teñen as seguintes matrices nas bases canónicas:

$$\begin{pmatrix}4 & 2 & -2\\ 2 & 5 & 1\\ -2 & 1 & 3\end{pmatrix},\ \begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 5 & 8 & 11\\ 3 & 8 & 14 & 20\\ 4 & 11 & 20 & 30\end{pmatrix}.$$

Solución. No primeiro caso, aplicando congruencia pivote chegamos a que unha forma reducida é $4\bar{x}^2+4\bar{y}^2+\bar{z}^2$ e que unha matriz de cambio de base é

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

O rango é 3, o índice de inercia positivo é 3, o negativo é 0 e a forma cuadrática é definida positiva.
No segundo caso, aplicando congruencia pivote chegamos a que a forma reducida é $\bar{x}^2+\bar{y}^2+\bar{z}^2+\bar{t}^2$ e que unha matriz de cambio de base é

$$S=\begin{pmatrix}1 & -2 & 1 & 0\\ 0 & 1 & -2 & 1\\ 0 & 0 & 1 & -2\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

O rango é 4, o índice de inercia positivo é 4, o negativo é 0 e a forma cuadrática é definida positiva.

Problema 4.45. Consideramos a forma cuadrática $q\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$ con $q(x,y,z)=2xy+2yz+2zx$.

(a) Encontrar unha base u de $\mathbb{R}^3$ na que a matriz de q sexa diagonal.

(b) Encontrar un subespazo vectorial $F\subset\mathbb{R}^3$ de dimensión máxima tal que $q|_F$ é definida negativa.

(c) Razoar se existe algún subespazo vectorial $G\subset\mathbb{R}^3$ de dimensión 2 e tal que $q|_G$ teña rango 1.

Solución. (a) A matriz da forma cuadrática é

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

Facendo congruencia pivote, temos que coa matriz de cambio

$$S = \begin{pmatrix} 1 & -1/2 & -1 \\ 1 & 1/2 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

a forma cuadrática é diagonal:

$$q(\bar{x}, \bar{y}, \bar{z}) = 2\bar{x}^2 - (1/2)\bar{y}^2 - 2\bar{z}^2.$$

(b) Consideramos $F = \langle (-1, 1, 0), (-1, -1, 1) \rangle$, que se corresponde cunha combinación lineal do segundo e do terceiro vector na base adaptada. A matriz da restrición da forma bilineal a F é

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

(c) É suficiente coller un vector isótropo, por exemplo $v_1 = (1, 0, 0)$, e completar cun v_2 de maneira que se ϕ é a forma bilineal asociada se teña que $\phi(v_1, v_2) = 0$, e $q(v_2) \neq 0$. Por exemplo, se collemos $v_2 = (1, -1, 0)$ observamos que a matriz de $q|G$ é

$$\begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}.$$

Observamos que para achar v_2 consideramos a aplicación lineal

$$\phi_{v_1} : E \longrightarrow K, \quad v \mapsto \langle v_1, v \rangle,$$

para a cal o núcleo ten dimensión 2, polo que é suficiente con coller un elemento que xunto con v_1 forme parte do núcleo.

Problema 4.46. Sexan $q, q' : \mathbb{R}^3 \to \mathbb{R}$ as formas cuadráticas

$$q(x, y, z) = 4xy + y^2 - 8xz - 4z^2$$
$$q'(x, y, z) = -x^2 + 2xy - z^2.$$

(a) Calcular o rango e a signatura de q. Diagonalizar q e obter a base asociada.

(b) Son (V, q) e (V, q') espazos cuadráticos isométricos? En caso afirmativo, obter unha isometría entre eles.

(c) Obter 4 subespazos de (V, q) de vectores isótropos de dimensión 1.

(d) Razoar se existen ou non subespazos de (V, q) de vectores isótropos de dimensión 2.

Solución. (a) A matriz de q na base canónica é

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

Aplicando congruencia-pivote, temos que a forma reducida é $q(\bar{x}, \bar{y}, \bar{z}) = \bar{y}^2 - 4\bar{z}^2$ e a matriz de cambio de base é

$$S = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Polo tanto, o rango é 2 e a sinatura é $(i_+, i_-) = (1, 1)$.

(b) Non son espazos isométricos xa que a matriz de q' é

$$B = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

e ten rango 3. A súa sinatura é $(i_+, i_-) = (1, -2)$.

(c) Observamos que $q(\bar{x}, \bar{y}, \bar{z}) = (\bar{y} + \bar{2}z)(\bar{y} - \bar{2}z)$, polo que podemos coller os subespazos xerados por vectores que na base dada polas columnas de S son da forma $(a, -2b, b)$ ou $(a, 2b, b)$, con $a, b \in \mathbb{R}$. Na base canónica, son os vectores da forma $(a, -2a - 2b, -a + b)$ e $(a, -2a + 2b, -a + b)$.

(d) Collemos por exemplo o plano $\bar{y} + \bar{2}z = 0$, que ten dimensión 2 e que unicamente consta de vectores isótropos. Na base canónica é o subespazo xerado por $(1, -2, -1)$ e por $(0, -2, 1)$. Alternativamente, é o subespazo de dimensión 2 con ecuación $4x + y + 2z = 0$.

Poderiamos coller tamén $\bar{y} - \bar{2}z = 0$. Nese caso, é o subespazo xerado por $(1, -2, -1)$ e $(0, 2, 1)$, cuxa ecuación é $y - 2z = 0$.

Problema 4.47. Consideremos en $\mathbb{R}^3$ as formas cuadráticas definidas polas matrices Q e Q',

$$Q = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}, \quad Q' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(a) Demostrar que as formas cuadráticas son equivalentes e calcular dous cambios de base diferentes que permitan pasar dunha a outra.

(b) Achar o radical e o cono isotrópico do espazo $(\mathbb{R}^3, Q)$.

(c) Existe un complementario ortogonal do subespazo

$$U = \{2x + y = 0\}$$

correspondente á forma bilineal definida por Q? (Cando temos unha forma bilineal, a noción de perpendicularidade é a mesma que no caso de produtos escalares.)

Solución. (a) Aplicando congruencia pivote, chegamos a que Q admite a mesma forma reducida que Q' empregando como matriz de cambio de base

$$S = \begin{pmatrix} -1 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 2 & 0 & \sqrt{2} \\ 1 & 0 & 0 \end{pmatrix}.$$

Os vectores columna de S forman unha base $v = (v_1, v_2, v_3)$ na que a matriz é diagonal e ten a forma de Q'. É inmediato ver que Q' admite a mesma forma diagonal se consideramos unha base da forma $v = (v_1, v_2 + av_1, v_3 + bv_1)$, onde $a, b \in \mathbb{R}$. Isto é así porque v_1 está no radical; alternativamente, podemos comprobar que

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Polo tanto, se $v_1 = (-1,2,1)$, $v_2 = \left(\frac{\sqrt{2}}{2},0,0\right)$, $v_3 = \left(-\frac{\sqrt{2}}{2},\sqrt{2},0\right)$, podemos coller as bases (v_1,v_2,v_3) ou (v_1,v_1+v_2,v_1+v_3) para pasar á base diagonal. Unha alternativa consiste en empregar $-\sqrt{2}$ en lugar de $\sqrt{2}$ no último paso do método de congruencia pivote; iso permite obter outra base diferente das anteriores.

(b) Para achar o radical é suficiente con pór

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Achamos por tanto que o radical é o subespazo xerado polo vector $(-1,2,1)$.

Para o cono isotrópico, traballando coa forma reducida, temos que se corresponde, nas coordenadas asociadas á base v, con

$$(\bar{y}+\bar{z})(\bar{y}-\bar{z}).$$

Polo tanto, é a unión de dous planos: $F_1 = \langle(v_1, v_2 - v_3)$ e $F_2 = (v_1, v_2+v_3)$. Máis explicitamente,

$$F_1 = \langle (0,1,1),(1,-1,0)\rangle, \quad F_2 = \langle (1,0,-1),(0,1,0)\rangle.$$

(c) Para achar un complementario ortogonal, temos que impor que $U^\perp$ estea en suma directa con U e que sexa ortogonal a el. Unha base de U está dada por $(0,0,1)$ e por $(1,-2,0)$. Entón, se impomos que un vector (a,b,c) sexa ortogonal a ambos, temos que $b-2c=0$. Iso é un subespazo de dimensión 2, e a intersección con U é o subespazo xerado por $(-1,2,1)$; polo tanto, é suficiente con coller calquera vector de $U^\perp$ que non sexa un múltiplo de $(-1,2,1)$. Por exemplo, collemos o $(1,0,0)$.

Problema 4.48. No corpo dos números reais, consideremos a forma cuadrática dada por

$$q(x,y,z,t) = x^2 + y^2 + 4z^2 - t^2 + 2xz + 2xt - 4yz + 4zt = 0.$$

(a) Achar unha forma reducida de q e indicar cal é a matriz do cambio de base. Indicar o rango e os índices de inercia de q. É definida positiva, definida negativa ou non definida?

(b) Dar catro subespazos isótropos de dimensión 1 para a forma cuadrática q (é dicir, subespazos nos que todos os vectores sexan isótropos).

(b) Dar un subespazo isótropo de dimensión 2 para a forma cuadrática q.

(d) Supoñamos agora que a forma cuadrática $q(x,y,z,t)$ está definida sobre un corpo xenérico. Dar tres exemplos de corpos, K_1, K_2 e K_3, nos que $q(x,y,z,t)$ sexa (afinmente) equivalente a $x^2+y^2+z^2+t^2$.

Solución. (a) Aplicando o método de congruencia pivote, temos que unha forma reducida é

$$q(\bar{x},\bar{y},\bar{z},\bar{t}) = \bar{x}^2 + \bar{y}^2 - \bar{z}^2 - \bar{t}^2.$$

A matriz do cambio de base é

$$S = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A partir da forma reducida temos o seguinte.

- **Rango.** $r = 4$.
- **Índice.** $i = 2$. En particular, $(i_+, i_-) = (2, 2)$.
- **Non definida.** Nunha forma diagonal hai tanto valores positivos como negativos.

(b) Chamamos (u_1, u_2, u_3, u_4) aos vectores columna de S. Podemos coller entón como vectores isótropos $u_1+u_3 = (0, 2, 1, 0)$, $u_1+u_4 = (-1, 2, 1, 1)$, $u_2+u_3 = (-1, 3, 1, 0)$ e $u_2 + u_4 = (-2, 3, 1, 1)$. Como os vectores son linealmente independentes, podemos coller os subespazos xerados por cada un deles.

(c) O subespazo que ten por ecuacións na base reducida $\bar{x} = \bar{z}$ e $\bar{y} = \bar{t}$ é un subespazo isótropo de dimensión 2. En termos de xeradores, é o subespazo

$$\langle (0, 2, 1, 0), (-2, 3, 1, 1) \rangle = \langle (0, 2, 1, 0), (-2, 1, 0, 1) \rangle.$$

(d) Ten que se cumprir que -1 sexa un cadrado en K. É suficiente con considerar os complexos, $\mathbb{Z}/5\mathbb{Z}$ e $\mathbb{Z}/13\mathbb{Z}$. Máis en xeral, serve calquera $\mathbb{Z}/p\mathbb{Z}$, con p primo da forma $p = 4k + 1$ para $k \geq 1$, $k \in \mathbb{Z}$. Convén observar que se $p = 2$ non é certo, xa que o método de congruencia pivote non ten por que funcionar e a forma cuadrática inicial é $q(x, y, z, t) = x^2 + y^2 + t^2$, polo que o seu rango é 3.

Aínda que isto non se pedía, imos incluír unha demostración do feito de que se p é un primo impar, entón -1 é un cadrado módulo p se, e soamente se, $p \equiv 1 \pmod{4}$. Cando $p \equiv 1 \pmod{4}$, é dicir, $p = 4k + 1$, é suficiente con considerar $x = (2k)!$ e observar que

$$x^2 = (2k)! \cdot (2k)! = (2k!) \cdot (-1)(2k) \cdot (-1)(2k-1) \ldots (-1)1 = (4k)! = (p-1)! \equiv -1,$$

onde a última congruencia é consecuencia do chamado teorema de Wilson, que afirma que se p é un primo, entón $(p - 1)! \equiv -1$ módulo p. A demostración do teorema, á súa vez, é practicamente inmediata: módulo p, os únicos números que son inversos deles mesmos son o 1 e o -1, xa que $x^2 \equiv 1 \pmod{p}$ só ten 2 solucións por tratarse dunha ecuación sobre un corpo. Polo tanto, podemos agrupar os números desde o 2 ao $p - 2$ co seu inverso, de xeito que ao facer o produto obtemos 1, e polo tanto $(p - 1)! = 1 \cdot (p - 1) \equiv -1 \pmod{p}$.

Por outro lado, supoñamos que $p = 4k+3$ e que existe x de maneira que $x^2 \equiv -1 \pmod{p}$. Entón, elevando ao cadrado, $x^4 \equiv 1 \pmod{p}$. Ao mesmo tempo, polo pequeno teorema de Fermat, $x^{4k+2} \equiv 1 \pmod{p}$. Pola identidade de Bézout, sabemos que existen enteiros a e b de maneira que $4a + (4k + 2)b = 2$. Entón,

$$x^2 = x^{4a} \cdot x^{(4k+2)b} \equiv 1 \cdot 1 = 1 \pmod{p},$$

que é unha contradición co feito de que $x^2 \equiv 1 \pmod{p}$.

Problema 4.49. No corpo dos números reais, consideramos $q(x,y,z,t) = 2(xy+yz+zt)$.

(a) Encontrar unha base na que admita unha forma reducida e dar os seus índices de inercia positivo e negativo.

(b) É única a base do apartado anterior? En caso negativo, dar outra base na que $q(x,y,z,t)$ admita unha forma reducida.

(c) Dar catro subespazos isótropos de dimensión 1 e catro de dimensión 2.

Solución. (a) A matriz da aplicación lineal asociada é

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Aplicando congruencia-pivote, tense que a forma reducida é

$$q(\bar{x},\bar{y},\bar{z},\bar{t}) = 2\bar{x}^2 - \frac{1}{2}\bar{y}^2 + 2\bar{z}^2 - \frac{1}{2}\bar{t}^2$$

e a base correspondente vén dada polas columnas da matriz

$$S = \begin{pmatrix} 1 & -1/2 & -1 & 1/2 \\ 1 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 1 & 1/2 \end{pmatrix}.$$

Os índices de inercia positivo e negativo son ambos iguais a 2.

(b) Non, a base non é única; por exemplo, ao facermos congruencia pivote podemos multiplicar calquera fila por -1 e aplicar a mesma transformación na columna correspondente. Isto deixa invariante a forma diagonal, pero modifica no signo un dos vectores da base. Por exemplo, podemos coller

$$S' = \begin{pmatrix} -1 & -1/2 & -1 & 1/2 \\ -1 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 1 & 1/2 \end{pmatrix}.$$

(c) Na base adaptada, temos que os subespazos dados por $\langle(1,2,0,0)\rangle$, $\langle(-1,2,0,0)\rangle$, $\langle(0,0,1,2)\rangle$ e $\langle(0,0,-1,2)\rangle$ son isótropos. Na base canónica, correspóndense cos subespazos

$$\begin{aligned} U_1 &= \langle(0,1,0,0)\rangle \\ U_2 &= \langle(1,0,0,0)\rangle \\ U_3 &= \langle(0,0,0,1)\rangle \\ U_4 &= \langle(1,0,-1,0)\rangle. \end{aligned}$$

De xeito similar, na base adaptada, para calquera elección de signos,

$$V_i = \langle(\pm1,2,0,0),(0,0,\pm1,2)\rangle$$

é un subespazo isótropo de dimensión dous, xa que

$$q(\pm a, 2a, \pm b, 2b) = 2a^2 - 2a^2 + 2b^2 - 2b^2 = 0.$$

Cómpre observar que non é suficiente coller dous vectores isótropos calquera, xa que nese caso a suma podería non selo. Na base canónica, os subespazos V_i veñen dados por

$$\begin{aligned} V_1 &= \langle (0,1,0,0), (0,0,0,1) \rangle \\ V_2 &= \langle (1,0,0,0), (0,0,0,1) \rangle \\ V_3 &= \langle (0,1,0,0), (1,0,-1,0) \rangle \\ V_4 &= \langle (1,0,0,0), (1,0,-1,0) \rangle. \end{aligned}$$

Problema 4.50. Sexa $\varphi : \mathcal{M}_2(\mathbb{R}) \times \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ a forma bilineal simétrica definida por $\varphi(A, B) = \operatorname{Tr}(AB) - \operatorname{Tr}(A)\operatorname{Tr}(B)$. Sexa $q = q_\varphi$ a forma cuadrática asociada a φ.

(a) Encontrar unha forma reducida de q e unha base q-ortogonal.

(b) Encontrar un subespazo vectorial $F \subset \mathcal{M}_2(\mathbb{R})$, de dimensión máxima, e tal que $\varphi(\mathbb{I}_2, B) = 0$, para todo $B \in F$. Dar o rango e os índices de inercia de $q|F$.

Solución. (a) Imos comezar comprobando que, efectivamente, se trata dunha aplicación bilineal e simétrica. Tendo en conta que a traza é unha aplicación bilineal tense que

$$\begin{aligned} \varphi(\lambda_1 A_1 + \lambda_2 A_2, B) &= \operatorname{Tr}((\lambda_1 A_1 + \lambda_2 A_2)B) - \operatorname{Tr}(\lambda_1 A_1 + \lambda_2 A_2)\operatorname{Tr}(B) \\ &= \lambda_1 \operatorname{Tr}(A_1 B) + \lambda_2 \operatorname{Tr}(A_2 B) \\ &\quad - \lambda_1 \operatorname{Tr}(A_1)\operatorname{Tr}(B) - \lambda_2 \operatorname{Tr}(A_2)\operatorname{Tr}(B) \\ &= \lambda_1 \varphi(A_1, B) + \lambda_2 \varphi(A_2, B). \end{aligned}$$

De xeito análogo temos que $\varphi(A, \lambda_1 B_1 + \lambda_2 B_2) = \lambda_1 \varphi(A, B_1) + \lambda_2 \varphi(A, B_2)$. A simetría é inmediata, xa que $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$, polo que $\varphi(A, B) = \varphi(B, A)$.

Temos que, por definición, $q : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ é a forma cuadrática

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

A matriz de q na base natural de $\mathcal{M}_2(\mathbb{R})$ é

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

Aplicando congruencia pivote, temos que unha forma reducida está dada por

$$q(\bar{x}, \bar{y}, \bar{z}, \bar{t}) = -2\bar{x}^2 + 2\bar{y}^2 - 2\bar{z}^2 + 2\bar{t}^2.$$

Unha base ordenada (v_1, v_2, v_3, v_4) na que a matriz é q-ortogonal está dada por

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ v_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ v_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \ v_4 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(b) Cúmprese que $\varphi(\mathbb{I}_2, B) = \mathrm{Tr}(\mathbb{I}_2, B) - \mathrm{Tr}(\mathbb{I}_2)\,\mathrm{Tr}(B) = \mathrm{Tr}(B) - 2\,\mathrm{Tr}(B)$. Polo tanto, $\varphi(\mathbb{I}_2, B) = 0$ se, e soamente se, $\mathrm{Tr}(B) = 0$. Polo tanto, o subespazo de dimensión máxima é

$$F = \{B \in \mathcal{M}_2(\mathbb{R}) \text{ con } \mathrm{Tr}(B) = 0\}.$$

Unha base de F é (u_1, u_2, u_3), con

$$u_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ u_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ u_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

A matriz da restrición $q|F$ é

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Os seus valores propios son 2, 1 e -1, polo que o rango é 3 os índices de inercia positivos e negativos son 2 e 1, respectivamente.

Problema 4.51. (a) Dada a forma cuadrática de $\mathbb{R}^3$ definida por $q_\alpha(x, y, z) = x^2 + 2y^2 + z^2 + 2(xy + yz + \alpha xz)$, encontrar o rango e os índices de inercia positivo e negativo de q_α en función do parámetro α. Atopar tamén unha base de $\mathbb{R}^3$ na que a forma bilineal asociada sexa diagonal.

(b) Sexa $q_\alpha(x, y, z)$ a forma cuadrática do apartado anterior. Para os valores de α que sexa posible, dar un subespazo $F \subset \mathbb{R}^3$, $F \neq 0$, de maneira que a forma bilineal asociada sexa identicamente cero en F. Cando non sexa posible, explicar por que.

(c) Existe algún valor de α para o cal sexa posible tomar o F do apartado anterior de dimensión 2?

Solución. (a) Aplicando o método de congruencia-pivote temos que unha forma reducida é

$$q_\alpha(\bar{x}, \bar{y}, \bar{z}) = \bar{x}^2 + \bar{y}^2 + 2\alpha(1 - \alpha)\bar{z}^2,$$

e unha matriz de cambio de base é

$$S = \begin{pmatrix} 1 & -1 & 1 - 2\alpha \\ 0 & 1 & \alpha - 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Polo tanto, temos o seguinte.

- O rango é 2 cando $\alpha = 0$ ou $\alpha = 1$ e 3 no resto dos casos.
- O índice de inercia positivo é 3 cando $\alpha \in (0, 1)$ e 2 cando $\alpha \in (-\infty, 0] \cup [1, +\infty)$.
- O índice de inercia negativo é 1 cando $\alpha \in (-\infty, 0) \cup (1, +\infty)$ e 0 no resto dos casos.

A matriz da forma bilineal é definida positiva cando $\alpha \in (0, 1)$ e semidefinida positiva se $\alpha = 0$ ou $\alpha = 1$. En caso contrario é non definida.

(b) A matriz da forma bilineal é definida positiva cando $\alpha \in (0, 1)$, polo que nese caso non existe ningún subespazo non trivial no que a forma bilineal asociada sexa dexenerada.

- Se $\alpha = 0$ ou $\alpha = 1$ temos que $q(1-2\alpha, \alpha-1, 1) = 0$, polo que $F = \langle(1-2\alpha, \alpha-1, 1)\rangle$.
- Se $\alpha \in (-\infty, 0) \cup (1, +\infty)$, o subespazo que queremos coller é o xerado, na base para a cal q_α é diagonal, polo vector $(0, \sqrt{-2\alpha(1-\alpha)}, 1)$. Na base canónica temos entón

$$F = \left\langle(-\sqrt{-2\alpha(1-\alpha)} + 1 - 2\alpha, \sqrt{-2\alpha(1-\alpha)} + \alpha - 1, 1)\right\rangle.$$

Por exemplo, se $\alpha = 9$, ese é o subespazo xerado polo vector $(-29, 20, 1)$.

(c) Supoñamos que F se puidese coller de dimensión 2. Sexa $\{v_1, v_2\}$ unha base de F, e $(\{v_1, v_2, v_3\}$ a ampliación a unha base de E. Entón, chamando ϕ_α á forma bilineal asociada, tense que

$$2\phi_\alpha(v_1, v_2) = q_\alpha(v_1 + v_2) - q_\alpha(v_1) - q_\alpha(v_2),$$

polo que a matriz de ϕ_α é da forma

$$\begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ a & b & c \end{pmatrix},$$

con $a, b, c \in \mathbb{R}$. En particular, o rango é como moito 2, o que quere dicir que $\alpha = 0$ ou $\alpha = 1$. Porén, nese caso a forma cuadrática é equivalente a $q_0(x, y, z) = x^2 + y^2$, e os únicos vectores (x, y, z) para os que $q_0(x, y, z)$ son os da forma $(0, 0, z)$. Isto demostra que nunca poden existir subespazos isótropos de dimensión 2.

Problema 4.52. Un problema habitual en física consiste na *diagonalización simultánea* de dúas formas cuadráticas sobre os reais, é dicir, en atopar unha base de vectores propios na que a matriz de ambas sexa diagonal. Por exemplo, en problemas de oscilacións é frecuente ter que resolver ecuacións diferenciais da forma $Mq'' = -Kq$, onde tanto M como K son matrices simétricas definidas positivas.

Sexan A e B dúas matrices simétricas e supoñamos tamén que B é definida positiva. Como B define un produto escalar, partindo da base canónica de $\mathbb{R}^n$, podemos atopar unha base $u = \{u_1, \dots, u_n\}$ que sexa ortonormal para ese produto escalar. A matriz de B na base u é a identidade, e se C é a matriz que pasa de u á base canónica, entón $\mathbb{I} = C^tBC$. Na nova base, a matriz de A é $\bar{A} = C^tAC$. Como A é unha matriz simétrica, polo teorema espectral sabemos que diagonaliza nunha base ortonormal $v = \{v_1, \dots, v_n\}$. Nesta nova base, a matriz de B segue a ser a identidade, xa que se trata dun cambio de base por unha matriz ortonormal.

Aplicar o procedemento discutido no apartado anterior para diagonalizar simultaneamente as formas cuadráticas dadas por

$$A(x, y) = -4xy, \quad B(x, y) = x^2 - 2xy + 4y^2.$$

Solución. A matriz de B,

$$B = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix},$$

define un produto escalar, polo que podemos comezar coa base canónica $\mathcal{B}_e = \{e_1, e_2\}$ e aplicar Gram–Schmidt para obter unha base ortonormal, $\mathcal{B}_u = \{u_1, u_2\}$. Temos que $u_1 = (1, 0)$ e para obter o segundo vector temos que normalizar

$$e_2 - \langle e_2, u_1\rangle u_1 = e_2 + u_1 = (1, 1),$$

polo que $u_2 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$. A matriz de A na base u é

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

Temos que $\mathrm{Char}(\bar{A}; X) = \frac{1}{3}(3X^2 + 4X - 4)$, que ten por raíces $X = -2$ e $X = \frac{2}{3}$. O vector propio correspondente a $X = -2$ ten coordenadas $(1, \sqrt{3})$ na base u, polo que se corresponde co vector

$$v_1 = \lambda(u_1 + \sqrt{3}u_2) = \lambda(e_1 + (e_1 + e_2)) = \lambda(2e_1 + e_2).$$

Como $B(2,1) = 4$ tomamos $\lambda = 1/2$ e $v_1 = e_1/2 + e_2$. Por outra banda, o vector propio correspondente a $X = 2/3$ ten coordenadas $(\sqrt{3}, -1)$ na base u, polo que se corresponde co vector

$$v_2 = \mu\left(\sqrt{3} - \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right).$$

Normalizando, tense que $v_2 = \frac{\sqrt{3}}{3}e_1 - \frac{\sqrt{3}}{6}e_2$. Polo tanto, a matriz de B na base $v = (v_1, v_2)$ é a identidade, mentres que a de A correspóndese coa forma cuadrática $-2\bar{x}^2 + \frac{2}{3}\bar{y}^2$.

Problema 4.53. Sexa $E = \mathbb{R}^4$. Consideramos a forma bilineal dada por

$$\begin{aligned} \phi(x, y) = &-x_1y_2 + 2x_1y_4 + x_2y_1 + 3x_2y_4 \\ &+ 2x_3y_4 - 2x_4y_4 - 3x_4y_2 - 2x_4y_3 \end{aligned}.$$

Demostrar que ϕ é unha forma simpléctica e achar unha base simpléctica para ϕ.

Solución. A matriz de ϕ é

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

Como é antisimétrica e non dexenerada, trátase da matriz dunha forma simpléctica. Comezando coa base canónica $\{e_1, e_2, e_3, e_4\}$ e aplicando o procedemento habitual, podemos coller $v_1 = (1, 0, 0, 0)$ e $v_2 = (0, -1, 0, 0)$, de xeito que $\phi(v_1, v_2) = 1$. Deste xeito,

$$e_3' = e_3 - \phi(e_3, v_2)v_1 + \phi(e_3, v_1)v_2 = (0, 0, 1, 0)$$

e

$$e_4' = e_4 - \phi(e_4, v_2)v_2 + \phi(e_4, v_1)v_2 = (-3, 2, 0, 1).$$

Como $\phi(e_3', e_4') = 2$, podemos coller $v_3 = (0, 0, 1, 0)$ e $v_4 = (-3/2, 1, 0, 1/2)$. Neste caso, (v_1, v_3, v_2, v_4) forman a base simpléctica buscada.

Problema 4.54. Sexa ϕ a forma bilineal que na base canónica de $\mathbb{R}^3$ ten por matriz

$$\begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}.$$

(a) É ϕ unha forma simpléctica?

(b) Sexa $F = \{v \in \mathbb{R}^3$ tales que $\phi(v, \cdot) = 0\}$ e sexa G un complementario de F. Demostrar que $\phi|G$ é unha forma simpléctica e achar unha base simpléctica.

Solución. (a) Non é unha forma simpléctica xa que a matriz é singular.

(b) Temos que $F = \langle(0, 2, 1)\rangle$, polo que podemos coller $G = \langle(1, 0, 0), (0, 1, 0)\rangle$. A matriz de $\phi|G$ é
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$
Se collemos entón $G' = \langle(1, 0, 0), (0, -1, 0)\rangle$, obtemos unha base simpléctica.

Capítulo 5

Tensores e álxebra tensorial

Ao longo deste capítulo, sexa E un espazo vectorial de dimensión finita sobre un corpo K, e sexa $E^* = \mathcal{L}(E, K)$ o seu espazo dual. Escribiremos $\mathcal{B}_v = \{v_1, \ldots, v_n\}$ para referirnos a unha base de E e $\mathcal{B}_v^* = \{v_1^*, \ldots, v_n^*\}$ para a súa base dual. Recordamos que entón $v_i^*(v_j) = \delta_{i,j}$.

5.1. Definicións e nocións básicas

Definición 5.1. Un *(p, q)-tensor*, ou un *tensor p-covariante, q-contravariante* é unha aplicación multilineal

$$f \colon E \times \overset{p}{\cdots} E \times E^* \times \overset{q}{\cdots} \times E^* \longrightarrow K.$$

O conxunto destes elementos denótase como $T_p^q(E)$. Se $q = 0$, pomos simplemente $T_p(E)$, e isto son os tensores *p-covariantes*; se $p = 0$, escribimos $T^q(E)$, e falamos de tensores *q-contravariantes.* Cando tanto p como q son diferentes de 0, fálase de *tensores mixtos.*

Para abreviar notación, ás veces poremos $E^{\times p}$ para referirnos a $E \times \overset{p}{\cdots} \times E$. Convén recordar que $T_1(E) = E^*$ é o espazo dual.
Un primeiro exemplo que non apareceu no estudo dos determinantes corresponde á situación de $T^1(E) = \mathcal{L}(E^*, K)$, que é o conxunto de aplicacións lineais $f \colon E^* \longrightarrow K$. Este espazo chámase o espazo bidual de E e denótase como E^{**}. A seguinte proposición afirma que é posible dar un isomorfismo entre E e E^{**} que non depende da elección dunha base; en xeral, un isomorfismo así entre E e E^* non existe, aínda que si sucedía no caso dos espazos euclidianos.

Proposición 5.1. A aplicación $i \colon E \to E^{**}$ que envía $v \mapsto \hat{v}$, onde $\hat{v} \colon E^* \to K$ está definida por $\hat{v}(\omega) = \omega(v)$. está ben definida, é lineal e inxectiva. Se E ten dimensión finita, entón é un isomorfismo.

Demostración. O elemento $\hat{v}$ pertence ao bidual por construción; ademais, a linealidade de i é evidente. Para ver a inxectividade, se $\hat{v}$ é 0, entón temos que $\omega(v) = 0$ para todo ω; se v fose diferente de 0, habería algunha aplicación lineal que ao avaliala en v dese un resultado non cero. Como non é o caso, temos que i é inxectiva e, se a dimensión é finita, entón é tamén un isomorfismo. $\square$

Exemplo. En dimensión infinita o morfismo segue a ser inxectivo, pero non é certo que sexa un isomorfismo. Aqueles espazos para os que i induce un isomorfismo co seu bidual chámanse *reflexivos*. Se entendemos como dual o conxunto de aplicacións lineais, sen impor ningunha condición extra (o chamado *dual alxébrico*), entón os únicos espazos reflexivos son os de dimensión finita. Os exemplos máis interesantes aparecen en análise funcional ao tratar con espazos de Banach, que son espazos vectoriais completos con respecto a unha norma; neste caso, acostúmase falar de *dual topolóxico*, xa que na definición pídese tamén que as aplicacións lineais sexan continuas con respecto á norma. Un caso importante ten que ver cos denominados espazos ℓ^p:

$$\ell^p = \Big\{(a_n)_{n\geq 1} \text{ con } a_n \in \mathbb{R} \text{ e } \sum_{n=1}^{\infty} |a_n|^p < \infty\Big\}.$$

Se $1 < p < q$, entón $\ell^p \subset \ell^q$ e a inclusión é estrita, é dicir, hai elementos en ℓ^q que non están en ℓ^p. O dual (topolóxico) de ℓ^p é ℓ^q, onde $q = \frac{p}{p-1}$. Porén, cúmprese que $(\ell^p)^{**} \simeq \ell^p$. Consideremos agora os espazos

$$c_0 = \Big\{(a_n)_{n\geq 1} \text{ con } a_n \in \mathbb{R} \text{ e } \lim_{n\to\infty} a_n = 0\Big\},$$
$$\ell^\infty = \Big\{(a_n)_{n\geq 1} \text{ con } a_n \in \mathbb{R} \text{ e } \sup |a_n| < \infty\Big\}.$$

Temos que $c_0 \subset \ell_1 \subset \ell^\infty$, e novamente as inclusións son estritas. Non é complicado ver que o dual de c_0 é isomorfo a ℓ_1, e á súa vez o dual deste último é isomorfo a ℓ^∞; é dicir, o bidual de c_0 é isomorfo a ℓ^∞, que non coincide con c_0, senón que simplemente o contén. Isto ilustra que en moitos contextos **non** é posible identificar un espazo co seu bidual.

Convén ter en conta que, usando a identificación entre un espazo vectorial e o seu bidual, temos que $T_p^q(E) \cong T_q^p(E^*)$. Deste xeito, calquera resultado que probemos para $T_p(E)$ pode trasladarse a $T^q(E) \cong T_q(E^*)$. Por outro lado, $T_2(E)$ é

$$T_2(E) = \{f\colon E \times E \to K \text{ con } f \text{ bilineal}\}.$$

Por exemplo, os produtos escalares son tensores 2-covariantes. Un último caso sinxelo que convén ter presente é o dos tensores mixtos de tipo $(1,1)$, como por exemplo $f\colon E \times E^* \to K$, con $f(x,\omega) = \omega(x)$.
Como sucedía no caso covariante, os tensores p-covariantes e q-contravariantes teñen estrutura de espazo vectorial, e ademais pódese definir un produto tensorial.

Definición 5.2. Sexan $f \in T_p^q(E)$ e $g \in T_r^s(E)$, de xeito que

$$f\colon E^{\times p} \times E^{*\times q} \longrightarrow K, \quad g\colon E^{\times r} \times E^{*\times s} \longrightarrow K.$$

O *produto tensorial* é a aplicación

$$f \otimes g\colon E^{\times(p+r)} \times E^{*\times(q+s)} \longrightarrow K$$

tal que, se $x = (x_1, \ldots, x_p) \in E^{\times p}$, $y = (y_1, \ldots, y_r) \in E^{\times r}$, $u = (u_1, \ldots, u_q) \in E^{*\times r}$ e $v = (v_1, \ldots, v_s) \in E^{*\times s}$, entón $(f \otimes g)(x, y, u, v) = f(x, u)g(y, v)$.

É inmediato comprobar a linealidade da aplicación definida a través do produto tensorial, é dicir, que $f \otimes g \in T_{p+r}^{q+s}(E)$. Tense ademais que o produto tensorial é asociativo e cumpre a propiedade distributiva. Máis concretamente, se $f, f' \in T_p^q(E)$, $g, g' \in T_r^s(E)$, $h \in T_u^v(E)$ e $\lambda \in K$, entón:

(a) $(f \otimes g) \otimes h = f \otimes (g \otimes h)$.

(b) $f \otimes (g + g') = f \otimes g + f \otimes g'$ e $(f + f') \otimes g = f \otimes g + f' \otimes g$.

(c) $(\lambda f) \otimes g = \lambda(f \otimes g) = f \otimes (\lambda g)$.

Como sucedía no caso covariante estudado ao introducir os determinantes, é posible dar unha base do espazo $T_p^q(E)$. Para iso, sexa $\mathcal{B}_v = (v_1, \ldots, v_n)$ unha base de E e sexa $\mathcal{B}_v^* = (v_1^*, \ldots, v_n^*)$ a súa base dual. De cara aos seguintes resultados, identificamos directamente un espazo vectorial co seu bidual por estarmos en dimensión finita.

Proposición 5.2. O espazo vectorial $T_p^q(E)$ ten dimensión n^{p+q} e unha base está dada por

$$\mathcal{B}_p^q = \{v_{i_1}^* \otimes \cdots \otimes v_{i_p}^* \otimes v_{j_1} \otimes \cdots \otimes v_{j_q} \mid 1 \leq i_1, \ldots, i_p, j_1, \ldots, j_q \leq n\}.$$

Demostración. É unha comprobación rutineira análoga á xa traballada no caso covariante. Primeiro, vese que é un conxunto xerador, xa que calquera tensor se pode expresar en termos de elementos de $\mathcal{B}_p^q$, e a maneira de conseguir o coeficiente correspondente é avaliando nos diferentes $v_{i_1}^* \otimes \cdots \otimes v_{i_p}^* \otimes v_{j_1} \otimes \cdots \otimes v_{j_q}$. É dicir, podemos pór

$$w = \sum_{i_1, \ldots, j_q = 1}^{n} \lambda_{i_1, \ldots, j_q} v_{i_1}^* \otimes \cdots \otimes v_{i_p}^* \otimes v_{j_1} \otimes \cdots \otimes v_{j_q},$$

con $\lambda_{i_1, \ldots, j_q} = w(v_{i_1}, \ldots, v_{j_q}^*)$.
Para ver que son linealmente independentes, faise a mesma avaliación para probar que todos os coeficientes teñen que ser 0: pondo

$$w = \sum_{i_1, \ldots, j_q = 1}^{n} \lambda_{i_1, \ldots, j_q} v_{i_1}^* \otimes \cdots \otimes v_{i_p}^* \otimes v_{j_1} \otimes \cdots \otimes v_{j_q} = 0,$$

ao considerar $w(v_{i_1}, \ldots, v_{j_q}^*)$ chegamos a que $\lambda_{i_1, \ldots, j_q} = 0$. □

5.2. Tensores simétricos e antisimétricos

Como xa introducimos ao estudar os determinantes, podemos dicir que o grupo simétrico *actúa* no espazo dos tensores.

Definición 5.3. Sexa G un grupo e X un conxunto. Unha *acción* de G en X é unha aplicación $G \times X \longrightarrow X$, que escribiremos $g \cdot x$, que cumpre as seguintes condicións.

(a) $h \cdot (g \cdot x) = (hg) \cdot x$ para todo $g, h \in G$ e $x \in X$.

(b) $1 \cdot x = x$ para todo x.

No caso dos tensores, definimos $\sigma \cdot f$, onde $\sigma \in S_p$, como o tensor tal que

$$\sigma \cdot f(v_1, \ldots, v_p) = f(v_{\sigma(1)}, \ldots, v_{\sigma(p)}).$$

Isto define unha acción do grupo simétrico S_p no espazo $T_p(E)$, que ademais respecta a suma e o produto por escalares, é dicir,

$$\sigma \cdot (\lambda u + \mu v) = \lambda \sigma \cdot u + \mu \tau \cdot v.$$

Adoita falarse dunha acción dun grupo nun espazo vectorial cando se cumpre esta propiedade, é dicir, cando se respecta a estrutura de espazo vectorial.
Ao longo desta materia atopamos outras accións dun grupo nun conxunto. Por exemplo, o grupo $\mathrm{GL}_n(K)$ de matrices invertibles actúa *por conxugación* no conxunto das matrices $\mathcal{M}_n(K)$, é dicir, se $P \in \mathrm{GL}_n(K)$ e $A \in \mathcal{M}_n(K)$, definimos $P \cdot A := PAP^{-1}$. Outro exemplo é o das isometrías: as matrices ortogonais actúan no conxunto de vectores de norma fixada; por exemplo, se v ten norma 1 e $M \in \mathcal{O}_n(\mathbb{R})$, entón $M \cdot v = Mv$, xa que $\|Mv\| = \|v\| = 1$. Dicimos entón que as isometrías actúan nos espazos euclidianos, xa que preservan tanto a estrutura de espazo vectorial como a norma.
As accións son interesantes para entender o comportamento e a estrutura dos grupos. Un concepto útil é o de *órbita*: dicimos que dous elementos $x, y \in X$ están na mesma órbita se existe $g \in G$ de maneira que $g \cdot x = y$.

Proposición 5.3. Sexan $u_1, \ldots, u_p \in E^*$ e $\sigma \in S_p$ con $\tau = \sigma^{-1}$. Entón, $u_1 \otimes \cdots \otimes u_p \in T_p(E)$ e

$$\sigma \cdot (u_1 \otimes \cdots \otimes u_p) = u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)}.$$

Demostración. Escribimos $\sigma(1) = i_1$, $\sigma(2) = i_2$) e así ata $\sigma(p) = i_p$. Se $\tau = \sigma^{-1}$, entón $\tau(i_1) = 1, \tau(i_2) = 2, \ldots, \tau(i_p) = p$. Deste xeito, dados $x_1, \ldots, x_p \in E$ temos que

$$\begin{aligned}\sigma(u_1 \otimes \cdots \otimes u_p)(x_1, \ldots, x_p) &= u_1(x_{\sigma(1)}) \cdots u_p(x_{\sigma(p)}) \\ &= u_{\tau(i_1)}(x_{i_1}) \cdots u_{\tau(i_p)}(x_{i_p}) \\ &= u_{\tau(1)}(x_1) \cdots u_{\tau(p)}(x_p) \\ &= (u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)})(x_1, \ldots, x_p).\end{aligned}$$

Polo tanto, $\sigma \cdot (u_1 \otimes \cdots \otimes u_p) = u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)}$. □

Exemplo. Se consideramos o ciclo $\sigma = (1, 2, 3)$, a súa acción sobre o tensor $u_1 \otimes u_2 \otimes u_3$ é

$$\sigma \cdot u_1 \otimes u_2 \otimes u_3 = u_3 \otimes u_1 \otimes u_2,$$

xa que $\sigma^{-1} = (1, 3, 2)$.

Recordamos a seguinte definición que xa traballamos cando presentamos os determinantes.

Definición 5.4. Un tensor $f \in T_p(E)$ dise que é *simétrico* se $\sigma \cdot f = f$ para todo $\sigma \in S_p$. O subespazo vectorial dos tensores simétricos denótase por $S_p(E)$.
Dise que $f \in T_p(E)$ é *antisimétrico* se $\sigma \cdot f = \mathrm{sgn}(\sigma) f$ para todo $\sigma \in S_p$. O subespazo vectorial dos tensores antisimétricos denótase por $A_p(E)$.

A partir de agora, imos supoñer que K ten característica cero, o que nos permitirá dividir por $p!$ na seguinte definición. Nese caso, os tensores antisimétricos chámanse tamén *alternados*, unha noción que xa se explorou e que equivale á de antisimétrico cando a característica é diferente de 2.

Definición 5.5. Sexan $S\colon T_p(E) \to T_p(E)$ e $A\colon T_p(E) \to T_p(E)$ as aplicacións definidas por

$$S(f) = \frac{1}{p!} \sum_{\sigma \in S_p} \sigma \cdot f, \quad A(f) = \frac{1}{p!} \sum_{\sigma \in S_p} \mathrm{sgn}(\sigma) \sigma \cdot f.$$

Chamámolos *operadores de simetrización* e *antisimetrización*.

Por exemplo, se $f \in T_2(E)$ temos que

$$S(f)(x,y) = \frac{f(x,y)+f(y,x)}{2}, \quad A(f)(x,y) = \frac{f(x,y)-f(y,x)}{2}.$$

É interesante comparar esta descomposición coa que sucede no caso das matrices: toda matriz $A \in \mathcal{M}_n(K)$ nun corpo de característica diferente de 2 se pode descompoñer como suma dunha simétrica e dunha antisimétrica do xeito

$$A = \Big(\frac{A+A^t}{2}\Big) + \Big(\frac{A-A^t}{2}\Big).$$

A seguinte proposición resume as propiedades máis relevantes dos operadores de simetrización e antisimetrización.

Proposición 5.4. Sexa $f \in T_p(E)$.

(a) Os operadores S, $A \colon T_p(E) \to T_p(E)$ son aplicacións lineais.

(b) $S(f) \in S_p(E)$ e $A(f) \in A_p(E)$.

(c) $S(f) = f$ se, e soamente se, $f \in S_p(E)$, e $A(f) = f$ se, e soamente se, $f \in A_p(E)$.

(d) As aplicacións lineais S e A son proxectores, é dicir, $S^2 = S$ e $A^2 = A$.

(e) Tense que $S_p(E) = \mathrm{im}(S)$ e $A_p(E) = \mathrm{im}(A)$. Ademais, $T_p(E) = S_p(E) \oplus \ker(S)$ e $T_p(E) = A_p(E) \oplus \ker(A)$.

Demostración. A linealidade é consecuencia directa da definición. Para ver que $S(f) = f$ é suficiente con demostrar que $\sigma \cdot S(f) = S(f)$; iso séguese do feito que

$$\sum_{\tau \in S_p} \tau \cdot (\sigma f) = \sum_{\tau \in S_p} (\tau\sigma) \cdot f = \sum_{\rho \in S_p} \rho \cdot f \in S_p,$$

xa que a multiplicación por σ induce un isomorfismo en S_p. O caso alternado é igual. Supoñamos agora que $S(f) = f$; entón, como $S(f) \in S_p(E)$, tamén ten que suceder que $f \in S_p(E)$, e o mesmo no caso alternado. Para ver que $S(S(f)) = S(f)$ simplemente se observa que cada sumando na fórmula de $S(f)$ é da forma $\sigma \cdot f$, e cúmprese que $S(\sigma \cdot f) = S(f)$ polo visto con anterioridade.
Un elemento está na imaxe se, e soamente se, é un tensor simétrico, polo tanto a última parte da proposición é evidente recordando que calquera endomorfismo f que cumpre $f^2 = f$ cumpre que $E = \mathrm{im}(f) \oplus \ker(f)$. O mesmo se cumpre para os alternados. □

Unha última propiedade importante de cara ao posterior estudo do produto exterior é a seguinte.

Proposición 5.5. Sexan $f \in T_p(E)$ e $\sigma \in S_p$. Entón $\sigma \cdot A(f) = A(\sigma \cdot f) = \mathrm{sgn}(\sigma) A(f)$.

Demostración. Sabemos que se $g \in A_p(E)$ entón $\sigma \cdot g = \mathrm{sgn}(\sigma) g$. Polo tanto, como $A(f) \in A_p(E)$ tense que $\sigma \cdot A(f) = \mathrm{sgn}(\sigma) A(f)$. Alternativamente, podemos facer a comprobación empregando o carácter multiplicativo do signo, é dicir, que $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$:

$$\begin{aligned}
\sigma \cdot A(f) &= \sigma \cdot \Big(\frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\tau)\tau\cdot f\Big) \\
&= \frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\tau)(\sigma\tau)\cdot f \\
&= \operatorname{sgn}(\sigma)\frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\sigma\tau)(\sigma\tau)\cdot f \\
&= \operatorname{sgn}(\sigma)\frac{1}{p!}\sum_{\rho\in S_p} \operatorname{sgn}(\rho)\rho\cdot f \\
&= \operatorname{sgn}(\sigma)A(f).
\end{aligned}$$

Observamos que se empregou que para un σ fixado, o conxunto $\{\sigma\tau \text{ con } \tau\in S_p\}$ coincide co grupo simétrico S_p.
Do mesmo xeito,

$$\begin{aligned}
A(\sigma\cdot f) &= \frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\tau)\tau\cdot(\sigma\cdot f) \\
&= \frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau\sigma)(\tau\sigma)\cdot f \\
&= \operatorname{sgn}(\sigma)\frac{1}{p!}\sum_{\tau\in S_p} \operatorname{sgn}(\tau\sigma)(\tau\sigma)\cdot f \\
&= \operatorname{sgn}(\sigma)A(f).
\end{aligned}$$

□

5.3. Produto exterior

O produto tensorial de tensores antisimétricos non é, polo xeral, antisimétrico (ao igual que sucede tamén no caso simétrico). Unha maneira de corrixir ese feito é facendo a seguinte definición.

Definición 5.6. Sexa $f \in T_p(E)$ e $g \in T_q(E)$. O *produto exterior* de f e g é o tensor antisimétrico

$$f \wedge g = \frac{(p+q)!}{p!q!}A(f\otimes g) \in A_{p+q}(E).$$

Alternativamente, podemos escribir

$$f\wedge g = \frac{1}{p!q!}\sum_{\sigma\in S_{p+q}} \operatorname{sgn}(\sigma)\sigma(f\otimes g).$$

O produto exterior cumpre determinadas propiedades que fan que a operación resulte útil para traballar cos tensores antisimétricos; en particular, é asociativa. Para establecer ese feito, usaremos o seguinte resultado previo.

Lema 5.1. Se $f \in T_p(E)$ e $g\in T_q(E)$, entón

$$A(A(f)\otimes g) = A(f\otimes g) = A(f\otimes A(g)).$$

Demostración. Podemos ver $S_p \subset S_{p+q}$ como unha permutación que fixa cada elemento do conxunto $\{p+1, \dots, p+q\}$. En particular, $(\sigma \cdot f) \otimes g = \sigma(f \otimes g)$. Usando a bilinealidade do produto tensorial, a linealidade do operador de antisimetrización e que $A(\sigma \cdot f) = \mathrm{sgn}(\sigma) A(f)$, obtemos a seguinte cadea de igualdades:

$$\begin{aligned}
A(A(f) \otimes g) &= A\Big(\Big(\frac{1}{p!}\sum_{\sigma \in S_p} \mathrm{sgn}(\sigma)\sigma \cdot f\Big) \otimes g\Big) \\
&= A\Big(\frac{1}{p!}\sum_{\sigma \in S_p} \mathrm{sgn}(\sigma)((\sigma \cdot f) \otimes g)\Big) \\
&= \frac{1}{p!}\sum_{\sigma \in S_p} \mathrm{sgn}(\sigma) A((\sigma \cdot f) \otimes g) \\
&= \frac{1}{p!}\sum_{\sigma \in S_p} \mathrm{sgn}(\sigma) A(\sigma \cdot (f \otimes g)) \\
&= \frac{1}{p!}\sum_{\sigma \in S_p} \mathrm{sgn}(\sigma)^2 A(f \otimes g) \\
&= \frac{1}{p!}\sum_{\sigma \in S_p} A(f \otimes g) = A(f \otimes g).
\end{aligned}$$

Analogamente tense que $A(f \otimes A(g)) = A(f \otimes g)$. □

O produto exterior é asociativo e distributivo, aínda que a diferenza do que sucede co produto tensorial, a asociatividade non é completamente obvia e require o uso do lema anterior. Por outro lado, tampouco é conmutativo, aínda que podemos relacionar dun xeito sinxelo $f \wedge g$ e $g \wedge f$.

Proposición 5.6. Sexan $f, f' \in T_p(E)$, $g, g' \in T_q(E)$, $h \in T_r(E)$ e $\lambda \in K$. Entón:

(a) $(f \wedge g) \wedge h = f \wedge (g \wedge h)$, que coincide con $\frac{(p+q+r)!}{p!q!r!} A(f \otimes g \otimes h)$.

(b) $f \wedge (g + g') = f \wedge g + f \wedge g'$ e $(f + f') \wedge g = f \wedge g + f' \wedge g$.

(c) $(\lambda f) \wedge g = \lambda(f \wedge g) = f \wedge (\lambda g)$.

(d) $g \wedge f = (-1)^{pq} f \wedge g$.

Demostración. Imos comezar demostrando a asociatividade empregando o lema anterior.

$$\begin{aligned}
(f \wedge g) \wedge h &= \frac{(p+q+r)!}{(p+q)!r!} A((f \wedge g) \otimes h) \\
&= \frac{(p+q+r)!}{(p+q)!r!} A\Big(\frac{(p+q)!}{p!q!} A(f \otimes g) \otimes h\Big) \\
&= \frac{(p+q+r)!}{(p+q)!r!} \frac{(p+q)!}{p!q!} A(A(f \otimes g) \otimes h) \\
&= \frac{(p+q+r)!}{p!q!r!} A((f \otimes g) \otimes h),
\end{aligned}$$

onde na última igualdade empregamos o lema anterior. De xeito análogo demóstrase que

$$f \wedge (g \wedge h) = \frac{(p+q+r)!}{p!q!r!} A(f \otimes A(g \otimes h)) = \frac{(p+q+r)!}{p!q!r!} A(f \otimes (g \otimes h)),$$

polo que o resultado é consecuencia da asociatividade do produto tensorial.
As propiedades (b) e (c) son consecuencia directa da linealidade do produto tensorial e do operador de antisimetrización.
Por último, para demostrar (d), consideramos a permutación $\tau \in S_{p+q}$ definida por

$$\tau = \begin{pmatrix} 1 & \dots & p & p+1 & \dots & p+q \\ q+1 & \dots & q+p & 1 & \dots & q \end{pmatrix} = (p+q, p+q-1, \dots, 3, 2, 1)^p.$$

Por ser a potencia p-ésima dun $(p+q)$-ciclo, o seu signo é $\operatorname{sgn}(\tau) = (-1)^{(p+q-1)p} = (-1)^{pq}$, xa que $(p-1)p$ sempre é par. Por outra banda, $g \otimes f = \tau(f \otimes g)$, xa que

$$\begin{aligned} (g \otimes f)(x_1, \dots, x_{p+q}) &= g(x_1, \dots, x_q) f(x_{q+1}, \dots, x_{q+p}) \\ &= f(x_{\tau(1)}, \dots, x_{\tau(p)}) g(x_{\tau(p+1)}, \dots, x_{\tau(p+q)}) \\ &= \tau(f \otimes g)(x_1, \dots, x_{p+q}). \end{aligned}$$

Polo tanto,

$$\begin{aligned} g \wedge f &= \frac{(p+q)!}{p!q!} A(g \otimes f) \\ &= \frac{(p+q)!}{p!q!} A(\tau(f \otimes g)) \\ &= \operatorname{sgn}(\tau) \frac{(p+q)!}{p!q!} A(f \otimes g) \\ &= (-1)^{pq} f \wedge g. \end{aligned}$$

□

A seguinte proposición resume algunhas propiedades do produto exterior que tamén nos resultarán útiles. Ademais, enfatiza a conexión coa teoría dos determinantes.

Proposición 5.7. Sexan $u_1, \dots, u_p \in E^*$. Cúmprense entón as seguintes propiedades.

(a) $u_1 \wedge \dots \wedge u_p = p! A(u_1 \otimes \dots \otimes u_p) = \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) u_{\sigma(1)} \otimes \dots \otimes u_{\sigma(p)}$.

(b) Se $\sigma \in S_p$ e $\tau = \sigma^{-1}$, entón $\sigma(u_1 \wedge \dots \wedge u_p) = u_{\tau(1)} \wedge \dots \wedge u_{\tau(p)} = \operatorname{sgn}(\sigma) u_1 \wedge \dots \wedge u_p$.

(c) Se $u_i = u_j$ para algún $i < j$, entón $u_1 \wedge \dots \wedge u_p = 0$.

(d) Se $x_1, \dots, x_p \in E$, entón $(u_1 \wedge \dots \wedge u_p)(x_1, \dots, x_p) = \det(u_i(x_j))$.

(e) $u_1 \wedge \dots \wedge u_p \neq 0$ se, e soamente se, $u_1, \dots, u_p$ son linealmente independentes.

(f) Sexa $\mathcal{B}_e$ unha base de E e $\mathcal{B}_e^*$ a súa base dual. Sexan $1 \leq i_1 < \dots < i_p \leq n$ e $1 \leq j_1 < \dots < j_p \leq n$. Entón,

$$(v_{i_1}^* \wedge \dots \wedge v_{i_p}^*)(v_{j_1}, \dots, v_{j_p}) = \begin{cases} 1 & \text{se } (i_1, \dots, i_p) = (j_1, \dots, j_p), \\ 0 & \text{se } (i_1, \dots, i_p) \neq (j_1, \dots, j_p). \end{cases}$$

Demostración. (a) A primeira parte é consecuencia da asociatividade e da seguinte observación:

$$A(u_1 \otimes \dots \otimes u_p) = \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)(u_{\sigma(1)} \otimes \dots \otimes u_{\sigma(p)}).$$

Para demostrar isto, simplemente observamos que

$$\begin{aligned}
A(u_1 \otimes \cdots \otimes u_p) &= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)\sigma(u_1 \otimes \cdots \otimes u_p) \\
&= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)(u_{\sigma^{-1}(1)} \otimes \cdots \otimes u_{\sigma^{-1}(p)}) \\
&= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma^{-1})(u_{\sigma^{-1}(1)} \otimes \cdots \otimes u_{\sigma^{-1}(p)}) \\
&= \frac{1}{p!} \sum_{\tau \in S_p} \operatorname{sgn}(\tau)(u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)}).
\end{aligned}$$

(b) Para a segunda parte, como $u_1 \wedge \cdots \wedge u_p$ é antisimétrico, temos que $\sigma(u_1 \wedge \cdots \wedge u_p) = \operatorname{sgn}(\sigma)(u_1 \wedge \cdots \wedge u_p)$. Usando o apartado (a), sabemos que $\sigma \cdot A(f) = A(\sigma \cdot f)$ e como $\sigma(u_1 \otimes \cdots \otimes u_p) = u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)}$, entón

$$\sigma(u_1 \wedge \cdots \wedge u_p) = p!\sigma A(u_1 \otimes \cdots \otimes u_p) = p!A(u_{\tau(1)} \otimes \cdots \otimes u_{\tau(p)}) = u_{\tau(1)} \wedge \cdots \wedge u_{\tau(p)}.$$

(c) Supoñamos que $u_i = u_j$, con $i < j$, e collamos $\sigma = (i,j)$, $\tau = \sigma^{-1}$. Polo apartado (b), $u_{\tau(1)} \wedge \cdots \wedge u_{\tau(p)} = \operatorname{sgn}(\sigma) u_1 \wedge \cdots u_p$. Como $\tau = (i,j)$, ao aplicar τ o produto exterior quédase igual, e como $\operatorname{sgn}(\sigma) = -1$, entón $u_1 \wedge \cdots \wedge u_p = 0$.

(d) Observamos que

$$\begin{aligned}
(u_1 \wedge \cdots \wedge u_p)(x_1, \ldots, x_p) &= \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma)(u_{\sigma(1)} \otimes \cdots \otimes u_{\sigma(p)})(x_1, \ldots, x_p) \\
&= \sum_{\sigma \in S_p} \operatorname{sgn}(\sigma) u_{\sigma(1)}(x_1) \cdots u_{\sigma(p)}(x_p) \\
&= \det(u_i(x_j)).
\end{aligned}$$

(e) Procedemos igual que no caso dos determinantes. Se $u_1, \ldots, u_p$ son linealmente independentes, completamos ata ter unha base $u_1, \ldots, u_n$ de E^*. Esta base é a dual dunha base de E, á que chamamos $v = v_1, \ldots, v_n$. Entón $(u_1 \wedge \cdots \wedge u_n)(v_1, \ldots, v_n) = 1$. Reciprocamente, supoñamos que $u_p = \sum_{j=1}^{p-1} \lambda_j u_j$. Entón, $u_1 \wedge \cdots \wedge u_p = 0$.

(f) Usando o apartado (d), temos que $(v_{i_1}^* \wedge \cdots \wedge v_{i_p}^*)(v_{j_1}, \ldots, v_{j_p}) = \det(v_{i_k}^*(v_{j_\ell}))$. Se $(i_1, \ldots, i_p) = (j_1, \ldots, j_p)$, entón o valor do determinante é 1. Senón, collemos o primeiro enteiro $1 \leq q \leq p$ tal que $i_q \neq j_q$. Se $i_q < j_q$ temos que

$$1 \leq i_1 = j_1 < \ldots < i_{q-1} = j_{q-1} < i_q < j_q < j_{q+1} < \ldots < j_p \leq n.$$

Polo tanto, $i_q \notin \{j_1, \ldots, j_p\}$ e $v_{i_q}^*(v_{j_1}) = \ldots = v_{i_q}^*(v_{j_p}) = 0$. Entón, unha fila da matriz $(v_{i_k}^*(v_{j_\ell}))$ é 0 polo que o determinante é cero. O caso no que $i_q > j_q$ é análogo: nese caso,

$$1 \leq j_1 = i_1 < \ldots < j_{q-1} = i_{q-1} < j_q < i_q < i_{q+1} < \ldots < i_p \leq n,$$

con $j_q \notin \{i_1, \ldots, i_p\}$, $v_{i_1}^*(v_{j_q}) = \ldots = v_{i_p}^*(v_{j_q}) = 0$. Por conseguinte, unha columna da matriz é cero e o determinante é 0.

□

A seguinte proposición usa as propiedades anteriores para establecer a dimensión de $A_p(E)$. Ao estudarmos os determinantes, xa demostramos que se $p = n$, entón $A_p(E)$ ten dimensión 1. Nese caso, a proba baseábase en dous feitos: (a) establecer que ao considerar produtos da forma $v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*$ nos que dous índices son iguais o resultado é cero, e máis en xeral, que se os índices son os mesmos pero noutra orde existe unha relación entre eles; (b) ver que a única combinación lineal de elementos da forma $v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*$ permitida por (a) é de feito un tensor antisimétrico. Imos xeralizar ese argumento para valores arbitrarios de n e p.

Proposición 5.8. Cúmprese que $\{v_{i_1}^* \wedge \cdots \wedge v_{i_p}^*\}_{1 \leq i_1 < \ldots < i_p \leq n}$ é unha base de $A_p(E)$. En particular, $\dim A_p(E) = \binom{n}{p}$.

Demostración. En primeiro lugar, observamos que os tensores da base proposta son xeradores: se hai dous índices iguais entón sabemos que o elemento é cero, e aplicando transposicións podemos conseguir que os índices estean ordenados en orde crecente (é dicir, o antisimetrizado de calquera elemento da base coincide, salvo escalar, con algún elemento do conxunto).
Por outra banda, se consideramos unha combinación lineal da forma

$$\sum_{1 \leq i_1 < \ldots < i_p \leq n} \lambda_{i_1,\ldots,i_p} v_{i_1}^* \wedge \cdots \wedge v_{i_p}^* = 0$$

e avaliamos en $(v_{i_1}, \ldots, v_{i_p})$, entón, aplicando o último epígrafe da proposición anterior, quédanos que $\lambda_{i_1,\ldots,i_p} = 0$. Polo tanto, o conxunto é tamén linealmente independente. $\square$

Exemplo. Sexa E un espazo vectorial de dimensión 3 con base $\{v_1, v_2, v_3\}$. Entón, $A_2(E)$ ten dimensión 3, e unha base está dada por $\{v_1 \wedge v_2, v_2 \wedge v_3, v_3 \wedge v_1\}$. Máis en xeral, se E ten dimensión n, $A_2(E)$ ten dimensión $\binom{n}{2}$, que coincide á súa vez coa dimensión do subespazo de matrices antisimétricas.

Unha xeralización interesante consiste en achar a dimensión do espazo dos tensores simétricos.

Proposición 5.9. Cúmprese que $\{S(v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*)\}_{1 \leq i_1 \leq \ldots \leq i_p \leq n}$ é unha base de $S_p(E)$. En particular, $\dim S_p(E) = \binom{n+p-1}{p}$.

Demostración. Aplicando unha permutación de S_p podemos conseguir que os índices estean ordenados en orde non decrecente, polo que o simetrizado de calquera elemento da base é igual a algún dos do conxunto proposto. Iso é suficiente para establecer que son xeradores, xa que o simetrizado de calquera elemento de $T_p(E)$ pódese escribir como combinación lineal destes.
Procedendo como na demostración anterior e avaliando no simetrizado de $v_{i_1} \otimes \cdots \otimes v_{i_p}$ chegamos a que o conxunto tamén é linealmente independente.
O resultado sobre a dimensión séguese de observar que un elemento da base o podemos identificar cunha n-tupla $(a_1, \ldots, a_n)$, onde a_i indica o número de elementos que son iguais a v_i^*. Empregando combinatoria elemental, sabemos que o número de solucións da ecuación

$$a_1 + \ldots + a_n = p, \text{ con } a_i \geq 0,$$

é $\binom{n+p-1}{n-1} = \binom{n+p-1}{p}$. $\square$

Exemplo. Sexa E un espazo vectorial de dimensión 3 con base $\{v_1, v_2, v_3\}$. Entón, $S_2(E)$ ten dimensión 6, e unha base está dada por

$$\{v_1 \otimes v_1, (v_1 \otimes v_2 + v_2 \otimes v_1)/2, (v_1 \otimes v_3 + v_3 \otimes v_1)/2, v_2 \otimes v_2, (v_2 \otimes v_3 + v_3 \otimes v_2)/2, v_3 \otimes v_3\}.$$

Máis en xeral, se E ten dimensión n, $S_2(E)$ ten dimensión $\binom{n+1}{2}$, que coincide á súa vez coa dimensión do subespazo de matrices simétricas.

Se $p = 2$, $A_2(E)$ e $S_2(E)$ están en suma directa, xa que a súa intersección é trivial e

$$\dim A_2(E) + \dim S_2(E) = \binom{n+1}{2} + \binom{n}{2} = \frac{(n+1)n + n(n-1)}{2} = n^2 = \dim T_2(E).$$

En cambio, isto non é certo se $p > 2$. Por exemplo, se E ten dimensión 3, $\dim T_3(E) = 27$, $\dim S_3(E) = 10$ e $\dim A_3(E) = 1$.
Imos introducir agora a noción de tensor descompoñible.

Definición 5.7. Un tensor $f \in A_p(E)$ dise que é *descompoñible* cando existen vector $u_1, \dots, u_p \in E^*$ de maneira que

$$f = u_1 \wedge \cdots \wedge u_p.$$

De xeito análogo, diremos tamén que $g \in A^q(E)$ é descompoñible cando existen vector $v_1, \dots, v_q \in E$ de maneira que

$$g = v_1 \wedge \cdots \wedge v_q.$$

Os seguintes resultados preséntanse no caso covariante, pero a súa adaptación ao caso contravariante é totalmente inmediata.

Proposición 5.10. Un elemento $f \in A_2(E)$ é descompoñible se, e soamente se, $f \wedge f = 0$ en $A_4(E)$.

Demostración. Supoñamos que $f = u \wedge v$ con $u, v \in E^*$. Entón,

$$f \wedge f = u \wedge v \wedge u \wedge v = -(u \wedge u) \wedge v \wedge v = 0.$$

Consideremos agora que $f \wedge f = 0$. Se a dimensión de E é igual a 2, o resultado é trivial xa que $A_2(E)$ é de dimensión 1 xerado por unha base $u_1 \wedge u_2$, onde $u = (u_1, u_2)$ é unha base de E^*. Se a dimensión de E é 3, consideramos o morfismo $E^* \to A_3(E)$ que envía v a $v \wedge f$. Como estamos en dimensión 3, podemos considerar unha base do núcleo, v_1, v_2, e estendela a unha base (v_1, v_2, v_3) de E^*. Entón

$$f = \lambda_1 v_2 \wedge v_3 + \lambda_2 v_3 \wedge v_1 + \lambda_3 v_1 \wedge v_2.$$

Temos agora que, por definición, $0 = v_1 \wedge f = \lambda_1 v_1 \wedge v_2 \wedge v_3$, polo que $\lambda_1 = 0$; do mesmo xeito, $\lambda_2 = 0$. Polo tanto, $f = \lambda_3 v_1 \wedge v_2$, que é descompoñible (observamos que de momento aínda non se usou que $f \wedge f = 0$).
Imos facer agora o caso xeral por indución, supondo o resultado certo cando a dimensión de E é como máximo $n - 1$, e considerar o caso $\dim E = n$. Collemos unha base $(v_1, \dots, v_n)$ de E^* e escribimos

$$f = \sum_{1 \le i < j \le n} a_{ij} v_i \wedge v_j = \Big(\sum_{i=1}^{n-1} a_{in} v_i \Big) \wedge v_n + \sum_{1 \le i < j}^{n-1} a_{ij} v_i \wedge v_j = u \wedge v_n + f',$$

onde se $U = \langle v_1, \dots, v_{n-1} \rangle$, temos que $u \in U$ e $f' \in A_2(U)$. Como f cumpre que $f \wedge f = 0$, temos que

$$0 = f \wedge f = (u \wedge v_n + f') \wedge (u \wedge v_n + f') = 2u \wedge f' \wedge v_n + f' \wedge f'.$$

Como v_n non aparece na expansión nin de $u \wedge f'$ nin na de $f' \wedge f'$, necesitamos simultaneamente que

$$u \wedge f' = 0, \quad f' \wedge f' = 0.$$

Pola hipótese de indución, como $f' \wedge f' = 0$ temos que $f' = a \wedge b$, polo que a primeira ecuación queda como $u \wedge a \wedge b = 0$. Iso quere dicir que existe unha relación de dependencia lineal da forma $\lambda u + \mu_1 a + \mu_2 b = 0$. Se $\lambda = 0$, entón a e b son linealmente dependentes, polo que $f' = 0$ e $f = u \wedge v_n$. En caso contrario, $u = (-\mu_1/\lambda)a + (-\mu_2/\lambda)b$, polo que

$$f = (-\mu_1/\lambda)a \wedge v_n + (-\mu_2/\lambda)b \wedge v_n + a \wedge b.$$

Este é precisamente o caso de dimensión 3, no que vimos que se cumpría a propiedade. Isto conclúe a demostración e establece que f sempre é descompoñible. $\square$

O caso de dimensión 3 pódese facer de xeito diferente, xustificando que calquera tensor da forma

$$f = a_{12} v_1 \wedge v_2 + a_{13} v_1 \wedge v_3 + a_{23} v_2 \wedge v_3$$

é descompoñible. Se $f \neq 0$, entón algún dos coeficientes é non nulo; sen perder xeneralidade, supoñamos que $a_{12} \neq 0$. Nese caso,

$$f = a_{12} \left(v_1 + \frac{a_{23}}{a_{12}} v_3 \right) \wedge \left(v_2 + \frac{a_{13}}{a_{12}} v_3 \right).$$

Este resultado admite a seguinte versión máis xeral.

Proposición 5.11. Se E ten dimensión n, todo tensor de $A_{n-1}(E)$ é descompoñible.

Demostración. Sexa $x \in A_{n-1}(E)$. Escribimos

$$x = \lambda_1 \hat{v}_1 \wedge v_2 \wedge \cdots \wedge v_n + \lambda_2 v_1 \wedge \hat{v}_2 \wedge \cdots \wedge v_n + \lambda_n v_1 \wedge v_2 \wedge \cdots \wedge \hat{v}_n.$$

Se $x = 0$ non hai nada que demostrar; en caso contrario, un dos λ_i é diferente de 0. Sen perder xeralidade, supoñemos que $\lambda_n \neq 0$. Polo tanto, podemos atopar solucións do sistema

$$x = \lambda_n (v_1 + \mu_1 v_n) \wedge (v_2 + \mu_2 v_n) \wedge \cdots \wedge (v_{n-1} + \mu_{n-1} v_n).$$

Ao desenvolver o produto exterior da dereita, unicamente nos queda o termo $\lambda_n v_1 \wedge v_2 \wedge \cdots \wedge \hat{v}_n$ (sen v_n) e os termos cun único v_n. En particular, temos $n-1$ ecuacións da forma

$$\lambda_i v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_n = (-1)^{n-1-i} \mu_i v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_n,$$

onde o signo depende do número de transposicións. Polo tanto, $\mu_i = (-1)^{n-1-i} \lambda_i$ e temos a descomposición desexada. $\square$

Exemplo. En $\mathbb{R}^4$, consideramos o elemento de $A_3(\mathbb{R}^4)$ dado por

$$f = v_1^* \wedge v_2^* \wedge v_3^* + 3v_1^* \wedge v_2^* \wedge v_4^* - 2v_1^* \wedge v_3^* \wedge v_4^* + v_2^* \wedge v_3^* \wedge v_4^*.$$

Como o coeficiente con $v_1^* \wedge v_2^* \wedge v_3^*$, consideramos unha descomposición da forma

$$f = \lambda(v_1^* + \mu_1 v_4^*) \wedge (v_2^* + \mu_2 v_4^*) \wedge (v_3^* + \mu_3 v_4^*).$$

Igualando termos, vemos que $\lambda = 1$, $\mu_1 = 1$, $\mu_2 = 2$ e $\mu_3 = 3$.

O estudo dos tensores descompoñibles ten importancia tamén no estudo da xeometría, e permite introducir un obxecto de gran relevancia, a *grassmaniana*. Porén, determinar se un tensor en $A_k(E)$ é descompoñible, cando $k > 2$, é un problema polo xeral máis complexo.

5.4. A álxebra tensorial

Os espazos vectoriais, polo xeral, non teñen unha estrutura de multiplicación. Porén, xa atopamos diferentes situacións nos que iso si sucedía; o exemplo máis representativo vímolo ao traballarmos os endomorfismos dun espazo vectorial. O caso dos tensores dá outro exemplo interesante. Na seguinte definición, identificamos $E^{\otimes k}$ con $T^k(E)$, xa que calquera k-tupla de elementos de E pode entenderse como unha aplicación lineal de k copias de E^* en K. Por convenio, entendemos que $E^{\otimes 0} = K$.

Definición 5.8. O K-espazo vectorial $T(E) = \oplus_{k \geq 0} E^{\otimes k}$ chámase a *K-álxebra tensorial* ou a *K-álxebra dos tensores*, co produto dado polo produto tensorial.

A definición anterior é correcta porque o produto tensorial é bilineal nas diferentes compoñentes, $1 \in E$ é a identidade para o produto e ademais a multiplicación é asociativa. Como espazo vectorial, $T(E)$ ten dimensión infinita, xa que a compoñente $E^{\otimes k}$ ten dimensión n^k.
Outros caso a ter en conta son as álxebras dos tensores simétricos e a dos alternados, neste último caso co produto dado polo produto exterior.

Definición 5.9. Sexa $C(E)$ a subálxebra de $T(E)$ xerada por todos os elementos da forma $u \otimes v - v \otimes u$, onde $u, v \in T(E)$. A *K-álxebra dos tensores simétricos*, $S(E)$, defínese como $T(E)/C(E)$.

Como espazo vectorial, $S(E)$ é isomorfo a $\oplus_{k \geq 0} S_k(E)$, pero non hai un isomorfismo de álxebras xa que ao igual que ocorre no caso antisimétrico, o produto tensorial de tensores simétricos non ten por que ser simétrico.

Definición 5.10. O K-espazo vectorial $A(E) = \oplus_{k \geq 0} A^k(E)$ é a K-álxebra dada pola suma directa

$$A(E) = \bigoplus_{k \geq 0} A^k(E) = K \oplus E \oplus A^2(E) \oplus \cdots,$$

coa multiplicación dada polo produto exterior. Adoita chamarse *K-álxebra exterior de E*.

Como espazo vectorial $A(E)$ ten dimensión 2^n xa que

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n.$$

(Esta igualdade é consecuencia directa do binomio de Newton).
A seguinte proposición amosa a relación entre dous conceptos que aparecerán ao estudarmos a teoría de integración e ademais é un primeiro exemplo de como se relacionan elementos de diferentes graos dentro da álxebra, caracterizando as solucións da ecuación $v \wedge \omega = 0$, onde $v \in E$ e $\omega \in A^k(E)$. En particular, afirma que fixado un vector $v \neq 0$, a condición de que o produto exterior cunha forma ω sexa 0 se pode interpretar en termos da existencia dun elemento η tal que $\omega = v \wedge \eta$.

Proposición 5.12. Sexa $v \in E$ con $v \neq 0$, e sexa $\omega \in A^k(E)$. Entón, $v \wedge \omega = 0$ se, e soamente se, $\omega = v \wedge \eta$ para algún $\eta \in A^{k-1}(V)$. En particular, se $\omega \in A(E)$, $v \wedge \omega = 0$ se, e soamente se, $\omega = v \wedge \eta$ para algún $\eta \in A(E)$.

Demostración. Usando a asociatividade, se $\omega = v \wedge \eta$, entón $v \wedge \omega = v \wedge (v \wedge \eta) = (v \wedge v) \wedge \eta = 0$. Imos agora demostrar o recíproco, é dicir, que se $v \wedge \omega = 0$, entón $\omega = v \wedge \eta$ para algún η. Se $k > n$, entón $\omega = 0$ e non hai nada que demostrar. Supoñamos entón que $1 \leq k \leq n$. Podemos estender v a unha base de todo o espazo E, $v_1, \ldots, v_n$ con $v = v_1$. Se $k = n$ a condición $v \wedge \omega = 0$ é automática; tamén no caso $k = n$, $A^n(E)$ ten como base $v_1 \wedge \cdots \wedge v_n$, polo que $\omega = c(v_1 \wedge \cdots \wedge v_n)$, para algún escalar c. Entón $\omega = v \wedge \eta$, con $\eta = cv_2 \wedge \cdots \wedge v_n$.
Supoñamos entón que $1 \leq k \leq n-1$, polo que $n \geq 2$. Usando a base anterior, podemos pór

$$\omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} c_{i_1,\ldots,i_k} v_{i_1} \wedge \cdots \wedge v_{i_k}.$$

Como $v_1 = v$, $v \wedge v_1 = 0$, polo que

$$v \wedge \omega = \sum_{2 \leq i_1 < \ldots < i_k \leq n} c_{i_1,\ldots,i_k} v_{i_1} \wedge \cdots \wedge v_{i_k}.$$

A parte da dereita é parte dunha base de $A^{k+1}(E)$, polo que de $v \wedge \omega = 0$ todos os coeficientes se anulan. Entón,

$$\omega = v \wedge \sum_{1 < i_2 < \ldots < i_k \leq n} c_{1,i_2,\ldots,i_k} v_{i_2} \wedge \cdots \wedge v_{i_k}.$$

Polo tanto, temos unha descomposición da forma $\omega = v \wedge \eta$.
Para o caso xeral, pomos $\omega = \sum_{k=0}^{n} \omega_k$, con $\omega_k \in A^k(E)$ é a parte de grao k de ω. Entón,

$$v \wedge \omega = \sum_{k=0}^{n} v \wedge \omega_k.$$

Como $v \wedge \omega_k \in A^{k+1}(E)$, os diferentes termos da suma están en diferentes partes de $A(E)$. Polo tanto, $v \wedge \omega_k = 0$ para todo k. Pola primeira parte, $\omega_0 = 0$ e para $k \geq 1$ temos $\omega_k = v \wedge \eta_{k-1}$, para algún $\eta_{k-1} \in A^{k-1}(E)$. Entón, $\omega = v \wedge \eta$, onde $\eta = \sum_{k=1}^{n} \eta_{k-1} \in A(E)$. □

A álxebra exterior tamén se pode caracterizar en termos dunha propiedade universal. Se A é unha K-álxebra de maneira que existe un endomorfismo $L \colon E \to A$ de maneira que $L(v)^2 = 0$ para todo $v \in E$. Entón, existe unha única extensión de L a unha aplicación de A-álxebras $\tilde{L} \colon A(E) \to A$.

5.5. Produto tensorial de espazos vectoriais

Antes de definir o produto tensorial de espazos vectoriais, imos introducir a noción de *produto de Kronecker de dúas matrices.*

Definición 5.11. Sexan $M = (a_{i,j})$ e $N = (b_{k,\ell})$ dúas matrices $m \times n$ e $p \times q$ respectivamente. O *produto tensorial de M por N* (ou *produto de Kronecker*), que denotamos $M \otimes N$, é a matriz $mp \times nq$ formada por bloques $m \times n$ de maneira que o bloque (i, j) é $a_{i,j}N$:

$$M \otimes N = \begin{pmatrix} a_{1,1}N & \dots & a_{1,n}N \\ \vdots & & \vdots \\ a_{m,1}N & \dots & a_{m,n}N. \end{pmatrix}$$

Unha propiedade que resultará útil, especialmente nos problemas, é entender a relación entre os valores propios de dúas matrices e dos seus produtos tensoriais.

Proposición 5.13. Sexan $A \in \mathcal{M}_m(K)$ e $B \in \mathcal{M}_n(K)$ dúas matrices con valores propios λ e μ en K. Entón, $A \otimes \mathbb{I}_n + \mathbb{I}_m \otimes B$ ten valor propio $\lambda + \mu$ e $A \otimes B$ ten valor propio $\lambda\mu$.

Demostración. Temos que $Av = \lambda v$ e $Bw = \mu w$ para vectores $v \in K^m$ e $w \in K^n$ non nulos. Entón

$$\begin{aligned} (A \otimes \mathbb{I}_n + \mathbb{I}_m \otimes B)(v \otimes w) &= Av \otimes w + v \otimes Bw \\ &= \lambda v \otimes w + v \otimes \mu w \\ &= (\lambda + \mu)(v \otimes w). \end{aligned}$$

De xeito similar,

$$(A \otimes B)(v \otimes w) = Av \otimes Bw = \lambda v \otimes \mu w = \lambda\mu(v \otimes w).$$

□

De cara a motivar as aplicacións do produto de Kronecker, imos dar unha definición provisional do produto tensorial de dous espazos vectoriais E e F, antes de formalizar o concepto mediante unha definición máis abstracta.

Definición 5.12. Sexan $\mathcal{B}_v = \{v_1, \dots, v_n\}$ e $\mathcal{B}'_v = \{v'_1, \dots, v'_m\}$ bases de E e G, respectivamente. Definimos $E \otimes G$ como o espazo vectorial de dimensión nm xerado polos elementos $v_i \otimes v'_j$; é dicir

$$E \otimes G = \Big\{ \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_{i,j} v_i \otimes v'_j \Big\}.$$

En xeral, se $u = \sum_{i=1}^{n} \lambda_i v_i$ e $v = \sum_{j=1}^{m} \mu_j v'_j$, definimos

$$u \otimes v = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \mu_j v_i \otimes v'_j.$$

O produto tensorial de $f : E \to F$ e $g : G \to H$ defínese como a aplicación $f \otimes g : E \otimes G \to F \otimes H$ que sobre unha base de $E \otimes G$ toma os valores

$$(f \otimes g)(v_i \otimes v'_j) = f(v_i) \otimes g(v'_j),$$

e que logo se estende por linealidade.

É un exercicio doado ver que a definición que demos de $f \otimes g$ non depende da base escollida. A seguinte proposición establece que o produto de Kronecker é precisamente a matriz asociada ao produto tensorial de dúas aplicacións lineais nas bases correspondentes.

Proposición 5.14. Sexan E, F, G e H espazos vectoriais de dimensión finita, e sexan u, v, w e t bases de E, F, G e H, respectivamente. Sexan $f\colon E \to F$ e $g\colon G \to H$ dúas aplicacións lineais. Sexa A a matriz de f nas bases u e v, e B a matriz de g nas bases w e t. Entón, $f \otimes g\colon E \otimes G \to F \otimes H$ é unha aplicación lineal cuxa matriz nas bases $u \otimes w$ e $v \otimes t$ é $A \otimes B$.

Demostración. Trátase de efectuar un cálculo rutineiro. Sexan m, n, p e q as dimensións dos espazos vectoriais E, F, G e H, respectivamente. Entón,

$$\begin{aligned}(f \otimes g)(u_i \otimes w_k) &= f(u_i) \otimes g(w_k) \\ &= \Big(\sum_{j=1}^{n} a_{ji} v_j\Big) \otimes \Big(\sum_{\ell=1}^{q} b_{\ell k} t_\ell\Big) \\ &= \sum_{j=1}^{n}\sum_{\ell=1}^{q} a_{ji} b_{\ell k} v_j \otimes t_\ell.\end{aligned}$$

Polo tanto, na entrada $(jq - q + \ell, ip - p + k) = (\ell + q(j-1), k + p(i-1))$ da matriz correspondente a $f \otimes g$ temos $a_{ji} b_{\ell k}$, que é precisamente a entrada correspondente en $A \otimes B$. $\square$

Imos traballar un exemplo para ver a intuición detrás da proposición anterior.

Exemplo. Consideramos os endomorfismos de K^2 que na base canónica se representan coas matrices de $M_2(K)$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},\ A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

Entón, $f \otimes g$ é un endomorfismo de $K^2 \otimes K^2$, que representamos por $A \otimes A'$. Escribindo $\{v_1, v_2\}$ para a base canónica de K^2, podemos calcular a matriz de $A \otimes A'$ con respecto á base $\{v_1 \otimes v_1, v_1 \otimes v_2, v_2 \otimes v_1, v_2 \otimes v_2\}$. Temos entón que

$$\begin{aligned}(A \otimes A')(v_1 \otimes v_1) &= Av_1 \otimes A'v_1 \\ &= (av_1 + cv_2) \otimes (a'v_1 + c'v_2) \\ &= aa'v_1 \otimes v_1 + ac'v_1 \otimes v_2 + ca'v_2 \otimes v_1 + cc'v_2 \otimes v_2,\end{aligned}$$

e un cálculo análogo dá resultados análogos para $(A \otimes A')(v_1 \otimes v_2)$, $(A \otimes A')(v_2 \otimes v_1)$ e $(A \otimes A')(v_2 \otimes v_2)$. Polo tanto, a matriz de $A \otimes A'$ é o produto de Kronecker

$$\begin{pmatrix} aA' & bA' \\ cA' & dA' \end{pmatrix}.$$

Por último, observamos que o produto de Kronecker tamén o podemos empregar para os cambios de base. A seguinte proposición expresa a matriz de cambio de base no contexto máis xeral posible.

Proposición 5.15. Sexan $\mathcal{B}_v = \{v_1, \dots, v_n\}$ e $\mathcal{B}_u = \{u_1, \dots, u_n\}$ dúas bases de E e sexan $\mathcal{B}_v^* = \{v_1^*, \dots, v_n^*\}$ e $\mathcal{B}_u^* = \{u_1^*, \dots, u_n^*\}$ as súas bases duais. Sexan $A = M_{u,e}$ e $B = M_{u^*,e^*} = (A^t)^{-1}$ as matrices de cambio de base. Sexan $\mathcal{B}_p^q$ e $\bar{\mathcal{B}}_p^q$ as bases de $T_p^q(E)$ asociadas a u e a e, respectivamente. A matriz C de cambio de base de $\bar{\mathcal{B}}_p^q$ a $\mathcal{B}_p^q$ é

$$C = B^{\otimes p} \otimes A^{\otimes q}.$$

Demostración. A proposición é un caso particular do resultado sobre o produto de Kronecker cando collemos as matrices correspondentes á identidade nas bases e e u ou e^* e u^*. □

Exemplo. Consideremos o caso dun espazo vectorial de dimensión 2, con base $\mathcal{B}_v = \{v_1, v_2\}$ e base dual $\mathcal{B}_v^* = \{v_1^*, v_2^*\}$. Consideramos outra base $\mathcal{B}_u = \{u_1, u_2\}$ e a correspondente base dual $\mathcal{B}_u^* = \{u_1^*, u_2^*\}$. Se A é a matriz de cambio de base de $\mathcal{B}_u$ a $\mathcal{B}_v$ e B é a matriz de cambio de base de $\mathcal{B}_u^*$ a $\mathcal{B}_v^*$, sabemos que $B = (A^t)^{-1}$. Pomos

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}.$$

En $T_1^1(E)$, a base vén dada por

$$\mathcal{B}_v^* \otimes \mathcal{B}_v := \{v_1^* \otimes v_1, v_1^* \otimes v_2, v_2^* \otimes v_1, v_2^* \otimes v_2,\}$$

e de xeito análogo consideramos unha base $\mathcal{B}_u^* \otimes \mathcal{B}_u$. Imos escribir $u_1^* \otimes u_1$ na base $v_i^* \otimes v_j$. Para iso, pomos $u_1^* \otimes u_1 = \sum_{i,j=1}^2 \lambda_{i,j} v_i^* \otimes v_j$ e avaliamos en $v_1 \otimes v_1^*$. Polo tanto,

$$u_1^* \otimes u_1(v_1, v_1^*) = u_1^*(v_1) u_1(v_1^*) = \lambda_{1,1}.$$

Como $u_1^*(v_1) = \alpha$ e $v_1^*(u_1) = a$, temos que $\lambda_{1,1} = \alpha a$.
Procedendo da mesma maneira, un cálculo rutineiro amosa que

$$\begin{aligned}
u_1^* \otimes u_1 &= (\alpha a) v_1^* \otimes v_1 + (\alpha b) v_1^* \otimes v_2 + (\beta a) v_2^* \otimes v_1 + (\beta b) v_2^* \otimes v_2 \\
u_1^* \otimes u_2 &= (\alpha c) v_1^* \otimes v_1 + (\alpha d) v_1^* \otimes v_2 + (\beta c) v_2^* \otimes v_1 + (\beta d) v_2^* \otimes v_2 \\
u_2^* \otimes u_1 &= (\gamma a) v_1^* \otimes v_1 + (\gamma b) v_1^* \otimes v_2 + (\delta a) v_2^* \otimes v_1 + (\delta b) v_2^* \otimes v_2 \\
u_2^* \otimes u_2 &= (\gamma c) v_1^* \otimes v_1 + (\gamma d) v_1^* \otimes v_2 + (\delta c) v_2^* \otimes v_1 + (\delta d) v_2^* \otimes v_2.
\end{aligned}$$

Polo tanto, a matriz de cambio de base é

$$\begin{pmatrix} \alpha a & \alpha c & \gamma a & \gamma c \\ \alpha b & \alpha d & \gamma b & \gamma d \\ \beta a & \beta c & \delta a & \delta c \\ \beta b & \beta d & \delta b & \delta d \end{pmatrix} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \otimes \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Cando estudamos o produto de Kronecker demos unha definición provisional de produto tensorial, que era práctica á hora de calcular, pero que resulta inapropiada para algúns propósitos. Co obxectivo de presentar eses resultados nun contexto máis axeitado, sexan E e F dous K-espazos vectoriais, e sexan $\{u_1, \dots, u_n\}$ e $\{v_1, \dots, v_m\}$ bases de E e F, respectivamente.
O produto tensorial $E \otimes F$, tal e como se definiu anteriormente, cumpre que

$$(u + u') \otimes v = u \otimes v + u' \otimes v, \quad u \otimes (v + v') = u \otimes v + u \otimes v',$$

e tamén respecta o produto por escalares. Porén, unha diferenza clara coa suma directa é que mentres que calquera elemento de $E \oplus F$ se pode escribir como un par (u, v), con $u \in E$ e $v \in F$, iso non é certo para o produto tensorial, e non podemos pór calquera elemento como $u \otimes v$. En xeral, un elemento de $u \otimes v$ é unha combinación lineal do xeito

$$\sum_{i=1}^{\dim E} \sum_{j=1}^{\dim F} \lambda_{i,j}(u_i \otimes v_j).$$

Por exemplo, mentres que é evidente saber se dous elementos (u, v), $(u', v') \in E \oplus F$ son iguais, para o caso do produto tensorial non o é tanto, senón que cómpre desenvolver a expresión en termos dunha base e ver se todos os coeficientes son iguais.
Nesta sección imos formalizar o concepto de propiedade universal que responde ao seguinte obxectivo. Construír un K-espazo vectorial G de unha aplicación bilineal $\varphi\colon E \times F \to G$ de maneira que calquera aplicación bilineal $E \times F \to H$ nun K-espazo vectorial H se pode escribir como unha composición

$$E \times F \xrightarrow{\varphi} G \xrightarrow{\psi} H,$$

sendo ψ unha aplicación lineal. O matemático Jeremy Kun describiu $E \otimes F$ como o *porteiro* de todas as aplicacións bilineais que saen de $E \times F$. É dicir, é un espazo vectorial que nos permite transformar as aplicacións bilineais de $E \times F$ en aplicacións lineais do novo espazo $E \otimes F$.
Observamos que a construción dun novo obxecto alxébrico a partir dun vello faise polo xeral introducindo a nova estrutura e definindo nela as operacións necesarias. Por exemplo, o espazo cociente E/F consiste de clases $[v]$, e a suma de dúas delas é $[v] + [v'] = [v + v']$. En cambio, ás veces é máis axeitado proceder dun xeito diferente, dicindo como o novo obxecto se relaciona, a través das súas aplicacións, con outros obxectos *similares*. Este acercamento, en termos do que coñecemos como *propiedade universal*, é común non só en álxebra, senón tamén en topoloxía ou en xeometría.
O obxectivo, polo tanto, é construír un par (G, φ), onde G é un K-espazo vectorial e $\varphi\colon E \times F \to G$ é unha aplicación bilineal, de maneira que se cumpre a seguinte propiedade universal.

Definición 5.13. Dise que o par (G, φ) cumpre a *propiedade universal* con respecto a E e F se para calquera (H, ψ), onde H é un espazo vectorial sobre K e $\psi\colon E \times F \to H$ é unha aplicación bilineal, existe unha única aplicación lineal $f\colon G \to H$ tal que $f \circ \varphi = \psi$. É dicir, o seguinte diagrama é conmutativo:

$$\begin{tikzcd} E \times F \arrow[r, "\varphi"] \arrow[dr, "\psi"'] & G \arrow[d, "f"] \\ & H. \end{tikzcd}$$

A seguinte proposición establece que existe un único par que cumpre a propiedade universal. Como é habitual, escribimos $\mathcal{L}(E^*, F^*)$ para o conxunto das aplicacións K-lineais de E a F.

Proposición 5.16. Existe un único par (G, φ), onde $\varphi\colon E \times F \to G$ é bilineal, cumprindo a propiedade universal.

Demostración. Imos comezar demostrando a unicidade. Se (G_1, φ_2) e (G_2, φ_2) cumpren a propiedade universal, existe unha aplicación lineal $f\colon G_1 \to G_2$ de maneira que $f \circ \varphi_1 = \varphi_2$, e outra aplicación lineal $g\colon G_2 \to G_1$ tal que $g \circ \varphi_2 = \varphi_1$. En particular, $(g \circ f) \circ \varphi_1 = g \circ \varphi_2 = \varphi_1$. Como $\mathrm{Id}_{G_1} \circ \varphi_1 = \varphi_1$, pola propiedade universal do par (G_1, φ_1), temos que $g \circ f = \mathrm{Id}_{G_1}$ (xa que na definición de propiedade universal esíxese a existencia dunha única aplicación que cumpra a igualdade). Do mesmo xeito, como $(f \circ g) \circ \varphi_2 = f \circ \varphi_1 = \varphi_2$, pola propiedade universal do par (G_2, ϕ_2) tense que $f \circ g = \mathrm{Id}_{G_2}$. Polo tanto, f e g son isomorfismos que conmutan coa aplicación bilineal:

$$\begin{tikzcd}
 & G_1 \arrow[d, shift right, "f"'] \\
E \times F \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & \\
 & G_2. \arrow[u, shift right, "g"']
\end{tikzcd}$$

Para probar a existencia, consideramos o espazo vectorial

$$\mathcal{L}(E^* \times F^*, K) = \{g\colon E^* \times F^* \to K \text{ con } g \text{ bilineal }\}.$$

Sexa $x \in E$ e $y \in F$. Para definir o elemento $x \otimes y := \varphi(x, y)$, farase o seguinte. A través dos isomorfismos canónicos entre E e F e os seus biduais, podemos definir unha aplicación $x \otimes y\colon E^* \times F^* \to K$, onde se $(\alpha, \beta) \in E^* \times F^*$,

$$(x \otimes y)(\alpha, \beta) = x(\alpha)y(\beta) = \alpha(x)\beta(y).$$

É inmediato ver que $x \otimes y \in \mathcal{L}(E^* \times F^*, K)$ e que $\varphi\colon E \times F \to \mathcal{L}(E^* \times F^*, K)$ definida por $\varphi(x, y) = x \otimes y$ é bilineal. Imos ver agora que

$$u \otimes v := \{u_i \otimes v_j \text{ con } 1 \le i \le n,\ 1 \le j \le m\}$$

é unha base de $\mathcal{L}(E^* \times F^*, K)$. Sexan $\alpha = \sum_{i=1}^n \alpha_i u_i^*$ e $\beta = \sum_{j=1}^m \beta_j v_j^*$. Temos que $(u_h \otimes v_k)(\alpha, \beta) = \alpha(u_h)\beta(v_k) = \alpha_h \beta_k$. Dado $\omega \in \mathcal{L}(E^* \times F^*, K)$ calquera, podemos escribir

$$\omega(\alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \omega(u_i^*, v_j^*) = \sum_{i=1}^n \sum_{j=1}^m \omega(u_i^*, v_j^*) \cdot (u_i \otimes v_j)(\alpha, \beta).$$

Como a igualdade se cumpre para calquera par $(\alpha, \beta) \in E^* \times F^*$, podemos pór $\omega = \sum_{i=1}^n \sum_{j=1}^m \omega(u_i^*, v_j^*) \cdot (u_i \otimes v_j)$. Polo tanto, $u \otimes v$ é un conxunto xerador. Para ver que son linealmente independentes, procedemos do xeito habitual, escribindo $\rho = \sum_{i=1}^n \sum_{j=1}^m \lambda_{i,j} u_i \otimes v_j$; avaliando en (u_h^*, v_k^*), obtemos que $0 = \rho(u_h^*, v_k^*) = \lambda_{h,k}$, polo que todos os coeficientes son cero.

Unicamente queda por ver que o par $(\mathcal{L}(E^* \times F^*, K), \varphi)$ cumpre a propiedade universal. Para iso, sexa (H, ψ) un par con $\psi\colon E \times F \to H$ bilineal. Calquera aplicación $f\colon G \to H$ tal que $f \circ \varphi = \psi$ ten que cumprir que $(f \circ \varphi)(u_i, v_j) = \psi(u_i, v_j)$, é dicir, $f(u_i \otimes v_j) = \psi(u_i, v_j)$. Como $u \otimes v$ é unha base de G, sabemos que existe unha única aplicación lineal $f\colon G \to H$ que cumpre $f(u_i \otimes v_j) = \psi(u_i, v_j)$ para todo i, j. Para esta única f, se $x = \sum_{i=1}^n x_i u_i \in E$ e $y = \sum_{j=1}^n y_j v_j \in F$, entón

$$\begin{aligned}(f\circ\varphi)(x,y) &= f\Big(\sum_{i=1}^n x_i y_j \varphi(u_i, v_j)\Big)\\ &= \sum_{i=1}^n\sum_{j=1}^m x_i y_j f(u_i\otimes v_j)\\ &= \sum_{i=1}^n\sum_{j=1}^m x_i y_j \psi(u_i, v_j)\\ &= \psi(x,y).\end{aligned}$$

Polo tanto, $f\circ\varphi = \psi$ e queda probado que existe unha única aplicación lineal $f\colon G \to H$ tal que $f\circ\varphi = \psi$. □

En termos de bases, temos a seguinte caracterización.

Proposición 5.17. Coas notacións anteriores, sexa (G, φ) un par e $\varphi\colon E\times F \to G$ bilineal. Entón, as seguintes propiedades son equivalentes:

(a) (G, φ) cumpre a propiedade universal.

(b) Existen bases u de E e v de F tal que $\{\varphi(u_i, v_j)$ con $1 \le i \le n; 1 \le j \le m\}$ é base de G.

(c) Para calquera elección de bases u de E e v de F, $\{\varphi(u_i, v_j)$ con $1 \le i \le n$; $1 \le j \le m\}$ é base de G.

Demostración. Vemos primeiro que (a) implica (c). Vimos xa que $(\mathcal{L}(E^*\times F^*, K), \varphi)$ cumpre a propiedade universal e se collemos bases u de E e v de F,

$$u\otimes v = \{\varphi(u_i, v_j) \text{ con } 1\le i\le n,\ 1\le j\le m\}$$

é base de $\mathcal{L}(E^*\times F^*, K)$. Por hipótese, (G,φ) cumpre a propiedade universal, e sabemos que un par que cumpra a propiedade universal é único salvo isomorfismo. Polo tanto, existe un isomorfismo $f\colon \mathcal{L}(E^*\times F^*, K) \to G$ de espazos vectoriais tal que $f\circ\varphi = \varphi$. Polo tanto, se fixamos bases u de E e v de F,

$$\{\varphi(u_i, v_j) \text{ con } 1\le i\le n,\ 1\le j\le m\} = f(\{\varphi(u_i, v_j) \text{ con } 1\le i\le n,\ 1\le j\le m\})$$

é unha base de G porque é a imaxe polo isomorfismo f dunha base de $\mathcal{L}(E^*\times F^*, K)$. O feito de que (c) implique (b) é evidente. Por último, podemos ver que (b) implica (a). Por hipótese, $\{\varphi(u_i, v_j)\}$ é unha base de G. Sexa (H,ψ) un par. Queremos ver que existe un único morfismo de pares $f\colon (G,\varphi) \to (H,\psi)$. Como $f\circ\varphi$ é igual a ψ, iso quere dicir que $f(\varphi(u_i, v_j)) = \psi(u_i, v_j)$. Como as aplicacións lineais quedan univocamente determinadas pola imaxe dunha base, deducimos que existe unha única aplicación lineal f tal que $f(\varphi(u_i, v_j)) = \psi(u_i, v_j)$ para todo i, j. □

Definición 5.14. Un *produto tensorial de E e F* é un par (G,φ) que cumpre a propiedade universal (e que polo visto agora, é único salvo isomorfismo). O par $(\mathcal{L}(E^*\times F^*, K), \varphi)$, con $\varphi\colon E\times F \to \mathcal{L}(E^*\times F^*, K)$ dado por $\varphi(x,y) = x\otimes y$ cumpre a propiedade universal, polo que se acostuma denotar *o* produto tensorial de E e F.

Proposición 5.18. (a) Sexa G un K-espazo vectorial. Entón

$$\phi\colon \mathcal{L}(E \times F, G) \longrightarrow \mathcal{L}(E \otimes F, G)$$

definida por $\phi(f) = g$, onde $f\colon E \otimes F \to G$, é a única aplicación lineal tal que $g \circ \varphi = f$ é un isomorfismo de espazos vectoriais. Aquí, $\varphi\colon E \times F \to E \otimes F$ é a aplicación dada pola propiedade universal.

(b) Existe un isomorfismo canónico $(E \otimes F)^* \cong E^* \otimes F^*$.

Demostración. Pola propiedade universal do produto tensorial, dada $f\colon E \times F \to G$, existe unha única aplicación lineal $g\colon E \otimes F \to G$ que maneira que $g \circ \varphi = f$. Polo tanto, $\phi(f) = g$, onde $g \circ \varphi = f$ está ben definida. A aplicación ϕ é inxectiva, porque se $\phi(f) = \phi(f')$, entón $f = \phi(f) \circ \varphi = \phi(g) \circ \varphi = g$. Do mesmo xeito, é sobrexectiva xa que dada $g\colon E \otimes F \to G$ podemos coller como antiimaxe $f := g \circ \varphi$. A linealidade é evidente, polo que temos o isomorfismo pedido.
Para a segunda parte, temos que $E^* \otimes F^* = \mathcal{L}(E^{**} \times F^{**}, K)$ Usando os isomorfismos canónica entre un espazo e o seu bidual, e collendo $G = K$ na primeira parte da proposición, obtemos

$$E^* \otimes F^* = \mathcal{L}(E^{**} \times F^{**}, K) \simeq \mathcal{L}(E \times F, K) \simeq \mathcal{L}(E \otimes F, K) = (E \otimes F)^*.$$

□

Unha das aplicacións habituais dos produtos tensoriais é proporcionar unha linguaxe na que entender *a extensión de escalares.* Por exemplo, dado un polinomio en $\mathbb{R}[X]$, frecuentemente pasamos a consideralo como un polinomio con coeficientes en $\mathbb{C}$; iso pódese representar en termos do isomorfismo

$$\mathbb{R}[X] \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathbb{C}[X],$$

onde o subíndice $\mathbb{R}$ indica que o produto tensorial se considera sobre os $\mathbb{R}$-espazos vectoriais.
Os produtos tensoriais son especialmente interesantes cando en vez de traballar sobre espazos vectoriais traballamos sobre *módulos.* É dicir, en lugar de considerar que o produto por escalares ocorre cos coeficientes nun corpo, pasamos a traballar cun anel R. A principal diferenza é que non podemos falar de bases do mesmo xeito que para espazos vectoriais. Por exemplo, no caso do $\mathbb{Z}$-módulo $\mathbb{Z}/n\mathbb{Z}$, o elemento 1 é un xerador, pero non é linealmente independente xa que $n \times 1 = 0$.
Outro exemplo habitual na liña anterior consiste en reducir módulo p un polinomio con coeficientes enteiros. Nese caso, vemos $(\mathbb{Z}/p\mathbb{Z})$ como un $\mathbb{Z}$-módulo e usamos que

$$\mathbb{Z}[X] \otimes_{\mathbb{Z}} \mathbb{Z}/p\mathbb{Z} \simeq (\mathbb{Z}/p\mathbb{Z})[X].$$

Un exemplo ilustrativo dun produto tensorial no que o resultado non é evidente é o seguinte, cuxa demostración omitimos.

Proposición 5.19. Sexan m e n enteiros positivos. Entón,

$$\mathbb{Z}/m\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/\gcd(m,n)\mathbb{Z}.$$

A idea esencial é que dado un elemento $x \otimes y$, con $x \in \mathbb{Z}/m\mathbb{Z}$ e $y \in \mathbb{Z}/n\mathbb{Z}$, podemos definir un elemento en $\mathbb{Z}/\gcd(m,n)\mathbb{Z}$ tomando o produto das reducións de x e y módulo d, é dicir,

$$\mathbb{Z}/m\mathbb{Z} \otimes \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/\gcd(m,n)\mathbb{Z}, \quad (x,y) \mapsto xy \pmod{\gcd(m,n)}.$$

Noutras palabras, as aplicacións bilineais que se poden definir a partir de $\mathbb{Z}/m\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z}$ son as que *factorizan* a través de $\mathbb{Z}/\gcd(m,n)\mathbb{Z}$.

5.6. Tensores e integración

Nesta sección, consideramos o espazo vectorial $\mathbb{R}^n$. En cada punto $p \in \mathbb{R}^n$, temos a noción de *espazo tanxente*, que denotamos como $\mathbb{R}^n_p$ e que se corresponde coas diferentes direccións que pode tomar o gradiente dunha curva diferenciable que pase por dito punto. Recordamos que $\mathbb{R}^n_p \simeq \mathbb{R}^n$ para todo p.

Definición 5.15. Unha *k-forma diferencial en* $\mathbb{R}^n$ (ou forma diferencial de grao k) é unha aplicación

$$\omega \colon \mathbb{R}^n \longrightarrow \bigcup_p \Big(\bigwedge^k (\mathbb{R}^n_p)^* \Big), \quad p \mapsto \omega(p),$$

onde

$$\omega(p) \colon \mathbb{R}^n_p \times \cdots \times \mathbb{R}^n_p \longrightarrow \mathbb{R} \quad (v_1, \ldots, v_k) \mapsto \omega(p)(v_1, \ldots, v_k).$$

En particular, podemos pór

$$\omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} \omega_{i_1 \ldots i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \tag{5.1}$$

onde $\omega_{i_1,\ldots,i_k} \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ son as funcións compoñentes de ω (que suporemos sempre de clase $\mathcal{C}^\infty$). As mesmas definicións funcionan cando cambiamos $\mathbb{R}^n$ por un aberto U.
O conxunto de formas diferenciais de grao k nun aberto $U \subset \mathbb{R}^n$ denótase como $\Omega^k(U)$.

Identificamos $\Omega^0(U)$ co conxunto de funcións diferenciables U, é dicir, con $\mathcal{C}^\infty(U)$. Como vimos neste tema, as formas diferenciais de grao k forman un $\mathbb{R}$-espazo vectorial; hai ademais unha operación de produto exterior entre as formas de grao k e grao ℓ,

$$\bigwedge \colon \Omega^k(U) \times \Omega^\ell(U) \longrightarrow \Omega^{k+\ell}(U).$$

Moitas das definicións realizarémolas por simplicidade en $\mathbb{R}^n$, pero as mesmas definicións serven para abertos calquera.
Na teoría da integración, un concepto fundamental é o de diferencial exterior, que dá unha aplicación lineal entre as formas diferenciais de grao k e as de grao $k+1$. É dicir, é unha aplicación lineal dentro da álxebra exterior, pero que non preserva o grao.

Definición 5.16. A *diferencial exterior* defínese como unha aplicación lineal entre as formas diferenciais de grao k e as de $k+1$. Se $f \in \Omega^0(\mathbb{R}^n)$, entón

$$df := \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i;$$

se $\omega \in \Omega^k(\mathbb{R}^n)$ e $k \geq 1$, coas notacións anteriores,

$$d\omega := \sum_{1 \leq i_1 < \ldots < i_k \leq n} d\omega_{i_1,\ldots,i_k} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

A seguinte proposición recolle algunhas das propiedades máis importantes da diferencial exterior.

Proposición 5.20. Supoñamos que $a, b \in \mathbb{R}$, ω_1, $\omega_2 \in \Omega^k(\mathbb{R}^n)$ i $\eta \in \Omega^\ell(\mathbb{R}^n)$.

(1) $d(a\omega_1 + b\omega_2) = a\, d\omega_1 + b\, d\omega_2$;

(2) $d(\omega_1 \wedge \eta) = d\omega_1 \wedge \eta + (-1)^k \omega_1 \wedge d\eta$;

(3) $d^2(\omega_1) = d(d\omega_1) = 0$.

Sexa $\varphi\colon \mathbb{R}^m \longrightarrow \mathbb{R}^n$ unha función diferenciable. Considérase a aplicación

$$\varphi^*\colon \Omega^k(\mathbb{R}^n) \longrightarrow \Omega^k(\mathbb{R}^m) \quad \omega \mapsto \varphi^*\omega$$

definida da seguinte maneira.

Definición 5.17. Sexa $f \in \Omega^0(\mathbb{R}^n)$; nese caso, defínese $\varphi^* f = f \circ \varphi$. Se $\omega \in \Omega^k(\mathbb{R}^n)$, con $k > 0$, $p \in \mathbb{R}^m$ e $v_1, \ldots, v_k \in \mathbb{R}^m_p$, entón

$$(\varphi^*\omega)_p(v_1, \ldots, v_k) := \omega_{\varphi(p)}(d\varphi_p(v_1), \ldots, d\varphi_p(v_k)).$$

A k-forma $\varphi^*\omega$ chámase *pull-back* de ω por φ.

Algunhas das propiedades que cumpre o pull-back son as seguintes.

Proposición 5.21. Supoñamos que $a, b \in \mathbb{R}$, ω_1, $\omega_2 \in \Omega^k(\mathbb{R}^n)$ e $\eta \in \Omega^\ell(\mathbb{R}^n)$.

(1) $\varphi^*(a\omega_1 + b\omega_2) = a\,\varphi^*\omega_1 + b\,\varphi^*\omega_2$;

(2) $\varphi^*(\omega_1 \wedge \eta) = \varphi^*\omega_1 \wedge \varphi^*\eta$;

(3) $\varphi^*(d\omega_1) = d(\varphi^*\omega_1) = 0$.

A diferencial dunha función f é unha aplicación lineal entre os espazos tanxentes en p e en $f(p)$. O pull-back pode interpretarse como a aplicación dual da diferencial. É dicir, en cada punto do espazo temos un espazo vectorial asociado (o espazo tanxente) e o seu espazo dual (o espazo cotanxente, formado polas formas diferenciais). Por exemplo, no caso da esfera $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$, o espazo tanxente en $(0, 0, 1)$ é o subespazo vectorial de $\mathbb{R}^3$ xerado por $(1, 0, 0)$ e $(0, 0, 1)$.
O pull-back dá unha aplicación lineal entre os espazos cotanxentes; como en $\mathbb{R}^n$ o único que será posible é integrar formas diferenciais de orde n, o procedemento para integrar nunha variedade de dimensión n en $\mathbb{R}^m$ é considerar unha parametrización $\varphi\colon U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ e dada unha m-forma en $\mathbb{R}^m$ considerar o seu pull-back $\varphi^*(\omega)$.

Exemplo. Na esfera S centrada na orixe e de raio R, consideramos a forma diferencial

$$\omega = \Big(\frac{x}{r^3} + a\Big) dy \wedge dz + \Big(\frac{y}{r^3} + b\Big) dz \wedge dx + \Big(\frac{z}{r^3} + c\Big) dx \wedge dy,$$

con $a, b, c \in \mathbb{R}$. Esta expresión é inconveniente de cara á teoría de integración, porque a esfera é unha variedade de dimensión 2 e para describila estamos empregando a súa descrición en $\mathbb{R}^3$, que ten dimensión 3. Consideramos a parametrización da esfera dada por

$$\sigma\colon D = [0, \pi) \times [0, 2\pi) \longrightarrow \mathbb{R}^3, \quad (\varphi, \theta) \mapsto (R \sin\varphi \cos\theta, R \sin\varphi \sin\theta, R \cos\varphi).$$

Podemos entón considerar o pull-back $\sigma^*\omega$, que será unha forma diferencial en D, que ten dimensión 2. En particular, temos que

$$dy \wedge dz = R^2 \sin^2\varphi \cos\theta, \quad dz \wedge dx = R^2 \sin^2\varphi \sin\theta, \quad dx \wedge dy = R^2 \sin\varphi \cos\varphi.$$

Na teoría de integración, defínese $\int_S \omega := \int_D \sigma^*\omega$, polo que neste caso poderiamos comprobar directamente que

$$\int_S \omega = 2\pi \int_0^\pi \sin\varphi \, d\varphi = 4\pi.$$

En particular, o resultado non depende de a, b e c.

Definición 5.18. Sexa $\omega \in \Omega^k(U)$ unha forma diferencial. Dise que ω é *pechada* se $d\omega = 0$. Dise que ω é *exacta* se existe $\eta \in \Omega^{k-1}(U)$ de maneira que $d\eta = \omega$.

Como $d^2\omega = 0$ para calquera $\omega \in \Omega^k(U)$, tense que toda forma exacta é pechada. Como as formas diferenciais exactas e pechadas forman $\mathbb{R}$-espazos vectoriais, ten sentido definir a *cohomoloxía de de Rham* de grao k como

$$H^k(U) = \frac{\{\omega \in \Omega^k(U) \text{ tal que } \omega \text{ é pechada}\}}{\{\omega \in \Omega^k(U) \text{ tal que } \omega \text{ é exacta}\}}.$$

O teorema de de Rham afirma que $H^k(U)$ unicamente depende da xeometría de U, é dicir, relaciona uns obxectos de carácter analítico (as formas diferenciais) con outros de carácter topolóxico.

5.7. Tensores en física: definición en termos de coordenadas

En física e en enxeñería, os tensores non se definen como aplicacións multilineais, senón en termos do seu comportamento por cambios de coordenadas. A seguinte definición de tensor é bastante habitual en libros de física.

Definición 5.19. Sexa E un espazo vectorial de dimensión $n \geq 1$. Un *tensor de rango 0* en E é un escalar. Para $k \geq 1$, un *tensor contravariante de rango k* en E é un obxecto T con n^k compoñentes en calquera sistema de coordenadas de E de maneira que, se $\{T^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i_k,\leq n}$ e $\{\tilde{T}^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i,\leq n}$ son as compoñentes de T en dous sistemas de coordenadas, entón

$$\tilde{T}^{i_1,\dots,i_k} = \sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k} a_{i_1j_1}\cdots a_{i_kj_k},$$

onde (a_{ij}) é a matriz correspondente a expresar o primeiro sistema de coordenadas de E en termos do segundo.

Esta é, precisamente, a definición dun elemento de $E^{\otimes k}$. Para demostrar iso, fixamos unha base $\{v_1,\dots,v_n\}$ de E, de maneira que T teña compoñentes $\{T^{i_1,\dots,i_k}\}_{1\leq i_1,\dots,i_k\leq n}$:

$$T = \sum_{1\leq i_1,\dots,i_k\leq n} T^{i_1,\dots,i_k} v_{i_1}\otimes\cdots\otimes v_{i_k} \in E^{\otimes k}.$$

Consideremos agora un segundo sistema de coordenadas $\{u_1,\dots,u_n\}$, e poñamos $v_j = \sum_{i=1}^n a_{ij}u_i$. Entón,

$$\begin{aligned}
T &= \sum_{1\leq i_1,\dots,i_k\leq n} T^{i_1,\dots,i_k} v_{i_1}\otimes\cdots\otimes v_{i_k}\\
&= \sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k}\Big(\sum_{i_1=1}^n a_{i_1j_1}u_{i_1}\Big)\otimes\cdots\otimes\Big(\sum_{i_k=1}^n a_{i_kj_k}u_{i_k}\Big)\\
&= \sum_{1\leq i_1,\dots,i_k\leq n}\Big(\sum_{1\leq j_1,\dots,j_k\leq n} T^{j_1,\dots,j_k} a_{i_1j_1}\cdots a_{i_kj_k}\Big)u_{i_1}\otimes\cdots\otimes u_{i_k}\\
&= \sum_{1\leq i_1,\dots,i_k\leq n} \tilde{T}^{i_1,\dots,i_k} u_{i_1}\otimes\cdots\otimes u_{i_k}.
\end{aligned}$$

Polo tanto, en física as compoñentes dun tensor contravariante de rango k en E son os coeficientes dun elemento de $E^{\otimes k}$ nalgunha base de $E^{\otimes k}$.
Isto todo pódese xeralizar ao caso do espazo dual. Por consistencia coas notacións habituais en física ou xeometría diferencial, imos escribir as bases duais de $\{v_1, \ldots, v_n\}$ e $\{u_1, \ldots, u_n\}$ como $\{v^1, \ldots, v^n\}$ e $\{u^1, \ldots, u^n\}$, respectivamente. Unha base de E preséntase con subíndices, mentres que os coeficientes levan superíndices; e no caso de E^* unha base escríbese con superíndices e os coeficientes con subíndices. Ademais, se a matriz que pasa da base $\{u_1, \ldots, u_n\}$ a $\{v_1, \ldots, v_n\}$ se denota por (a_{ij}), a inversa escríbese como (a^{ij}), a trasposta como (a_{ji}) e a inversa da trasposta como (a^{ji}).

Definición 5.20. Para $k \geq 1$, un *tensor covariante de rango ℓ* en E é un obxecto T con n^ℓ compoñentes en calquera sistema de coordenadas de E de maneira que, se $\{T_{i_1,\ldots,i_\ell}\}_{1 \leq i_1,\ldots,i_\ell, \leq n}$ e $\{\tilde{T}_{i_1,\ldots,i_\ell}\}_{1 \leq j_1,\ldots,j_\ell, \leq n}$ son as compoñentes de T en dous sistemas de coordenadas, entón

$$\tilde{T}_{i_1,\ldots,i_\ell} = \sum_{1 \leq j_1,\ldots,j_\ell \leq n} T_{j_1,\ldots,j_\ell} a^{j_1 i_1} \cdots a^{j_\ell i_\ell},$$

onde (a^{ji}) é a inversa da transposta da matriz correspondente a expresar o primeiro sistema de coordenadas de E en termos do segundo.

A comprobación neste caso é análoga á anterior:

$$\begin{aligned}
T &= \sum_{1 \leq j_1,\ldots,j_\ell \leq n} T_{j_1,\ldots,j_\ell} e^{j_1} \otimes \cdots \otimes e^{j_\ell} \\
&= \sum_{1 \leq j_1,\ldots,j_\ell \leq n} T_{j_1,\ldots,j_\ell} \Big(\sum_{i_1=1}^{n} a^{j_1 i_1} f^{i_1} \Big) \otimes \cdots \otimes \Big(\sum_{i_\ell=1}^{n} a^{j_\ell i_\ell} f^{i_\ell} \Big) \\
&= \sum_{1 \leq i_1,\ldots,i_\ell \leq n} \Big(\sum_{1 \leq j_1,\ldots,j_\ell \leq n} T_{j_1,\ldots,j_\ell} a^{j_1 i_1} \cdots a_{j_\ell i_\ell} \Big) f^{i_1} \otimes \cdots \otimes f^{i^\ell} \\
&= \sum_{1 \leq i_1,\ldots,i_\ell \leq n} \tilde{T}_{i_1,\ldots,i_\ell} f^{i_1} \otimes \cdots \otimes f^{i_\ell}.
\end{aligned}$$

Por último, as definicións esténdense de xeito análogo ao caso de tensores mixtos.

Exemplo. Imos discutir as diferentes transformacións para tensores de rango 2. Para iso, mantemos as notacións das definicións anteriores.

- No caso de tensores de tipo $(2, 0)$, temos que

$$(\tilde{T}^{i_1 i_2}) = \sum_{j_1, j_2} T^{j_1 j_2} a_{i_1 j_1} a_{i_2 j_2}.$$

- No caso de tensores de tipo $(0, 2)$, temos que

$$(\tilde{T}_{i_1 i_2}) = \sum_{j_1, j_2} (T_{j_1 j_2}) a^{j_1 i_1} a^{j_2 i_2}.$$

- No caso de tensores de tipo $(1, 1)$, temos que

$$(\tilde{T}^{i_1}_{i_2}) = \sum_{j_1, j_2} T^{j_1}_{j_2} a_{i_1 j_1} a^{j_2 i_2}.$$

O caso de tipo $(0,2)$ correspóndese coa lei de transformación dunha forma bilineal por cambios de base, e os de tipo $(1,1)$ amosan a transformación dunha aplicación linear. Iso permítenos pensar nunha aplicación bilineal como un tensor de tipo $(0,2)$ e nunha aplicación linear como nun tensor de tipo $(1,1)$, xa que $E \otimes E^* \simeq \mathrm{End}(E)$.

Esta sección pódese resumir coa frase *os físicos pensan nos tensores como sistemas de compoñentes organizados por un ou máis índices que se transforman de acordo a un conxunto específico de regras.*

5.8. Problemas

Definicións e primeiras propiedades.

Problema 5.1. Sexa E un espazo vectorial sobre un corpo K. Sexa $\varphi\colon E \times E^* \to K$ definida por $\varphi(x,\omega) = \omega(x)$.

(a) Demostrar que φ é unha forma lineal en ambas compoñentes.

(b) Sexa F un subespazo vectorial de E e H un subespazo vectorial de E^*. Probar que $F^\perp = \{\omega \in E^* \mid \varphi(x,\omega) = 0 \text{ para todo } x \text{ de } F\}$ e $H^\perp = \{x \in E \mid \varphi(x,\omega) = 0 \text{ para todo } \omega \text{ de } H\}$ son subespazos vectoriais de E^* e E, respectivamente, e que $\dim F + \dim F^\perp = \dim E$ e $\dim H + \dim H^\perp = \dim E$.

(c) Demostrar que $F^{\perp\perp} = F$ e que $H^{\perp\perp} = H$.

(d) Demostrar que se $F_1 \subset F_2$, entón $F_1^\perp \supset F_2^\perp$ e que se $H_1 \subset H_2$, $H_1^\perp \supset H_2^\perp$.

(e) Demostrar que $(F_1 + F_2)^\perp = F_1^\perp \cap F_2^\perp$ e $(H_1 + H_2)^\perp = H_1^\perp \cap H_2^\perp$.

(f) Demostrar que $(F_1 \cap F_2)^\perp = F_1^\perp + F_2^\perp$ e $(H_1 \cap H_2)^\perp = H_1^\perp + H_2^\perp$.

(g) Demostrar que existe unha bixección entre os subespazos de E e o de E^* dada por $F \mapsto F^\perp$ e $H \mapsto H^\perp$.

Solución. (a) Hai que comprobar que φ é bilineal nas dúas variables. Ímolo comprobar na primeira. Para a suma, temos que

$$\varphi(x+y,\omega) = \omega(x+y) = \omega(x) + \omega(y) = \varphi(x,\omega) + \varphi(y,\omega);$$

para o produto por escalares,

$$\varphi(\lambda x,\omega) = \omega(\lambda x) = \lambda\omega(x) = \lambda\varphi(x,\omega).$$

Na segunda variable é análoga, vendo que $\varphi(x,\omega_1+\omega_2) = \varphi(x,\omega_1) + \varphi(x,\omega_2)$ e $\varphi(x,\lambda\omega) = \lambda\varphi(x,\omega)$.

(b) No caso de $F^\perp$, se $\varphi(x,\omega_1) = 0$ e $\varphi(x,\omega_2)$, entón $\varphi(x,\omega_1+\omega_2) = 0$ e $\varphi(x,\lambda\omega) = 0$, onde empregamos a linealidade do apartado anterior. O mesmo razoamento funciona para probar que $H^\perp$ é un subespazo.

Supoñamos que $r = \dim F$, e fixamos unha base $\mathcal{B}_v = \{v_1,\dots,v_r\}$ de F, que podemos completar a unha base $\{v_1,\dots,v_n\}$ de E. Sexa $\{v_1^*,\dots,v_n^*\}$ a base dual. Temos que $\langle v_{r+1}^*,\dots,v_n^*\rangle \subset F^\perp$, porque todas esas formas lineais anulan todos os vectores de E. Para ver a inclusión oposta, sexa $\omega = \sum_{i=1}^n \lambda_i v_i^* \in F$, de maneira que para algún i, con $1 \leq i \leq r$, se cumpre que $\lambda_i \neq 0$. Entón, $\varphi(e_i,\omega) = \lambda_i \neq 0$, que é unha contradición. Polo tanto, $\dim F^\perp = n - r$. O caso de H demóstrase de xeito análogo.

(c) Comezamos observando que $F \subset F^{\perp\perp}$. Por definición,

$$F^{\perp\perp} = \{x \in E \mid \omega(x) = 0 \text{ para todo } \omega \in F^{\perp}\},$$

e se ω cumpre que $\omega(x) = 0$ para todo $x \in F$, entón tense que $F \subset F^{\perp\perp}$. Por outro lado, temos que as dimensións coinciden:

$$\dim F^{\perp\perp} = \dim E - \dim F^{\perp} = \dim E - (\dim E - \dim F) = \dim F.$$

Se temos unha inclusión e as dimensións coinciden, necesariamente son iguais. O caso de H é análogo.

(d) Supoñamos que $F_1 \subset F_2$. Para demostrar que $F_2^{\perp} \subset F_1^{\perp}$, sexa $\omega \in F_2^{\perp}$. Polo tanto, $\varphi(x, \omega) = 0$ para todo $x \in F_2$. Como $F_1 \subset F_2$, tense en particular que $\varphi(x, \omega) = 0$ para todo $x \in F_1$, polo que $\omega \in F_1^{\perp}$. O caso de H_1 e H_2 é análogo.

(e) Por definición,

$$(F_1 + F_2)^{\perp} = \{\omega \in E^* \mid \varphi(x + y, \omega) = 0 \text{ para todo } x \in F_1,\, y \in F_2\}.$$

Imos ver que $F_1^{\perp} \cap F_2^{\perp} \subset (F_1 + F_2)^{\perp}$. Para iso, temos que probar que se $\omega \in F_1^{\perp} \cap F_2^{\perp}$, entón $\varphi(x + y, \omega) = 0$ para todo $x \in F_1$ e $y \in F_2$. Temos que

$$\varphi(x + y, \omega) = \varphi(x, \omega) + \varphi(y, \omega) = 0 + 0 = 0,$$

onde usamos que $\varphi(x, \omega) = 0$ porque $\omega \in F_1^{\perp}$ e $x \in F_1$ e $\varphi(y, \omega)$ xa que $\omega \in F_2^{\perp}$ e $y \in F_2$. Para ver a inclusión oposta, sexa $\omega \in (F_1 + F_2)^{\perp}$. Temos que comprobar que $\omega \in F_1^{\perp}$ e $\omega \in F_2^{\perp}$. Se $\varphi(x + y, \omega) = 0$ para todo $x \in F_1$ e $y \in F_2$, pondo $y = 0$ temos que $\varphi(x, \omega) = 0$ para todo $x \in F_1$, polo que $\omega \in F_1^{\perp}$, e polo mesmo argumento, $\omega \in F_2^{\perp}$. O caso de H_1 e H_2 é análogo.

(f) Comezamos vendo que $F_1^{\perp} + F_2^{\perp} \subset (F_1 \cap F_2)^{\perp}$. Para iso, sexa $\omega_1 + \omega_2 \in F_1^{\perp} + F_2^{\perp}$ e sexa $x \in F_1 \cap F_2$. Entón,

$$\varphi(x, \omega_1 + \omega_2) = \varphi(x, \omega_1) + \varphi(x, \omega_2) = 0 + 0 = 0,$$

onde empregamos que $\varphi(x, \omega_1) = 0$ xa que $x \in F_1$ e $\omega_1 \in F_1^{\perp}$ e $\varphi(x, \omega_2) = 0$ xa que $x \in F_2$ e $\omega_2 \in F_2^{\perp}$. Do apartado anterior sabemos que $\dim(F_1 + F_2)^{\perp} = \dim F_1^{\perp} \cap F_2^{\perp}$. Aplicando a fórmula de Grassmann, iso é equivalente a

$$n - \dim(F_1 + F_2) = \dim F_1^{\perp} + \dim F_2^{\perp} - \dim(F_1^{\perp} + F_2^{\perp}).$$

Reorganizando termos, quédanos que

$$\begin{aligned} \dim(F_1^{\perp} + F_2^{\perp}) &= n + \dim(F_1 + F_2) - \dim F_1 - \dim F_2 \\ &= n - \dim(F_1 \cap F_2) = (F_1 \cap F_2)^{\perp}. \end{aligned}$$

Polo tanto, as dimensións son iguais e ao termos unha inclusión podemos concluír.

Unha maneira alternativa pasa por traballar en termos de bases. Sexa $\{v_1, \ldots, v_r\}$ unha base de $F_1 \cap F_2$; completámola de maneira que $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_s\}$ sexa base de F_1 e $\{v_1, \ldots, v_r, v_{s+1}, \ldots, v_t\}$ sexa base de F_2. Temos entón que

$\langle v_{r+1},\dots,v_s\rangle \cap \langle v_{s+1},\dots,v_t\rangle = \{0\}$. Podemos logo considerar unha base de E da forma $\{v_1,\dots,v_r,\dots,v_s,\dots,v_t,\dots,v_n\}$. Entón,

$$\begin{aligned} F_1^\perp + F_2^\perp &= \langle v_{s+1}^*,\dots,v_n^*\rangle + \langle v_{r+1}^*,\dots,v_s^*,v_{t+1}^*,\dots,v_n^*\rangle \\ &= \langle v_{r+1}^*,\dots,v_n^*\rangle \\ &= (F_1\cap F_2)^\perp. \end{aligned}$$

O caso de H_1 e H_2 é análogo.

(g) Consideramos a aplicación

$$\{\text{subespazos de } E\} \longrightarrow \{\text{subespazos de } E^*\}, \quad F \mapsto F^\perp.$$

A aplicación está ben definida xa que vimos que $F^\perp$ é un subespazo vectorial de E^*. Para establecer que é unha bixección, é suficiente con propoñer unha inversa e ver que a composición en ambos sentidos é a identidade. A inversa é a dada no enunciado, $H \mapsto H^\perp$. Para ver a bixección pedida, é logo suficiente ver que $F^{\perp\perp} = F$ e $H^{\perp\perp} = H$, pero iso séguese do apartado (c).

Problema 5.2. Sexa E un espazo vectorial de dimensión n e $\{v_1,\dots,v_n\}$ unha base de E. Definimos a aplicación contración $C\colon T_1^1(E)\to\mathbb{R}$ por $C(F)=\sum_{j=1}^n F(v_j,v_j^*)$.

(a) Demostrar que C está ben definida (é dicir, non depende da elección de base).

(b) Demostrar que C é lineal.

(c) Se $v\in E$ e $\omega\in E^*$, demostrar que $C(\omega\otimes v)=\omega(v)$.

Solución. (a) Consideramos outra base $\{u_1,\dots,u_n\}$ de maneira que $u_r=\sum_{i=1}^n a_{ir}v_i$ e $u_s^*=\sum_{j=1}^n b_{js}v_j^*$, onde as matrices $A=(a_{ir})$ e $B=(b_{js})$ cumpren que $A^tB=\mathbb{I}_n$. En concreto, iso quere dicir que $\sum_{k=1}^n a_{ik}b_{jk}=\delta_{ij}$. Polo tanto,

$$\begin{aligned} \sum_{t=1}^n F(u_t,u_t^*) &= \sum_{t=1}^n F\Big(\sum_{i=1}^n a_{it}v_i, \sum_{j=1}^n b_{jt}v_j^*\Big) \\ &= \sum_{i=1}^n\sum_{j=1}^n F(v_i,v_j^*)\cdot\Big(\sum_{t=1}^n a_{ti}b_{tj}\Big) \\ &= \sum_{i=1}^n\sum_{j=1}^n F(v_i,v_j^*)\delta_{ij} \\ &= \sum_{t=1}^n F(v_t,v_t^*). \end{aligned}$$

(b) Temos que comprobar que $C(F+G)=C(F)+C(G)$ e que $C(\lambda F)=\lambda C(F)$, para $F,G\in T_1^1(E)$ e $\lambda\in\mathbb{R}$. Para iso, simplemente observamos que

$$\begin{aligned} C(F+G) &= \sum_{j=1}^n (F+G)(v_j,v_j^*) \\ &= \sum_{j=1}^n F(v_j,v_j^*) + \sum_{j=1}^n G(v_j,v_j^*) \\ &= C(F)+C(G) \end{aligned}$$

e que

$$\begin{aligned} C(\lambda F) &= \sum_{j=1}^{n} (\lambda F)(v_j, v_j^*) \\ &= \sum_{j=1}^{n} \lambda F(v_j, v_j^*) \\ &= \lambda C(F). \end{aligned}$$

(c) Escribimos $v = \sum_{i=1}^{n} \lambda_i v_i$ e $\omega = \sum_{j=1}^{n} \mu_j v_j^*$. Entón,

$$\omega(v) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \mu_j v_j^*(v_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \mu_i \delta_{ij} = \sum_{i=1}^{n} \lambda_i \mu_i.$$

Por outro lado,

$$C(\omega \otimes v) = \sum_{j=1}^{n} (\omega \otimes v)(v_j, v_j^*) = \sum_{j=1}^{n} \omega(v_j) v(v_j^*) = \sum_{j=1}^{n} \mu_j \lambda_j.$$

Problema 5.3. Sexa $\mathcal{B}_v = \{v_1, v_2, v_3, v_4)$ a base dun espazo vectorial E. Calcular o simetrizado e o antisimetrizado dos seguintes tensores:

(a) $v_1 \otimes v_3 + 2v_2 \otimes v_4 - v_3 \otimes v_4$.

(b) $2v_1 \otimes v_1 \otimes v_2 - v_1 \otimes v_1 \otimes v_3 - 2v_1 \otimes v_4 \otimes v_2 + v_1 \otimes v_4 \otimes v_3$.

Solución. (a) O simetrizado do primeiro tensor é

$$\frac{v_1 \otimes v_3 + v_3 \otimes v_1 + 2v_2 \otimes v_4 + 2v_4 \otimes v_2 - v_3 \otimes v_4 - v_4 \otimes v_3}{2},$$

mentres que o antisimetrizado é

$$\frac{v_1 \otimes v_3 - v_3 \otimes v_1 + 2v_2 \otimes v_4 - 2v_4 \otimes v_2 - v_3 \otimes v_4 + v_4 \otimes v_3}{2}.$$

(b) No segundo caso procedemos de xeito análogo, tendo en conta as diferenzas entre os dous primeiros termos e os dous seguintes. Por exemplo, o simetrizado do primeiro sumando é

$$\frac{2v_1 \otimes v_1 \otimes v_2 + 2v_1 \otimes v_2 \otimes v_1 + 2v_2 \otimes v_1 \otimes v_1}{3},$$

xa que hai un índice repetido. O simetrizado do terceiro sería, en cambio sumar sobre as seis permutacións posibles e dividir por 6. No caso do antisimetrizado, o resultado de aplicalo aos dous primeiros é 0 xa que hai índices iguais; para os dous seguintes, hai que sumar sobre as 6 permutacións colocando o signo positivo ou negativo segundo sexa par ou impar.

Produto exterior e tensores descompoñibles

Problema 5.4. En $T_2(\mathbb{R}^3)$ considéranse os tensores

$$\begin{aligned} f_1(x, y) &= 2x_1y_3 - 3x_2y_3 + x_2y_1 - x_3y_2 \\ f_2(x, y) &= 5x_2y_2 - x_1y_3 + 4x_3y_2 + 2x_3y_3. \end{aligned}$$

Calcular os seus antisimetrizados e facer o produto exterior.

Solución. Sexa $\{e_1, e_2, e_3\}$ a base canónica de $\mathbb{R}^3$ e $\{e_1^*, e_2^*, e_3^*\}$ a súa base dual. Temos que

$$\begin{aligned} A(f_1(x,y)) &= e_1^* \wedge e_3^* - \frac{3}{2}e_2^* \wedge e_3^* + \frac{1}{2}e_2^* \wedge e_1^* - \frac{1}{2}e_3^* \wedge e_2^* \\ &= e_1^* \wedge e_3^* - e_2^* \wedge e_3^* + \frac{1}{2}e_2^* \wedge e_1^*, \end{aligned}$$

e

$$A(f_2(x,y)) = -\frac{1}{2}e_1^* \wedge e_3^* + 2e_3^* \wedge e_2^*.$$

O produto exterior $A(f_1(x,y)) \wedge A(f_2(x,y))$ é 0 porque en todos os sumandos teremos sempre produtos exteriores da forma $e_{i_1}^* \wedge e_{i_2}^* \wedge e_{i_3}^* \wedge e_{i_4}^*$, con $i_1, i_2, i_3, i_4 \in \{1,2,3\}$. Polo tanto, dous dos subíndices teñen que ser iguais.

Problema 5.5. Sexa $\mathcal{B}_v = (v_1, v_2, v_3)$ a base canónica de $\mathbb{R}^3$. Considerar as aplicacións lineais dadas por

$$f(x,y,z) = 3x - y, \quad g(x,y,z) = x + y + z, \quad h(x,y,z) = 2z.$$

(a) Escribir f, g e h na base $\mathcal{B}_e^* = \{e_1^*, e_2^*, e_3^*)$ de $T_1(\mathbb{R}^3)$.

(b) Calcular $f \wedge g$, $f \wedge h$, $g \wedge h$ e $f \wedge g \wedge h$.

(c) Dados os vectores $u = (1,-1,2)$, $v = (0,3,-1)$ e $w = (-1,0,1)$, calcular $(f \wedge g \wedge h)(u,v,w)$.

Solución. (a) Temos que $f = 3e_1^* - e_2^*$, $g = e_1^* + e_2^* + e_3^*$ e $h = 2e_3^*$.

(b) É un cálculo rutineiro:

$$\begin{aligned} f \wedge g &= 4e_1^* \wedge e_2^* + 3e_1^* \wedge e_3^* - e_2^* \wedge e_3^*, \\ f \wedge h &= 6e_1^* \wedge e_3^* - 2e_2^* \wedge e_3^* \\ g \wedge h &= 2e_1^* \wedge e_3^* + 2e_2^* \wedge e_3^* \\ f \wedge g \wedge h &= 6e_1^* \wedge e_2^* \wedge e_3^* - 2e_2^* \wedge e_1^* \wedge e_3^* = 8e_1^* \wedge e_2^* \wedge e_3^* \end{aligned}$$

(c) Temos que $\det = e_1^* \wedge e_2^* \wedge e_3^*$, polo que $f \wedge g \wedge h = 8\det$. Neste caso, $\det(u,v,w) = 8$, polo que $(f \wedge g \wedge h)(u,v,w) = 64$.

Problema 5.6. Sexa $\mathcal{B}_v = (v_1, v_2, v_3, v_4)$ unha base dun espazo vectorial E. Dicir cal dos seguintes tensores son descompoñibles e, cando o sexan, dar unha descomposición:

(a) $v_1 \wedge v_2 + v_3 \wedge v_4$.

(b) $v_1 \wedge v_2 + v_2 \wedge v_3$.

(c) $v_1 \wedge v_2 \wedge v_3 + v_1 \wedge v_2 \wedge v_4 - 2v_1 \wedge v_3 \wedge v_4 + 3v_2 \wedge v_3 \wedge v_4$.

(d) $(2v_1 \wedge v_4 + 3v_2 \wedge v_3) \wedge (v_1 \wedge v_3 - v_1 \wedge v_4 - 2v_2 \wedge v_3 + v_2 \wedge v_4)$.

Solución. (a) Non é descompoñible xa que

$$(v_1 \wedge v_2 + v_3 \wedge v_4) \wedge (v_1 \wedge v_2 + v_3 \wedge v_4) = 2v_1 \wedge v_2 \wedge v_3 \wedge v_4,$$

que non é cero.

(b) Podemos poñer simplemente $v_2 \wedge (-v_1 + v_3)$.

(c) Consideramos unha descomposición da forma

$$\lambda(v_1+\alpha v_4)\wedge(v_2+\beta v_4)\wedge(v_3+\gamma v_4).$$

Igualando o termo con $v_1\wedge v_2\wedge v_3$ quédanos que $\lambda=1$. Igualando para $v_1\wedge v_2\wedge v_4$ vemos que $\gamma=1$; para $v_1\wedge v_3\wedge v_4$, temos que $\beta=2$; e para $v_2\wedge v_3\wedge v_4$, que $\alpha=3$.

A decomposición non é única. Poderiamos ter considerado, por exemplo, unha da forma

$$\lambda(v_1+\alpha v_3)\wedge(v_2+\beta v_3)\wedge(v_4+\gamma v_3).$$

Neste caso, igualando o termo con $v_1\wedge v_2\wedge v_4$ quédanos que $\lambda=1$. Igualando para $v_1\wedge v_2\wedge v_3$, vemos que $\gamma=1$; para $v_1\wedge v_3\wedge v_4$, temos que $\beta=2$; e para $v_2\wedge v_3\wedge v_4$, que $\alpha=-3$.

(d) Operando, temos que o tensor é igual a $-7v_1\wedge v_2\wedge v_3\wedge v_4$.

Problema 5.7. Sexa $E=\mathbb{R}_2[X]$ o espazo vectorial dos polinomios de grao menor ou igual que 2. Sexa $\mathcal{B}_e=\{e_1,e_2,e_3\}=\{1,X,X^2\}$ a base canónica e $\mathcal{B}_e^*=\{e_1^*,e_2^*,e_3^*\}$ a dual. Definimos $f_1,f_2,f_3:E\longrightarrow\mathbb{R}$ da seguinte maneira:

$$\begin{aligned} f_1(P(X)) &= P(-1)+P(0)+P(1),\\ f_2(P(X)) &= -P(-1)+P(1), \qquad .\\ f_3(P(X)) &= P(-1)+P(1). \end{aligned}$$

(a) Probar que $\{f_1,f_2,f_3\}$ é unha base de E^* e escribir eses elementos na base e^*.

(b) Expresar $f_1\wedge f_2$ e o simetrizado de $f_2\otimes f_3$ na base natural de $T_2(E)$.

(c) Existirán números reais α,β,γ de maneira que $\alpha\alpha f_1\wedge f_2+\beta f_2\wedge f_3+\gamma f_3\wedge f_1$ non sexa descompoñible? Dar unha descomposición de

$$f_1\wedge f_2-f_2\wedge f_3-f_3\wedge f_1.$$

Solución. (a) Comezamos observando que f_1, f_2 e f_3 son elementos do dual, e empregando as descricións do enunciado tense que

$$f_1=3e_1^*+2e_3^*,\quad f_2=2e_2^*,\quad f_3=2e_1^*+2e_3^*.$$

Como se ten que

$$\begin{vmatrix}3&0&2\\0&2&0\\2&0&2\end{vmatrix}=4\neq0,$$

temos que f_1, f_2 e f_3 son tres elementos linealmente independentes en E^*, que ten dimensión 3. Polo tanto, tamén son base.

(b) Comezamos observando que $T_2(E)$ é un espazo vectorial de dimensión 9 que ten como base natural $\{e_i^*\otimes e_j^*\}_{1\leq i,j\leq3}$. Un cálculo sinxelo amosa que

$$f_1\wedge f_2=6e_1^*\wedge e_2^*+4e_3^*\wedge e_2^*.$$

Usando que $e_i^*\wedge e_j^*=e_i^*\otimes e_j^*-e_j^*\otimes e_i^*$, obtense que

$$f_1\wedge f_2=6e_1^*\otimes e_2^*-6e_2^*\otimes e_1^*+4e_3^*\otimes e_2^*-4e_2^*\otimes e_3^*.$$

Por outra banda, tendo en conta que $S(f_2\otimes f_3)=\frac{1}{2}(f_2\otimes f_3+f_3\otimes f_2)$, quédanos que

$$S(f_2\otimes f_3)=2e_2^*\otimes e_1^*+2e_2^*\otimes e_3^*+2e_1^*\otimes e_2^*+2e_3^*\otimes e_2^*.$$

(c) Calquera tensor de $A_2(\mathbb{R}^3)$ é descompoñible, xa que, máis en xeral, se E ten dimensión n, un tensor en $A_{n-1}(E)$ é sempre descompoñible. Temos que

$$f_1 \wedge f_2 - f_2 \wedge f_3 - f_3 \wedge f_1 = 10e_1^* \wedge e_2^* + 2e_1^* \wedge e_3^* - 8e_2^* \wedge e_3^*.$$

Buscamos unha descomposición da forma

$$10e_1^* \wedge e_2^* + 2e_1^* \wedge e_3^* - 8e_2^* \wedge e_3^* = \lambda(\alpha e_1^* + e_2^*) \wedge (\beta e_1^* + e_3^*);$$

igualando termos resulta $\lambda = -8$, $-\lambda\beta = 10$ e $\lambda\alpha = 2$. Polo tanto, $\alpha = -1/4$ e $\beta = 5/4$. Polo tanto,

$$f_1 \wedge f_2 - f_2 \wedge f_3 - f_3 \wedge f_1 = -8\Big(-\frac{1}{4}e_1^* + e_2^*\Big) \wedge \Big(\frac{5}{4}e_1^* + e_3^*\Big).$$

Problema 5.8. Sexa $u \times v$ o produto vectorial en $\mathbb{R}^3$. Probar que a aplicación

$$\wedge^2\mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad u \wedge v \mapsto u \times v$$

é un isomorfismo.

Solución. Comezamos recordando que se $u = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$, entón

$$u \times v = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1).$$

Ambos espazos teñen dimensión 3, polo que é suficiente comprobar que se trata dunha aplicación lineal inxectiva. A linealidade é inmediata a partir da linealidade do produto vectorial. Para ver que é inxectiva, observamos que $u \times v$ se, e soamente se, $\langle u, v\rangle$ ten dimensión 1. Nese caso, $u \wedge v = 0$, polo que o único vector con imaxe 0 é o 0.

Problema 5.9. Sexa E un $\mathbb{R}$-espazo vectorial de dimensión 4 e $\{v_0, v_1, v_2, v_3\}$ unha base de E.

(a) Se $F = \langle u, v\rangle \subset E$ é un subespazo de dimensión 2, demostrar que o subespazo $\langle u \wedge v\rangle \subset \wedge^2 E$ non depende da elección da base u, v de F.

(b) Sexa $q\colon \wedge^2 E \to \mathbb{R}$ a forma cuadrática definida en coordenadas por

$$q(w) = x_{0,1}x_{2,3} - x_{0,2}x_{1,3} + x_{0,3}x_{1,2}.$$

Demostrar que $q(w) = 0$ se, e soamente se, $w \wedge w = 0$ en $\wedge^4 E$.

Solución. (a) Collamos outra base de F, $u' = au + bv$, $v' = cu + dv$, con $ad - bc \neq 0$. Entón,

$$u' \wedge v' = adu \wedge v + bcv \wedge u = (ad - bc)u \wedge v.$$

(b) Pomos

$$w = x_{0,1}v_0 \wedge v_1 + x_{0,2}v_0 \wedge v_2 + x_{0,3}v_0 \wedge v_3 + x_{1,2}v_1 \wedge v_2 + x_{1,3}v_1 \wedge v_3 + x_{2,3}v_2 \wedge v_3.$$

Temos que o produto exterior de w por el mesmo é

$$w \wedge w = 2(x_{0,1}x_{2,3} - x_{0,2}x_{1,3} + x_{0,3}x_{1,2})v_0 \wedge v_1 \wedge v_2 \wedge v_3.$$

Polo tanto, $w \wedge w = 0$ se, e soamente se, $q(w) = 0$.

Problema 5.10. Sexa E un K-espazo vectorial de dimensión $n \geq 2$. Sexa $f \in A_2(E)$. Dicimos que f ten rango $r = 2k$ se existen $w_1, \ldots, w_r \in E^*$ de maneira que

$$f = \omega_1 \wedge w_2 + \omega_3 \wedge \omega_4 + \ldots + \omega_{r-1} \wedge \omega_r$$

e r é o número mínimo co que se pode facer unha descomposición como esta.

(a) Demostrar que se f ten rango r e se ten unha descomposición da forma

$$f = \omega_1 \wedge w_2 + \omega_3 \wedge \omega_4 + \ldots + \omega_{r-1} \wedge \omega_r,$$

entón $\{\omega_1, \ldots, \omega_r\}$ son linealmente independentes.

(b) Demostrar que se f ten rango $r = 2k$, entón para todo $1 \leq m \leq k$ se cumpre que $f \wedge \ldots \wedge f$, onde o produto exterior se realiza m veces, é diferente de 0; e que se o produto exterior se realiza $k+1$ veces, entón é 0.

(c) En $E = \mathbb{R}^4$ consideramos a base canónica $\mathcal{B}_e = \{e_1, e_2, e_3, e_4\}$. Calcular o rango do tensor

$$f = e_1^* \wedge e_2^* + e_1^* \wedge e_3^* + e_1^* \wedge e_4^* + e_2^* \wedge e_3^* + e_2^* \wedge e_4^* + e_3^* \wedge e_4^*.$$

(d) Se f ten rango r e temos unha descomposición como a do enunciado, ampliamos os vectores $\{\omega_1, \ldots, \omega_r\}$ ata ter unha base $B = \{\omega_1, \ldots, \omega_n\}$ de E^*. Calcular a matriz A de f na base B, considerándoa como unha forma bilineal en E. Que relación ten o rango de f co rango de A?

Solución. (a) Supoñamos que non son independentes, de maneira que existe unha combinación lineal en termos dos ω_i. Sen perder xeralidade, podemos supor que ω_r participa na combinación lineal e que polo tanto se pode escribir

$$\omega_r = a_1\omega_1 + \ldots + a_{r-1}\omega_{r-1}.$$

Observamos agora que o termo $\omega_{r-1} \wedge (a_1\omega_1 + a_2\omega_2)$ se pode incorporar ao primeiro termo da suma, do xeito

$$\omega_1 \wedge \omega_2 + \omega_{r-1} \wedge (a_1\omega_1 + a_2\omega_2) = (\omega_1 + a_2\omega_{r-1}) \wedge (\omega_2 - a_1\omega_{r-1}).$$

Procedendo de xeito análogo cos coeficientes (a_3, a_4) e así ata (a_{2k-3}, a_{2k-2}), chegamos a que teriamos unha descomposición con $r-2$ elementos ω_i.

(b) Se facemos o produto exterior m veces, o termo $\omega_1 \wedge \ldots \wedge \omega_{2m}$ aparece multiplicado polo coeficiente $m!$, xa que todas as permutacións que o involucran son pares; alternativamente, estamos considerando permutacións da forma $\eta_1 \wedge \eta_2$, onde $\eta_1, \eta_2 \in A_2(E)$, e nese caso tense que $\eta_1 \wedge \eta_2 = \eta_2 \wedge \eta_1$. Porén, ao facer o produto exterior $k+1$ veces temos o produto exterior de $r+2$ termos da forma ω_i, con $1 \leq i \leq 2r$. Como só hai r diferentes, necesariamente algún deles terá que ser igual.

(c) Temos que $f \wedge f = 2e_1^* \wedge e_2^* \wedge e_3^* \wedge e_4^*$, mentres que $f \wedge f \wedge f$. Polo tanto, o rango é 4.

(d) Sexa $\{v_1, \ldots, v_n\}$ a base de E que ten por dual $\{\omega_1, \ldots, \omega_n\}$. Observamos que $\omega_1 \wedge \omega_2 = \omega_1 \otimes \omega_2 - \omega_2 \otimes \omega_1$, polo que $\omega_1 \wedge \omega_2(v_i, v_j)$ é 1 se $(i,j) = (1,2)$, -1 se $(i,j) = (2,1)$ e 0 en calquera outro caso. Polo tanto, temos que a matriz de f é unha matriz cos bloques diagonais 2×2 da forma

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

En particular, o seu rango é r.

Problema 5.11. Sexa $w \in E \otimes E$, con $E = \mathbb{R}^4$.

(a) Probar que, se $w \in A^2(E)$, entón existe unha base $\{v_1, v_2, v_3, v_4\}$ de E tal que $w \in \langle v_1 \wedge v_2, v_3 \wedge v_4 \rangle$.

(b) Sexa $\varphi : E^* \times E^* \to \mathbb{R}$ o tensor 2-contravariante asociado ao elemento $w \in E \otimes E$. Demostrar que $w \neq 0$ é descompoñible se, e soamente se, para calquera base u de E^*, a matriz da forma bilineal φ na base u é antisimétrica e ten rango 2.

Solución. (a) Se $w \neq 0$, podemos escoller dous vectores linealmente independentes $v_1, v_2 \in E$ de maneira que $w \wedge v_1 \wedge v_2 \neq 0$. Como a dimensión de $\wedge^4 E$ é igual a 1, existe $\lambda \in \mathbb{R}$ de maneira que $w \wedge w = 2\lambda w \wedge v_1 \wedge v_2$. Polo tanto,

$$0 = w \wedge (w - 2\lambda v_1 \wedge v_2) = (w - \lambda v_1 \wedge v_2) \wedge (w - \lambda v_1 \wedge v_2),$$

e iso quere dicir que $w - \lambda v_1 \wedge v_2$ é $\wedge$-descompoñible, isto é, existen $v_3, v_4 \in E$ de maneira que

$$w - \lambda v_1 \wedge v_2 = v_3 \wedge v_4.$$

En particular, $w \in \langle v_1 \wedge v_2, v_3 \wedge v_4 \rangle$. Finalmente, $w \wedge v_1 \wedge v_2 \neq 0$, polo que $v_3 \wedge v_4 \wedge v_1 \wedge v_2 \neq 0$, e iso proba que $\{v_1, v_2, v_3, v_4\}$ son linealmente independentes.

Imos dar agora unha solución alternativa e máis computacional. Na base canónica, escribimos

$$w = a_{12} e_1 \wedge e_2 + a_{13} e_1 \wedge e_3 + a_{14} e_1 \wedge e_4 + a_{23} e_2 \wedge e_3 + a_{24} e_2 \wedge e_4 + a_{34} e_3 \wedge e_4.$$

Entón, pomos $v_1 = e_1$ e $v_2 = a_{12} e_2 + a_{13} e_3 + a_{14} e_4$. Se $a_{23} = a_{24} = a_{34} = 0$, entón $w = v_1 \wedge v_2$ e rematamos. Supoñamos agora que un deses tres coeficientes é non cero; podemos pór, sen perda de xeralidade, $a_{23} \neq 0$. Entón, observamos que

$$a_{23} e_2 \wedge e_3 + a_{24} e_2 \wedge e_4 + a_{34} e_3 \wedge e_4 = a_{23}\Big(e_2 - \frac{a_{34}}{a_{23}} e_4\Big) \wedge \Big(e_3 + \frac{a_{24}}{a_{23}}\Big).$$

Polo tanto, pondo $v_3 = a_{23}\Big(e_2 - \frac{a_{34}}{a_{23}} e_4\Big)$ e $v_4 = e_3 + \frac{a_{24}}{a_{23}}$, temos que

$$w = v_1 \wedge v_2 + v_3 \wedge v_4.$$

Se $\{v_1, v_2, v_3, v_4\}$ son linealmente independentes, xa acabamos. En caso contrario, supomos que hai unha combinación lineal da forma $v_4 = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$. Nese caso,

$$w = v_1 \wedge v_2 - \lambda_1 v_1 \wedge v_3 - \lambda_2 v_2 \wedge v_3 = (v_1 + \lambda_2 v_3) \wedge (v_2 - \lambda_1 v_3),$$

polo que podemos coller $u_1 = v_1 + \lambda_2 v_2$, $u_2 = v_2 - \lambda_1 v_3$ e completar a base con dous vectores u_3 e u_4, de maneira que $\{u_1, u_2, u_3, u_4\}$ cumpre as condicións.

(b) Observamos que o feito de que a matriz de φ sexa antisimétrica e de rango 2 non depende da base, xa que se a matriz nunha certa base, B, é antisimétrica e de rango 2, entón noutra base será $A = S^tBS$, con S invertible. Tense entón que $A^t = S^tB^tS = -S^tBS$. Por outro lado, se $\mathcal{B}_v = \{v_1, v_2, v_3, v_4\}$ é unha base e $w = \sum_{i,j=1}^4 a_{ij} v_i \otimes v_j$, entón a matriz de φ é $A = (a_{ij})$.

Para a primeira implicación, se $w = v_1 \wedge v_2 \neq 0$ tense que v_1 e v_2 son linealmente independentes, e ampliándoos a unha base de todo o espazo, a matriz correspondente é

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

que é antisimétrica e de rango 2.

Para a outra implicación, se a matriz de φ é $A = (a_{ij})$, con $a_{ij} = -a_{ji}$, tense que

$$w = \sum_{i<j} (a_{ij} w_i \otimes w_j - a_{ij} w_j \otimes w_i) = \sum_{i,j=1}^{4} a_{ij} w_i \wedge w_j = A(w),$$

polo que $w \in A^2(E)$. Ademais, polo apartado anterior podemos supor que a matriz é da forma

$$\begin{pmatrix} 0 & \lambda & 0 & 0 \\ -\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu \\ 0 & 0 & -\mu & 0 \end{pmatrix}.$$

Se o rango é 2, entón $\lambda = 0$ ou $\mu = 0$, polo que $w = \lambda v_1 \wedge v_2$ ou $\lambda = \mu v_3 \wedge v_4$, polo que en calquera caso é descompoñible.

Produto tensorial de espazos vectoriais.

Problema 5.12. Sexan E e F dous K espazos vectoriais de dimensión finita. Dar un isomorfismo

$$E^* \otimes F \simeq \mathcal{L}(E, F)$$

que non dependa da elección de bases.

Solución. Imos definir o isomorfismo

$$\varphi \colon E^* \otimes F \longrightarrow \mathcal{L}(E, F), \quad u \otimes v \mapsto \psi_{u \otimes v},$$

onde $\psi_{u\otimes v} \colon E \to F$ é a aplicación lineal que envía w a $u(w) \cdot v$. Logo, estendemos esta aplicación por linealidade ao resto de elementos de $E^* \otimes F$.
A aplicación φ está ben definida. En particular,

$$\varphi(u \otimes (v_1 + v_2)) = \psi_{u\otimes(v_1+v_2)} = \psi_{u\otimes v_1} + \psi_{u \otimes v_2} = \varphi(u \otimes v_1) + \varphi(u \otimes v_2),$$

e de xeito similar compróbase que respecta a suma na variable de E^* e o produto por escalar en ambas.
Ademais, φ é inxectiva, xa que $u(w) \cdot v$ para todo w se, e soamente se $v = 0$ ou $u = 0$. Como ambos espazos son de dimensión $\dim E \cdot \dim F$, a inxectividade implica que φ é un isomorfismo.

Problema 5.13. Sexan E_1, E_2, E_3 $\mathbb{R}$-espazos vectoriais de dimensión finita. Dar un isomorfismo

$$\mathcal{L}(E_1 \times E_2 \times E_3, \mathbb{R}) \simeq \mathcal{L}(E_1, \mathcal{L}(E_2 \times E_3, \mathbb{R}))$$

que non dependa da elección de bases.

Solución. Imos definir o isomorfismo

$$\varphi\colon L(E_1 \times E_2 \times E_3, \mathbb{R}) \longrightarrow \mathcal{L}(E_1, \mathcal{L}(E_2 \times E_3), \mathbb{R}),$$

que envía unha aplicación lineal $\psi\colon E_1 \times E_2 \times E_3 \longrightarrow \mathbb{R}$ á aplicación lineal $\varphi(\psi)$

$$\varphi(\psi)\colon E_1 \longrightarrow (E_2 \times E_3 \longrightarrow \mathbb{R}), \quad \varphi(\psi)(a)(b,c) = \psi(a,b,c).$$

A aplicación φ é lineal xa que $\varphi(\psi_1+\psi_2) = \varphi(\psi_1)+\varphi(\psi_2)$ e $\varphi(\lambda\psi) = \lambda\varphi(\psi)$. Por outro lado, é inxectiva xa que $\varphi(\psi) = 0$ se, e soamente se, ψ é igual a cero (xa que senón existirá unha terna (a,b,c) para a cal $\psi(a,b,c) \neq 0$).
Alternativamente, é posible construír explicitamente unha inversa de φ. Sexa $g \in \mathcal{L}(E_1, \mathcal{L}(E_2 \times E_3, \mathbb{R}))$, é dicir, de maneira que para cada $v \in E_1$, $g(v)\colon E_2 \times E_3 \longrightarrow \mathbb{R}$. Definimos entón $\psi(g)\colon E_1 \times E_2 \times E_3 \longrightarrow \mathbb{R}$ como a aplicación que envía (x,y,z) a $\psi(g)(x,y,z) = g(x)(y,z)$. É unha comprobación rutineira probar que é linear e que é a inversa de φ.

Problema 5.14. Probar que en $\mathbb{R}^2 \otimes \mathbb{R}^2$ se cumpre $(1,1)\otimes(1,4)+(1,-2)\otimes(-1,2) = 6e_1 \otimes e_2 + 3e_2 \otimes e_1$, onde $e_1 = (1,0)$ e $e_2 = (0,1)$.

Solución. Temos que

$$\begin{aligned}(1,1)\otimes(1,4)+(1,-2)\otimes(-1,2) &= e_1\otimes e_1 + 4e_1\otimes e_2 + e_2\otimes e_1 + 4e_2\otimes e_2 \\ &\quad - e_1\otimes e_1 + 2e_1\otimes e_2 + 2e_2\otimes e_1 - 4e_2\otimes e_2 \\ &= 6e_1\otimes e_2 + 3e_2\otimes e_1.\end{aligned}$$

Problema 5.15. Sexan $E_1, E_2, F_1, F_2, G_1, G_2$ K-espazos vectoriais de dimensión finita.

(a) Dadas aplicacións lineais $f_i \in \mathcal{L}(E_i, F_i)$, demostrar que existe unha única aplicación lineal $\varphi \in \mathcal{L}(E_1 \otimes E_2, F_1 \otimes F_2)$ tal que

$$\varphi(u_1 \otimes u_2) = f_1(u_1) \otimes f_2(u_2)$$

para todo $u_i \in E_i$.

(b) Demostrar que a aplicación $f_1 \otimes f_2 \mapsto \varphi$ define un isomorfismo

$$\mathcal{L}(E_1, F_1) \otimes \mathcal{L}(E_2, F_2) \simeq \mathcal{L}(E_1 \otimes E_2, F_1 \otimes F_2)$$

que non depende da elección de bases.

(c) Demostrar que, dadas aplicacións $f_i \in \mathcal{L}(E_i, F_i)$ e $g_i \in \mathcal{L}(F_i, H_i)$, tense que

$$(g_1 \otimes g_2) \circ (f_1 \otimes f_2) = (g_1 \circ f_1) \otimes (g_2 \otimes f_2).$$

Solución. (a) A aplicación dada por $\varphi(u_1 \otimes u_2) = f_1(u_1) \otimes f_2(u_2)$ é un elemento de $\mathcal{L}(E_1 \otimes E_2, F_1 \otimes F_2)$, pois é unha aplicación lineal de $E_1 \otimes E_2$ a $F_1 \otimes F_2$. En particular, como esta descrición determina os valores nunha base $e_i \otimes e_j$, ten que ser única.

(b) A linealidade da aplicación é inmediata. Por outro lado, se $\varphi = 0$ temos que $f_1(u_1) \otimes f_2(u_2) = (f_1 \otimes f_2)(u_1 \otimes u_2) = 0$ para todo $u_1 \in E_1$ e $u_2 \in E_2$, de onde se deduce que $f_1 \otimes f_2 = 0$. A dimensión de ambos espazos é o produto de dimensións $\dim(E_1)\dim(E_2)\dim(F_1)\dim(F_2)$, polo que unha vez temos demostrada a inxectividade, sabemos que é un isomorfismo.

(c) É suficiente demostrar a igualdade sobre unha base $e_i \otimes e_j$ de $E_1 \otimes E_2$. Unha comprobación rutineira amosa que ambos lados da igualdade dan $g_1(f_1(e_i)) \otimes g_2(f_2(e_j))$.

Problema 5.16. Sexa $\otimes^r f = f \otimes \cdots \otimes f$ o único endomorfismo tal que

$$\otimes^r f(u_1 \otimes \cdots \otimes u_r) = f(u_1) \otimes \cdots \otimes f(u_r)$$

para todo $u_1, \ldots, u_r \in E$.

(a) Probar que se $\sigma \in \operatorname{End}(\otimes^r E)$ é o endomorfismo asociado a unha permutación $\sigma \in S_r$, entón

$$\sigma \circ (\otimes^r f) = (\otimes^r f) \circ \sigma.$$

(b) Probar que os subespazos $S^r(E)$ e $A^r(E)$ son invariantes por $\otimes^r f$.

(c) Sexa $\wedge^r f \in \operatorname{End}(A^r E)$ a restrición de $\otimes^r f$ ao subespazo alternado. Demostrar que

$$\wedge^r f(u_1 \wedge \ldots \wedge u_r) = f(u_1) \wedge \cdots \wedge f(u_r)$$

para todo $u_1, \ldots, u_r \in E$.

(d) Probar que se o subespazo $F = \langle u_1, \ldots, u_r \rangle \subset E$ é invariante por f, entón

$$\wedge^r f(u_1 \wedge \cdots \wedge u_r) = \det(f|F) u_1 \wedge \cdots \wedge u_r.$$

Solución. (a) Temos que ver que ambos endomorfismos coinciden ao avalialos en elementos da forma $u_1 \otimes \ldots \otimes u_r$; o resultado despois pode estenderse por linealidade. Entón,

$$\begin{aligned}\sigma \circ (\otimes^r f)(v_1, \ldots, v_r) &= \sigma \circ (f(v_1) \otimes \ldots \otimes f(v_r)) \\ &= f(v_{\sigma(1)}) \otimes \ldots \otimes f(v_{\sigma(r)}).\end{aligned}$$

Por outro lado,

$$\begin{aligned}(\otimes^r f) \circ \sigma(v_1, \ldots, v_r) &= (\otimes^r f)(v_{\sigma(1)}, \ldots, v_{\sigma(r)}) \\ &= f(v_{\sigma(1)}) \otimes \ldots \otimes f(v_{\sigma(r)}).\end{aligned}$$

(b) Sexa $v \in S^r(E)$. Hai que demostrar que $\otimes^r f(v) \in S^r(E)$, é dicir, que $\otimes^r f(v) \in S^r(E)$. Como v é simétrico, entón $\sigma \cdot v = v$ para todo σ. Usando o apartado anterior, temos a seguinte cadea de igualdades:

$$\sigma \cdot \otimes^r f(v) = \sigma \otimes (\otimes^r f)(v) = (\otimes^r f)(\sigma \cdot v) = \otimes^r f(v).$$

Isto demostra a invariancia por $S^r(E)$. O caso alternado é análogo.

(c) Temos que $u_1 \wedge \cdots \wedge u_r = \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) u_{\sigma(1)} \otimes \ldots \otimes u_{\sigma(r)}$. Aplicando a linealidade de $\otimes^r$ quédanos que

$$\begin{aligned}\wedge^r f(u_1 \wedge \ldots \wedge u_r) &= \wedge^r f(\sum_{\sigma \in S_r} u_{\sigma(1)} \cdots u_{\sigma(r)}) \\ &= \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) f(u_{\sigma(1)}) \otimes \cdots \otimes f(u_{\sigma(r)}) \\ &= f(u_1) \wedge \cdots \wedge f(u_r).\end{aligned}$$

(d) Escribimos $f(u_i) = \sum_{j=1}^r a_{ji} u_j$. Polas propiedades do produto exterior,

$$f(u_1) \wedge \ldots \wedge f(u_r) = \Big(\sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) \prod_{i=1}^r a_{i\sigma(i)}\Big) u_1 \wedge \ldots \wedge u_r = \det(f|F) u_1 \wedge \ldots \wedge u_r.$$

Problema 5.17. Sexan $f \in \operatorname{End}(\mathbb{R}^2)$ e $g \in \mathbb{R}^3$ dous endomorfismos. Supoñamos que, respecto á base canónica e de $\mathbb{R}^2$ e á base canónica e' de $\mathbb{R}^3$, as matrices de f e g son

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

Achar a matriz de $f \otimes g \in \operatorname{End}(\mathbb{R}^2 \otimes \mathbb{R}^3)$ na base $\{e_1 \otimes e'_1, e_1 \otimes e'_2, \ldots, e_2 \otimes e'_3\}$.

Solución. A matriz é

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & -1 & -1 & -1 \\ 2 & 2 & 2 & -2 & -2 & -2 \\ 3 & 3 & 3 & -3 & -3 & -3 \\ \hline 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 0 & 6 & 6 & 6 \end{array}\right).$$

Problema 5.18. Consideramos dous endomorfismos $f, g \in \operatorname{End}(\mathbb{R}^3)$ que teñen por matriz, na base canónica,

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} -1 & 3 & 0 \\ -3 & 5 & 0 \\ 1 & -1 & 2 \end{pmatrix},$$

respectivamente.

(a) Encontrar a forma reducida de Jordan e unha base de Jordan para $f \otimes f \in \operatorname{End}(\mathbb{R}^3 \otimes \mathbb{R}^3)$.

(b) Encontrar os polinomios mínimos de $f \otimes f$, $f \otimes g$, $g \otimes f$ e $g \otimes g$.

Solución. (a) Comezamos achando a forma de Jordan de A. O polinomio característico é $\operatorname{Char}(A; X) = -(X-2)^3$ e a dimensión do núcleo de $A - 2\mathbb{I}$ é 2. Collemos un vector u_2 que non estea no núcleo, como $u_2 = (1, 0, 0)$, e pomos $u_1 = (A - 2\mathbb{I})u_2 = (-1, 0, -1)$. Finalmente, completamos a base cun vector u_3 que estea no núcleo de $A - 2\mathbb{I}$, como $u_3 = (0, 1, 0)$. Polo tanto, (u_1, u_2, u_3) é unha base de Jordan na que a matriz A ten por matriz

$$J_A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

A matriz de $f \otimes f$ na base dada por $v_i \otimes v_j$ é polo tanto

$$J_A \otimes J_A = \begin{pmatrix} 2J_A & J_A & 0 \\ 0 & 2J_A & 0 \\ 0 & 0 & 2J_A \end{pmatrix}.$$

Trátase dunha matriz 9×9 que é triangular superior e que unicamente ten o número 4 na diagonal principal, polo que o polinomio característico é $\mathrm{Char}(J_A \otimes J_A; X) = -(X-4)^9$. Por outro lado, temos que

$$J_A \otimes J_A - 4\mathbb{I}_9 = \begin{pmatrix} 0 & 2 & 0 & 2 & 1 & 0 & & & \\ 0 & 0 & 0 & 0 & 2 & 0 & & & \\ 0 & 0 & 2 & 0 & 0 & 0 & & & \\ & & & 0 & 2 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & & & & 0 & 2 & 0 \\ & & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 \end{pmatrix}$$

ten rango 4, mentres que

$$(J_A \otimes J_A - 4\mathbb{I}_9)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 8 & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & 0 & 0 & 0 & & & \\ & & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 \end{pmatrix}$$

ten rango 1.

A táboa de Jordan cos vectores por ciclos ten a forma seguinte.

v_3				
v_2	v_5	v_7		
v_1	v_4	v_6	v_8	v_9

Polo tanto, a forma de Jordan consta dun bloque de tamaño 3, dous bloques de tamaño 2 e dous bloques de tamaño 1, todos eles asociados ao valor propio 4.

Para dar unha base de Jordan, collemos $v_3 = u_2 \otimes u_2$, polo que

$$\begin{aligned} v_2 &= (J_A \otimes J_A - 4\mathbb{I}_9)(u_2 \otimes u_2) \\ &= (2u_2 + u_1) \otimes (2u_2 + u_1) - 4u_2 \otimes u_2 \\ &= 2u_1 \otimes u_2 + 2u_2 \otimes u_1 + u_1 \otimes u_1, \end{aligned}$$

e do mesmo xeito, $v_1 = 8u_1 \otimes u_1$. Podemos completar o segundo piso collendo $v_5 = u_2 \otimes u_3$ e $v_7 = u_3 \otimes u_2$, polo que $v_4 = 2u_1 \otimes u_3$ e $v_6 = 2u_3 \otimes u_1$. Finalmente, completamos a base de Jordan collendo $v_8 = u_2 \otimes u_1 - u_1 \otimes u_2$ e $v_9 = u_3 \otimes u_3$.

(b) Vemos que f e g teñen a mesma forma de Jordan, polo que son equivalentes, é dicir, existen bases nas que teñen a mesma representación canónica. Polo tanto, tamén o son $f \otimes f$, $f \otimes g$, $g \otimes f$ e $g \otimes g$, polo que os catro endomorfismos teñen o mesmo polinomio mínimo, que vimos que é $(X-4)^3$.

Bibliografía

[A23] S. Axler, *Linear Algebra Done Right*. Undergraduate Texts in Mathematics, Springer, 2023.

[CL00] M. Castellet; I. Llerena, *Àlgebra lineal i geometria*. Manuals Matemàtiques **1**, Universitat Autònoma de Barcelona, 2000.

[C23a] K. Conrad, *Exterior powers*. Notas dispoñibles en `https://kconrad.math.uconn.edu/blurbs/linmultialg/extmod.pdf`.

[C23b] K. Conrad, *Tensor products*. Notas dispoñibles en `https://kconrad.math.uconn.edu/blurbs/linmultialg/tensorprod.pdf`.

[DC94] M. do Carmo, *Differential Forms and Applications*. Universitext, Springer, 1994.

[G22] R. González Rodríguez, *Álxebra linear. Historia, teoría e práctica*. Universidade de Vigo, 2022.

[H98] E. Hernández, *Álgebra y geometría*. Addison–Wesley/Universidad Autónoma de Madrid, 1998.

[K15] T. Kapitula, *Ordinary Differential Equations and Linear Algebra*. SIAM, 2015.

[LM15] J. Liesen; V. Mehrmann, *Linear algebra*, Springer Undergraduate Mathematics Series, 2015.

[L06] D. Logan, *A First Course in Differential Equations*. Undergraduate Texts in Mathematics, Springer, 2006.

[MS21] L. Merino; E. Santos, *Álgebra lineal con métodos elementales*, Paraninfo, 2021.

[P14] F. Planas, *Àlgebra Multilineal i Geometria. Apunts de classe*. Apuntamentos do curso *Àlgebra Multilineal i Geometria* impartido na Universitat Politècnica de Catalunya, 2014. Dispoñibles en `https://web.mat.upc.edu/francesc.planas/AMG-apunts.pdf`.

[P05] F. Puerta, *Álgebra lineal*. Edicions UPC, 2005.

[Q13] J. Quer, *Àlgebra Lineal FME*. Apuntamentos do curso *Àlgebra Lineal* impartido na Universitat Politècnica de Catalunya, 2013.

[Ro08] S. Roman, *Advanced Linear Algebra*. Graduate Texts in Mathematics, Springer, 2008.