Curs d'Àlgebra Lineal

Rafel Amer Ramon

Francesc Carreras Escobar

Departament de Matemàtica Aplicada II

ETSEIT

Departament de Matemàtica Aplicada II
Escola Tècnica Superior d'Enginyeria
Industrial de Terrassa
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Rafel Amer i Francesc Carreras
Escola Tècnica Superiors d'Enginyeria de Terrassa
Colom, 11
08222 Terrassa

algebra@ruth.upc.es

Introducció

Aquest llibre recull l'experiència docent dels autors durant els darrers anys com a professors d'àlgebra lineal a l'Escola Tècnica Superior d'Enginyers Industrials de Terrassa.

El nou pla d'estudis d'aquest centre, implantat a partir del curs 93–94, ha reduït considerablement el nombre d'hores de classe destinades a aquesta assignatura. Això ha exigit, més que retallar el programa, exposar–lo d'una forma diferent que pot provocar distorsions en l'adquisició d'hàbits científics indispensables per a un tècnic.

Així doncs, hem cregut convenient proporcionar a l'alumne un text ajustat al programa oficial. Pensem que l'ajudarà a assimilar els conceptes i resultats i ens permetrà dedicar les classes més a la discussió d'idees que no pas a la pura transcripció d'apunts.

Hem volgut posar de manifest els aspectes formatius de l'àlgebra lineal i la utilitat de les seves aplicacions pràctiques. El temari inclou, d'una banda, els fonaments: sistemes d'equacions lineals, càlcul matricial, determinants, espais vectorials i transformacions lineals. De l'altra, aplicacions de caire geomètric: producte escalar i nocions derivades, teoria de la diagonalització, tensors i l'estudi analític de les figures de primer i segon grau en el pla i a l'espai tridimensional.

Fem constar que no hem inclòs exercicis. Les referències [1] i [2] contenen col·leccions de problemes, resolts i proposats respectivament, ajustades al nostre temari. En canvi, ens ha semblat oportú afegir, al final, un diccionari amb tots els termes definits en el llibre per facilitar–ne la localització.

Assumim la responsabilitat exclusiva dels errors que, malgrat les revisions successives, romanen encara en el text. Dit això, agraïm els comentaris constructius de diversos companys de la Sec-

ció de Terrassa del Departament de Matemàtica Aplicada II. També volem donar les gràcies a Cardellach Còpies S.A. que s'ha encarregat de l'edició i posterior difusió d'aquest llibre.

Esperem que aquesta obra sigui útil als estudiants. Que els estalviï el temps i l'esforç necessaris per cercar cada tema en una bibliografia heterogènia quant a nomenclatura, notació, contingut i punts de vista. I que, després d'aquesta primera funció, es converteixi en un llibre de prestatge, és a dir, sempre a mà, per trobar amb facilitat aquella idea o aquell procediment que, ai las!, tan sovint s'escapen de la memòria.

Setembre de 2003 — Rafel Amer i Francesc Carreras

Índex

1

Sistemes d'equacions i matrius

Les qüestions plantejades per l'àlgebra lineal i les seves aplicacions es formulen sovint en llenguatge matricial; sovint, també, condueixen a l'estudi de sistemes d'equacions lineals. La primera secció d'aquest capítol recull les definicions bàsiques sobre sistemes i una de les tècniques principals per resoldre'ls: el mètode de Gauss, automatitzat per la regla del pivot, amb la variant anomenada mètode de Gauss-Jordan (l'altre procediment habitual, basat en el càlcul de determinants, s'exposa en el capítol 2). La segona secció es dedica a les nocions fonamentals sobre matrius i càlcul matricial.

1.1 Sistemes d'equacions lineals

Un *sistema d'equacions lineals* és una col·lecció d'equacions de la forma

$$\left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \multicolumn{3}{c}{\dotfill} \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}\right\}$$

on els *coeficients* a_{ij}, els *termes independents* $b_1, b_2, \ldots, b_m$ i els valors que han de prendre les *incògnites* $x_1, x_2, \ldots, x_n$ són nombres reals (o, més en general, complexos). El primer subíndex de cada coeficient indica l'equació on apareix, i el segon la incògnita a la que multiplica, mentre que el subíndex de cada terme independent es refereix a l'equació. El sistema anterior és un *sistema amb m equacions i n incògnites* o, simplement, un sistema $m \times n$.

Un sistema és *compatible* si admet alguna solució (una o més), i és *incompatible* si no n'admet cap. Un sistema compatible és *determinat* si només admet una solució, i és *indeterminat* si en té més d'una (això voldrà dir que en té infinites). *Resoldre* un sistema compatible significa, en el cas determinat, donar el valor únic de cada incògnita; en el cas indeterminat, donar una expressió explícita de les incògnites en funció de certs *paràmetres lliures* (que habitualment coincideixen amb algunes de les mateixes incògnites), de manera que assignant valors arbitraris a aquests paràmetres es trobin totes les solucions.

Dos sistemes amb les mateixes incògnites (però potser amb diferent nombre d'equacions) són *equivalents* si admeten les mateixes solucions (que poden ser cap, una o més d'una).

Proposició 1.1 *(Operacions elementals.) Les operacions següents transformen un sistema en un altre sistema equivalent:*

(a) Sumar o restar una equació a una altra.

(b) Multiplicar o dividir una equació per un nombre diferent de zero.

(c) Intercanviar dues equacions.

(d) Canviar l'ordre de les incògnites en el sistema.

DEMOSTRACIÓ. Tot és de comprovació immediata. □

Mètode de Gauss. Tot sistema d'equacions lineals es pot transformar, mitjançant operacions elementals, en un sistema equivalent *de forma triangular*, és a dir, tal que el primer coeficient no nul de cada equació correspongui a una incògnita posterior a la del primer coeficient no nul de l'equació anterior. El mètode de Gauss s'aplica a cegues, sense saber de quin tipus és el sistema abans de començar els càlculs. En el capítol 3, el teorema de Rouché-Frobenius establirà els criteris per classificar a priori un sistema. D'altra banda, el nombre d'incògnites que queden lliures en resoldre un sistema indeterminat no depèn de la manera d'aplicar el mètode de Gauss (o un altre procediment): això quedarà demostrat també en el capítol 3 i permetrà anomenar-lo el *nombre de graus de llibertat del sistema*. Una sistematització dels càlculs que comporta el mètode de Gauss, força útil a la pràctica, s'exposa tot seguit.

Proposició 1.2 *(Regla del pivot.) Si $a_{ij} \neq 0$ és un coeficient tal que $a_{ih} = 0$ per a tot $h < j$, i es* *canvien els coeficients i el terme independent d'una altra equació, la k-èssima, prenent*

$$a'_{k\ell} = a_{ij}a_{k\ell} - a_{kj}a_{i\ell} \quad \text{per a} \quad \ell = 1, 2, \ldots, n,$$
$$b'_k = a_{ij}b_k - a_{kj}b_i\ ,$$

resulta un sistema equivalent a l'original (es diu que a_{ij} ha estat el pivot d'aquesta transformació). Aplicant reiteradament aquesta idea s'arriba a un sistema de forma triangular.

DEMOSTRACIÓ. Només cal observar que els canvis proposats equivalen a multiplicar l'equació k-èssima per a_{ij} i restar-li l'equació i-èssima multiplicada per a_{kj}, que els zeros obtinguts prèviament en columnes $h < j$ es conserven i que, ara, el coeficient de la incògnita x_j a l'equació k-èssima ja és també zero. □

Mètode de Gauss-Jordan. Transformant amb la regla del pivot no només les equacions posteriors sinó també les precedents, tot sistema d'equacions lineals es converteix en un sistema equivalent *de forma diagonal*, és a dir, on tots els coeficients de les columnes dels pivots (excepte ells mateixos) siguin zero. Això ens estalvia la progressiva substitució de les incògnites ja determinades en les equacions precedents del sistema triangular.

Exemples 1.3 (a) Considerem el sistema

$$\left.\begin{aligned} x + 4y + 3z &= 1 \\ 2x + 5y + 4z &= 4 \\ x - 2y - 2z &= 3 \end{aligned}\right\} .$$

Escrivint només, com és habitual, els coeficients i els termes independents en la disposició files–columnes que més endavant denominarem matriu, i aplicant repetidament operacions elementals seguint el mètode de Gauss, tenim la successió de sistemes equivalents següent:

$$\left(\begin{array}{rrr|r} 1 & 4 & 3 & 1 \\ 2 & 5 & 4 & 4 \\ 1 & -2 & -2 & 3 \end{array}\right) \simeq \left(\begin{array}{rrr|r} 1 & 4 & 3 & 1 \\ 0 & -3 & -2 & 2 \\ 0 & -6 & -5 & 2 \end{array}\right) \simeq \left(\begin{array}{rrr|r} 1 & 4 & 3 & 1 \\ 0 & 3 & 2 & -2 \\ 0 & 0 & 3 & 6 \end{array}\right) .$$

En el darrer sistema resulta, de la tercera equació, $z = 2$; substituint a l'equació anterior, $3y+4 = -2$, és a dir, $y = -2$; i substituint, finalment, a la primera, $x - 8 + 6 = 1$, d'on $x = 3$. Per tant,

$$x = 3\,, \quad y = -2\,, \quad z = 2$$

és la solució (única) del sistema original.

Si volem aplicar el mètode de Gauss-Jordan (fent també ús de la regla del pivot), la successió de sistemes, on marquem els pivots que fem servir, és ara la següent:

$$\left(\begin{array}{ccc|c} \boxed{1} & 4 & 3 & 1 \\ 2 & 5 & 4 & 4 \\ 1 & -2 & -2 & 3 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 4 & 3 & 1 \\ 0 & \boxed{-3} & -2 & 2 \\ 0 & -6 & -5 & 2 \end{array}\right) \simeq$$

$$\left(\begin{array}{ccc|c} -3 & 0 & -1 & -11 \\ 0 & -3 & -2 & 2 \\ 0 & 0 & \boxed{3} & 6 \end{array}\right) \simeq \left(\begin{array}{ccc|c} -9 & 0 & 0 & -27 \\ 0 & -9 & 0 & 18 \\ 0 & 0 & 3 & 6 \end{array}\right).$$

Així obtenim $-9x = -27$, $-9y = 18$, $3z = 6$, d'on $x = 3$, $y = -2$ i $z = 2$.

(b) Resoldrem ara el sistema

$$\left.\begin{aligned} x + y + z &= 0 \\ 2x + 3y - 2z &= 0 \\ x + 2y - 3z &= 0 \end{aligned}\right\} .$$

Aplicant novament el mètode de Gauss tenim

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 3 & -2 & 0 \\ 1 & 2 & -3 & 0 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -4 & 0 \\ 0 & 1 & -4 & 0 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

De la segona equació del sistema final surt $y = 4z$, i substituint a la primera queda $x+4z+z = 0$, d'on $x = -5z$. Les expressions

$$x = -5z\,, \quad y = 4z\,, \quad z = z$$

donen la *solució general* del sistema original, en la que la incògnita z fa el paper simultani de paràmetre lliure: donant valors a z obtenim les infinites solucions del sistema.

Fent servir el mètode de Gauss-Jordan i la regla del pivot com a l'exemple anterior, tenim

$$\left(\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 0 \\ 2 & 3 & -2 & 0 \\ 1 & 2 & -3 & 0 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & \boxed{1} & -4 & 0 \\ 0 & 1 & -4 & 0 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

De les equacions $x+5z = 0$ i $y-4z = 0$ resulta novament $x = -5z$ i $y = 4z$ ($z = z$, paràmetre lliure).

(c) Estudiem finalment el sistema

$$\left.\begin{aligned} x + y + z &= 1 \\ x + 2y + z &= 3 \\ 2x + 3y + 2z &= 3 \end{aligned}\right\} .$$

Aplicant el mètode de Gauss obtenim

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 \\ 2 & 3 & 2 & 3 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \end{array}\right) \simeq \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & -1 \end{array}\right).$$

La darrera equació del sistema final és $0 = -1$, per tant aquest sistema és incompatible i l'original també.

Es diu que un sistema és *homogeni* si

$$b_1 = b_2 = \cdots = b_m = 0\,.$$

Tot sistema homogeni és compatible, ja que admet la solució (anomenada *solució trivial*)

$$x_1 = x_2 = \cdots = x_n = 0\,.$$

Els sistemes homogenis seran els que haurem de resoldre amb més freqüència, i els més interessants són evidentment els indeterminats: l'últim resultat d'aquesta secció estableix una propietat bàsica dels sistemes homogenis.

Teorema 1.4 *Tot sistema homogeni amb més incògnites que equacions és indeterminat.*

DEMOSTRACIÓ. Si $n > m$, el nombre de pivots necessaris per a triangular el sistema serà $r \leq m < n$, per tant, al menys una incògnita passarà a ser paràmetre lliure. □

1.2 Càlcul matricial

Una *matriu de m files i n columnes amb coeficients reals* és una funció

$$A : \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \longrightarrow \mathbb{R},$$

que assigna a cada parell (i, j) un nombre real a_{ij} (la definició per a matrius amb coeficients complexos és anàloga). És habitual representar una matriu A com

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

posant el coeficient a_{ij} a la fila i i la columna j. Es diu que A és una *matriu del tipus* $m \times n$.

Atenent a la definició, cal observar que una matriu A del tipus $m \times n$ i una matriu B del tipus $p \times q$ són *iguals* (escriurem $A = B$) si, i només si, $m = p$, $n = q$ i $a_{ij} = b_{ij}$ per a tot i i tot j. El conjunt de totes les matrius del tipus $m \times n$ amb coeficients reals es representa per $M_{m\times n}(\mathbb{R})$ (i per $M_{m\times n}(\mathbb{C})$ el de les de coeficients complexos). Són especialment freqüents les *matrius columna* (tipus $m \times 1$), les *matrius fila* (tipus $1 \times n$) i les *matrius quadrades* (tipus $n \times n$ o *d'ordre* n).

En cada un dels conjunts $M_{m\times n}(\mathbb{R})$ es defineixen dues operacions internes: la *suma*, on $S = A + B$ és la matriu donada per $s_{ij} = a_{ij} + b_{ij}$, i el *producte per un nombre real*, on $P = tA$ és la matriu donada per $p_{ij} = ta_{ij}$, amb $t \in \mathbb{R}$. Les propietats essencials d'aquestes operacions, totes fàcils de comprovar, són les següents:

(a) Associativa de la suma: $A + (B + C) = (A + B) + C$.

(b) Commutativa de la suma: $A + B = B + A$.

(c) Neutre de la suma: existeix una matriu 0 tal que $A + 0 = A$ per a tota A.

(d) Oposats per a la suma: per a cada matriu A, existeix una matriu $-A$ tal que $A+(-A) = 0$.

(e) Distributiva del producte per a la suma de matrius: $t(A + B) = tA + tB$.

(f) Distributiva del producte per a la suma de nombres: $(t + s)A = tA + sA$.

(g) Associativitat dels productes: $(ts)A = t(sA)$.

Es defineix la *diferència* de dues matrius A i B com $A-B = A+(-B)$. Les propietats d'aquesta operació són les derivades de (a)–(g).

La tercera operació fonamental del càlcul matricial és la *multiplicació*: si $A \in M_{m\times n}(\mathbb{R})$ i $B \in M_{n\times p}(\mathbb{R})$, el seu *producte* és la matriu $P = AB \in M_{m\times p}(\mathbb{R})$ definida per les fórmules

$$p_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}\ .$$

Fem notar que el nombre de columnes de A ha de coincidir amb el de files de B, que la matriu resultant AB té tantes files com A i tantes columnes com B i, finalment, que l'operació consisteix en multiplicar "escalarment" les files de A per les columnes de B.

Exemples 1.5 (a) Tot sistema d'equacions lineals es pot representar matricialment en la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix},$$

és a dir, com $AX = B$.

(b) Si es defineix per a cada cada n la *matriu unitat* I_n com la matriu quadrada $n \times n$ de coeficients

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

és immediat comprovar que $I_m A = A = A I_n$ per a tota $A \in M_{m\times n}(\mathbb{R})$: es diu que les matrius unitat són *semineutres laterals* per al producte.

(c) Considerem les matrius

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{i} \quad C = \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}.$$

Amb elles podem il·lustrar dos “defectes” de la multiplicació de matrius que cal afegir al fet que no sempre és possible multiplicar (de vegades es pot calcular AB però no BA). En primer lloc, $AB \neq BA$, i això ens diu que no es compleix, en general, la propietat commutativa. D'altra banda, $AC = 0$, tot i que $A \neq 0$ i $C \neq 0$; per tant, d'una relació del tipus $AX = AY$ no es pot deduir que $X = Y$, encara que sigui $A \neq 0$ i $A(X - Y) = 0$.

Proposició 1.6 *Quan són factibles els productes que hi intervenen, es compleixen les propietats següents:*

(a) Associativa: $A(BC) = (AB)C$.

(b) Distributives: $A(B + C) = AB + AC, \quad (B + C)A = BA + CA$.

(c) Compatibilitat: $(tA)B = t(AB) = A(tB)$.

DEMOSTRACIÓ. Ens limitarem a justificar la part (a) com a il·lustració. Observem que si A és del tipus $m \times n$, B ha de ser del tipus $n \times p$ i C ha de ser $p \times q$. Si posem $P = BC$, $Q = AP$, $R = AB$ i $S = RC$, caldrà provar que $Q = S$. En primer lloc, tant Q com S són matrius del mateix tipus, $m \times q$. A més, els coeficients corresponents coincideixen:

$$\begin{aligned} q_{ij} &= \sum_{\ell=1}^{n} a_{i\ell} p_{\ell j} = \sum_{\ell=1}^{n} a_{i\ell} \left(\sum_{k=1}^{p} b_{\ell k} c_{kj} \right) = \sum_{\ell=1}^{n} \sum_{k=1}^{p} a_{i\ell} b_{\ell k} c_{kj} \\ &= \sum_{k=1}^{p} \sum_{\ell=1}^{n} a_{i\ell} b_{\ell k} c_{kj} = \sum_{k=1}^{p} \left(\sum_{\ell=1}^{n} a_{i\ell} b_{\ell k} \right) c_{kj} = \sum_{k=1}^{p} r_{ik} c_{kj} = s_{ij}. \quad \square \end{aligned}$$

La *transposada* d'una matriu $A \in M_{m\times n}(\mathbb{R})$ és la matriu $A^t \in M_{n\times m}(\mathbb{R})$ de coeficients $a^t_{ij} = a_{ji}$. És obvi que $(A^t)^t = A$, i el lector pot demostrar, aplicant la tècnica de la proposició anterior, les relacions

$$(A + B)^t = A^t + B^t, \qquad (\lambda A)^t = \lambda A^t \quad \text{i} \quad (AB)^t = B^t A^t.$$

Matrius quadrades

Ens concentrarem ara en el conjunt de les matrius quadrades $M_{n\times n}(\mathbb{R})$. En primer lloc, aquí sempre és possible fer el producte AB, i el resultat és intern. La matriu unitat I_n (que denotarem per I quan no hi hagi confusió sobre n) és l'element neutre —bilateral— del producte: $AI = A = IA$ per a tota A.

Una matriu quadrada A és *simètrica* si $A^t = A$, i *ortogonal* si $A^t A = I$. Les matrius simètriques i les ortogonals són les més interessants en les aplicacions, i seran estudiades amb detall més endavant.

Una matriu quadrada A és *regular* (o *invertible*) si existeix B tal que $AB = I = BA$. De fet, si $AB = I = CA$ tenim que

$$B = IB = CAB = CI = C\,,$$

per tant, si A és regular existeix una única matriu, escrita A^{-1} i denominada *inversa* de A, tal que $AA^{-1} = I = A^{-1}A$. A partir d'això és obvi que A^{-1} és regular i que $(A^{-1})^{-1} = A$. També és immediat provar que si A i B són regulars també ho és AB, i $(AB)^{-1} = B^{-1}A^{-1}$. A més, tota matriu regular A és *simplificable*, en el sentit que de la relació $AB = AC$ (i anàlogament de $BA = CA$) es dedueix que $B = C$, ja que

$$A^{-1}(AB) = A^{-1}(AC)\,, \qquad (A^{-1}A)B = (A^{-1}A)C\,, \qquad IB = IC$$

i per tant $B = C$.

Exemple 1.7 Una matriu del tipus 2×2

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

és regular si, i només si, $ad - bc \neq 0$. En aquest cas, la seva inversa és

$$A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

En efecte: si busquem una matriu B tal que $AB = I$ plantejarem

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} x & z \\ y & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

que condueix a dos sistemes 2×2 amb els mateixos coeficients (els de la matriu A), l'un per a x, y i l'altre per a z, t:

$$\left.\begin{aligned} ax + cy &= 1 \\ bx + dy &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} az + ct &= 0 \\ bz + dt &= 1 \end{aligned}\right\}.$$

Resolent-los per reducció arribem a les relacions

$$(ad-bc)x = d\,, \qquad (ad-bc)y = -b\,, \qquad (ad-bc)z = -c \quad \text{i} \quad (ad-bc)t = a\,,$$

d'on es dedueix que si $ad - bc \neq 0$ existeix una única matriu B que compleix $AB = I$, i és

$$B = \frac{1}{ad-bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

A més, es pot comprovar fàcilment que $BA = I$, és a dir, B és la inversa de A.

En canvi, si $ad - bc = 0$, la matriu

$$B = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

compleix $AB = 0$, i això és suficient per veure que A no té inversa.

El fet que certes matrius no nul·les no tinguin inversa (matrius *singulars*) és el darrer "defecte" que cal fer notar abans de comentar que, en general, el càlcul matricial presenta moltes analogies amb el càlcul ordinari amb nombres reals (i complexos) que el lector ja deu haver notat.

Càlcul de la inversa per Gauss–Jordan. Generalitzant l'exemple anterior, considerem el problema de trobar (si existeix) la inversa d'una matriu quadrada $A \in M_{n\times n}(\mathbb{R})$. Busquem una matriu B tal que $AB = I$; aquesta igualtat condueix a n sistemes $n \times n$, tots ells amb els mateixos coeficients (els de la matriu A) però afectant cada sistema només als elements d'una columna de B —com a incógnites— i als de la corresponent columna de la matriu I —com a termes independents—. Es pot fer un estudi simultani dels n sistemes: només cal aplicar el mètode de Gauss-Jordan, operant amb la regla del pivot sobre les files de la matriu $A \,|\, I$ obtinguda per juxtaposició.

Si hem trobat n pivots, haurem obtingut una matriu $I \mid B$ i tindrem que $AB = I$: en el capítol següent veurem que aquesta relació implica que A és regular i B és la inversa de A. En canvi, si han aparegut menys de n pivots, algun dels sistemes és incompatible i la matriu A no té inversa.

Exemples 1.8 (a) Busquem la inversa de la matriu

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

Aplicant el mètode de Gauss-Jordan a la matriu $A \,|\, I$ tenim

$$\left(\begin{array}{ccc|ccc} \boxed{1} & 2 & 4 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 & 0 & 1 \end{array}\right) \simeq \left(\begin{array}{ccc|ccc} 1 & 2 & 4 & 1 & 0 & 0 \\ 0 & \boxed{-3} & -4 & -2 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\simeq \left(\begin{array}{rrr|rrr} -3 & 0 & -4 & 1 & -2 & 0 \\ 0 & -3 & -4 & -2 & 1 & 0 \\ 0 & 0 & \boxed{-1} & 1 & 1 & -3 \end{array}\right) \simeq \left(\begin{array}{rrr|rrr} 3 & 0 & 0 & 3 & 6 & -12 \\ 0 & 3 & 0 & 6 & 3 & -12 \\ 0 & 0 & -1 & 1 & 1 & -3 \end{array}\right)$$

$$\simeq \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 2 & -4 \\ 0 & 1 & 0 & 2 & 1 & -4 \\ 0 & 0 & 1 & -1 & -1 & 3 \end{array}\right) .$$

Per tant A és regular i

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

(b) Estudiem ara la matriu

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 2 & 3 & 5 \end{pmatrix} .$$

Aplicant el mètode de Gauss-Jordan a la matriu $A \mid I$ tenim

$$\left(\begin{array}{rrr|rrr} \boxed{1} & 2 & 3 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 2 & 3 & 5 & 0 & 0 & 1 \end{array}\right) \simeq \left(\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & \boxed{-1} & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \end{array}\right)$$

$$\simeq \left(\begin{array}{rrr|rrr} -1 & 0 & -1 & 1 & -2 & 0 \\ 0 & -1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{array}\right) ,$$

i la manca d'un tercer pivot (aparició d'una fila de zeros) ens diu que A no és regular.

2

Determinants

El concepte de determinant és un instrument de gran interès teòric i pràctic per a l'àlgebra lineal. Tot i la seva aparent complexitat formal, representa una alternativa als procediments de càlcul basats en el mètode de Gauss, als quals supera sovint en elegància i profunditat.

Farem servir en aquest capítol algunes nocions i resultats sobre permutacions, que hem preferit recollir en un altre lloc (apèndix A). Per fer ben nítida la notació, serà convenient escriure com a_i^j el coeficient situat a la fila i i la columna j d'una matriu genèrica $A \in M_{n\times n}(\mathbb{R})$, i posar, quan sigui necessari especificar les files de A,

$$A = (A_1 \quad A_2 \ \cdots \ A_n)\,.$$

2.1 Definició i propietats

Donada una matriu quadrada $A \in M_{n\times n}(\mathbb{R})$,

$$A = \begin{pmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \dots & \dots & \dots & \dots \\ a_n^1 & a_n^2 & \cdots & a_n^n \end{pmatrix}\,,$$

es defineix el seu *determinant* com el nombre real

$$\det(A) = \begin{vmatrix} a_1^1 & a_1^2 & \cdots & a_1^n \\ a_2^1 & a_2^2 & \cdots & a_2^n \\ \multicolumn{4}{c}{\dotfill} \\ a_n^1 & a_n^2 & \cdots & a_n^n \end{vmatrix} = \sum_{\sigma \in \mathcal{P}_n} \epsilon(\sigma) a_1^{\sigma_1} a_2^{\sigma_2} \cdots a_n^{\sigma_n} ,$$

de manera que $\det(A)$ és la suma de $n!$ termes, cada un dels quals és el producte de n coeficients de la matriu (un de cada fila i, al mateix temps, un de cada columna) amb un signe $+$ o $-$. En els casos $n = 2$ i $n = 3$ fem servir aquesta definició directament; a partir de $n = 4$ cal recórrer a procediments indirectes que deduirem de les propietats del determinant.

Exemples 2.1 (a) Quan $n = 2$ només hi ha dues permutacions, la identitat i la transposició $\tau = (1, 2)$. Per tant, el determinant d'una matriu 2×2 és

$$\begin{vmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{vmatrix} = a_1^1 a_2^2 - a_1^2 a_2^1 .$$

(b) (*Regla de Sarrus.*) Per a $n = 3$ hi ha sis permutacions, tres amb signe positiu,

$$id = (1, 2, 3) , \qquad (2, 3, 1) , \qquad (3, 1, 2),$$

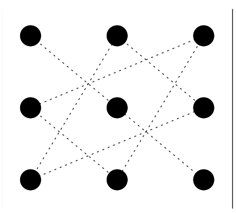
i tres amb signe negatiu,

$$(3, 2, 1) , \qquad (2, 1, 3) , \qquad (1, 3, 2).$$

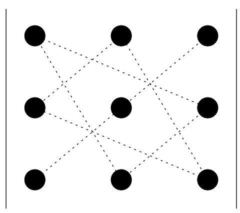
Per tant, el determinant d'una matriu 3×3 és

$$\begin{vmatrix} a_1^1 & a_1^2 & a_1^3 \\ a_2^1 & a_2^2 & a_2^3 \\ a_3^1 & a_3^2 & a_3^3 \end{vmatrix} = a_1^1 a_2^2 a_3^3 + a_1^2 a_2^3 a_3^1 + a_1^3 a_2^1 a_3^2 - a_1^3 a_2^2 a_3^1 - a_1^2 a_2^1 a_3^3 - a_1^1 a_2^3 a_3^2 .$$

Una manera de recordar aquests termes és l'esquema següent:



Signe +



Signe –

(c) Si una matriu A té una fila de zeros, $\det(A) = 0$. En efecte: si la fila i és tota nul·la, el terme corresponent a cada $\sigma \in \mathcal{P}_n$ conté el factor $a_i^{\sigma_i} = 0$, per tant tots els termes són nuls.

(d) Si A és una matriu *triangular inferior*, és a dir, si $a_i^j = 0$ quan $i < j$, el seu determinant és el producte dels elements de la *diagonal principal*:

$$\det(A) = a_1^1 a_2^2 \cdots a_n^n \,.$$

En efecte: per tal que el terme corresponent a una $\sigma \in \mathcal{P}_n$ no contingui cap dels zeros situats sobre la diagonal principal, ha de ser $\sigma_1 = 1$, $\sigma_2 = 2$, i així successivament fins a $\sigma_n = n$; en definitiva, només pot ser el terme de $\sigma = id$, i d'aquí el valor de $\det(A)$. En particular, això també és vàlid si A és una matriu *diagonal*, és a dir, si $a_i^j = 0$ quan $i \neq j$. Per exemple, $\det(I_n) = 1$ per a tot n.

Teorema 2.2 *(a) El determinant és una funció multilineal de les files, és a dir:*

(1) Si una fila es descompon en una suma, $A_i = B_i + C_i$, aleshores

$$\det(A_1 \cdots A_i \cdots A_n) = \det(A_1 \cdots B_i \cdots A_n) + \det(A_1 \cdots C_i \cdots A_n)\,.$$

(2) Si en una fila es treu un factor comú, $A_i = t\,B_i$, aleshores

$$\det(A_1 \cdots A_i \cdots A_n) = t\,\det(A_1 \cdots B_i \cdots A_n)\,.$$

(b) El determinant és una funció alternada de les files, és a dir, si dues files són iguals, $A_i = A_j$, el determinant és zero:

$$\det(A_1 \cdots A_i \cdots A_j \cdots A_n) = 0\,.$$

DEMOSTRACIÓ. (a) Com que $A_i = B_i + C_i$, tenim que $a_i^j = b_i^j + c_i^j$ per a qualsevol $j = 1, 2 \ldots, n$; per tant, $a_i^{\sigma_i} = b_i^{\sigma_i} + c_i^{\sigma_i}$ per a tota permutació σ. Aplicant això a cada terme arribem immediatament a obtenir la propietat (1). El raonament per a la (2) és anàleg.

(b) Si $\tau = (i, j)$ sabem que $\sigma \mapsto \sigma\tau$ transforma biunívocament el conjunt $\mathcal{P}_n^+$ de totes les permutacions de signe positiu en el conjunt $\mathcal{P}_n^-$ de totes les permutacions de signe negatiu. Tenint, a més, en compte que $a_i^k = a_j^k$ per a tot k, resulta:

$$\begin{aligned}
&\det(A_1 \cdots A_i \cdots A_j \cdots A_n) = \\
&= \sum_{\sigma\in\mathcal{P}_n^+} \left[\epsilon(\sigma) a_1^{\sigma_1} \cdots a_i^{\sigma_i} \cdots a_j^{\sigma_j} \cdots a_n^{\sigma_n} + \epsilon(\sigma\tau) a_1^{(\sigma\tau)_1} \cdots a_i^{(\sigma\tau)_i} \cdots a_j^{(\sigma\tau)_j} \cdots a_n^{(\sigma\tau)_n}\right] \\
&= \sum_{\sigma\in\mathcal{P}_n^+} \left[\epsilon(\sigma) a_1^{\sigma_1} \cdots a_i^{\sigma_i} \cdots a_j^{\sigma_j} \cdots a_n^{\sigma_n} - \epsilon(\sigma) a_1^{\sigma_1} \cdots a_i^{\sigma_j} \cdots a_j^{\sigma_i} \cdots a_n^{\sigma_n}\right] \\
&= \sum_{\sigma\in\mathcal{P}_n^+} \epsilon(\sigma) \left[a_1^{\sigma_1} \cdots a_i^{\sigma_i} \cdots a_j^{\sigma_j} \cdots a_n^{\sigma_n} - a_1^{\sigma_1} \cdots a_i^{\sigma_j} \cdots a_j^{\sigma_i} \cdots a_n^{\sigma_n}\right] = 0\,. \quad \square
\end{aligned}$$

Corol·lari 2.3 *(a) Si permutem dues files d'una matriu A, el determinant canvia de signe:*

$$\det(A_1 \cdots A_i \cdots A_j \cdots A_n) = -\det(A_1 \cdots A_j \cdots A_i \cdots A_n)\,.$$

(b) Més en general, per a qualsevol permutació σ,

$$\det(A_{\sigma_1} \quad A_{\sigma_2} \ \cdots \ A_{\sigma_n}) = \epsilon(\sigma)\det(A_1 \quad A_2 \ \cdots \ A_n)\,.$$

(c) Si a una fila li sumem o restem un múltiple d'una altra fila, el determinant no varia:

$$\det(A_1 \cdots A_i + tA_j \cdots A_j \cdots A_n) = \det(A_1 \cdots A_i \cdots A_j \cdots A_n)\,.$$

DEMOSTRACIÓ. (a) Partint de la relació

$$\det(A_1 \cdots A_i + A_j \cdots A_i + A_j \cdots A_n) = 0$$

i aplicant la multilinealitat, tenim que

$$\det(A_1 \cdots A_i \cdots A_i \cdots A_n) + \det(A_1 \cdots A_i \cdots A_j \cdots A_n) +$$

$$+\det(A_1 \cdots A_j \cdots A_i \cdots A_n) + \det(A_1 \cdots A_j \cdots A_j \cdots A_n) = 0\,,$$

és a dir,

$$\det(A_1 \cdots A_i \cdots A_j \cdots A_n) = -\det(A_1 \cdots A_j \cdots A_i \cdots A_n)\,,$$

ja que els altres dos determinants són nuls.

(b) Com que tota permutació és producte de transposicions, $\sigma = \tau_1\tau_2\cdots\tau_p$, i $\epsilon(\sigma) = (-1)^p$, posant $A_\sigma = (A_{\sigma_1} \quad A_{\sigma_2} \ \cdots \ A_{\sigma_n})$ podem escriure

$$\det(A_\sigma) = \det(A_{\tau_1\tau_2\cdots\tau_p}) = (-1)\det(A_{\tau_2\cdots\tau_p}) = \cdots = (-1)^p\det(A) = \epsilon(\sigma)\det(A)\,.$$

(c) Aplicant novament el teorema anterior,

$$\begin{aligned}
&\det(A_1 \cdots A_i + tA_j \cdots A_j \cdots A_n) = \\
&= \det(A_1 \cdots A_i \cdots A_j \cdots A_n) + \det(A_1 \cdots tA_j \cdots A_j \cdots A_n) \\
&= \det(A_1 \cdots A_i \cdots A_j \cdots A_n) + t\det(A_1 \cdots A_j \cdots A_j \cdots A_n) \\
&= \det(A_1 \cdots A_i \cdots A_j \cdots A_n)\,. \quad \square
\end{aligned}$$

Teorema 2.4 $\det(A^t) = \det(A)$. *En conseqüència, totes les propietats del determinant enunciades per files són vàlides també per a columnes.*

DEMOSTRACIÓ. Si a_i^j és el coeficient genèric de A i b_i^j el de A^t, sabem que $b_i^j = a_j^i$. Aleshores

$$\det(A^t) = \sum_{\sigma\in\mathcal{P}_n} \epsilon(\sigma) b_1^{\sigma_1} b_2^{\sigma_2} \cdots b_n^{\sigma_n} .$$

Reordenant els factors de cada un dels termes segons el superíndex tenim que

$$\begin{aligned}\det(A^t) &= \sum_{\sigma\in\mathcal{P}_n} \epsilon(\sigma) b^1_{\sigma^{-1}{}_1} b^2_{\sigma^{-1}{}_2} \cdots b^n_{\sigma^{-1}{}_n} \\ &= \sum_{\sigma\in\mathcal{P}_n} \epsilon(\sigma^{-1}) a_1^{\sigma^{-1}{}_1} a_2^{\sigma^{-1}{}_2} \cdots a_n^{\sigma^{-1}{}_n} ,\end{aligned}$$

ja que $\epsilon(\sigma) = \epsilon(\sigma^{-1})$. Posem $\tau = \sigma^{-1}$. Com que el sumatori anterior consta dels $n!$ termes corresponents a totes les possibles permutacions, podem escriure

$$\det(A^t) = \sum_{\tau\in\mathcal{P}_n} \epsilon(\tau) a_1^{\tau_1} a_2^{\tau_2} \cdots a_n^{\tau_n} = \det(A) . \quad \square$$

Teorema 2.5 *(Fórmula del producte.)* $\det(AB) = \det(A)\det(B)$.

DEMOSTRACIÓ. Si $C = AB = (C_1 \quad C_2 \ \cdots \ C_n)$, cada una d'aquestes files es pot escriure com

$$C_i = \sum_{k=1}^{n} a_i^k B_k ,$$

és a dir, les files de AB són el que s'anomena "combinacions lineals" de les files de B. Representem per $\mathcal{R}_n$ el conjunt de totes les permutacions amb repetició de $1, 2, \ldots, n$. Aleshores

$$\begin{aligned}\det(AB) &= \det(C_1 \quad C_2 \ \cdots \ C_n) \\ &= \det\left(\sum_{k=1}^{n} a_1^k B_k \quad \sum_{k=1}^{n} a_2^k B_k \ \cdots \ \sum_{k=1}^{n} a_n^k B_k\right) .\end{aligned}$$

Aplicant la multilinealitat del determinant queda

$$\begin{aligned}\det(AB) &= \sum_{r\in\mathcal{R}_n} \det(a_1^{r_1} B_{r_1} \quad a_2^{r_2} B_{r_2} \ \cdots \ a_n^{r_n} B_{r_n}) \\ &= \sum_{r\in\mathcal{R}_n} a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n} \det(B_{r_1} \quad B_{r_2} \cdots B_{r_n}) .\end{aligned}$$

Dels determinants que intervenen en aquest sumatori, podem eliminar tots els que tenen alguna fila repetida, i deixar només aquells per als quals r és una permutació sense repeticions, és a dir,

$r \in \mathcal{P}_n$. Per tant,

$$\begin{aligned}\det(AB) &= \sum_{r\in\mathcal{R}_n} a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n} \det(B_{r_1} \quad B_{r_2} \ \cdots \ B_{r_n}) \\ &= \sum_{r\in\mathcal{P}_n} \epsilon(r) a_1^{r_1} a_2^{r_2} \cdots a_n^{r_n} \det(B_1 \quad B_2 \ \cdots \ B_n) = \det(A)\det(B)\,. \quad \square\end{aligned}$$

2.2 Càlcul de determinants

Quan $n \geq 4$, el nombre de termes del sumatori que defineix el determinant és més gran o igual que 24 i creix molt ràpidament, fent impracticable la fórmula original. Estudiarem ara dos procediments de càlcul alternatius.

Mètode de Gauss. Operant per files es pot arribar sempre a una matriu triangular, el determinant de la qual és, segons hem vist, el producte dels coeficients de la diagonal principal. Només cal tenir en compte tres punts:

(a) Intercanviar dues files canvia el signe del determinant.

(b) Cada vegada que fem servir la regla del pivot per transformar una fila, cal dividir el determinant resultant pel pivot utilitzat.

(c) Si apareix una fila de zeros no cal seguir, perquè el determinant val zero.

Exemple 2.6 Càlcul del determinant

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

Marcant els pivots tenim

$$\begin{vmatrix} \boxed{2} & 1 & -3 & 2 \\ 3 & -2 & 1 & 4 \\ 2 & -3 & 1 & -2 \\ 2 & -1 & 0 & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 2 & 1 & -3 & 2 \\ 0 & \boxed{-7} & 11 & 2 \\ 0 & -4 & 4 & -4 \\ 0 & -2 & 3 & -1 \end{vmatrix} = \frac{1}{2}\frac{1}{7^2}\begin{vmatrix} 2 & 1 & -3 & 2 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & \boxed{16} & 36 \\ 0 & 0 & 1 & 11 \end{vmatrix}$$

$$= \frac{1}{2}\frac{1}{7^2}\frac{1}{16}\begin{vmatrix} 2 & 1 & -3 & 2 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & 16 & 36 \\ 0 & 0 & 0 & 140 \end{vmatrix} = \frac{2\cdot(-7)\cdot 16\cdot 140}{2\cdot 7^2\cdot 16} = -20\,.$$

Donada una matriu $A \in M_{n\times n}(\mathbb{R})$ i fixat un coeficient a_i^j, denotarem per A_{i-}^{j-} la submatriu d'ordre $n-1$ que s'obté eliminant la fila i i la columna j de la matriu A. Anomenarem *adjunt* de a_i^j al determinant d'aquesta submatriu multiplicat per $(-1)^{i+j}$:

$$A_i^j = (-1)^{i+j}\det(A_{i-}^{j-})\,.$$

És fàcil recordar el signe corresponent a cada coeficient amb l'ajut de l'esquema següent:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Lema 2.7 *Si tots els coeficients de la fila i diferents de a_i^j són zero, aleshores*

$$\det(A) = a_i^j A_i^j\,.$$

DEMOSTRACIÓ. Fent $i-1$ canvis de fila i $j-1$ canvis de columna podem situar el coeficient a_i^j a la posició $(1,1)$, conservant la fila i, que passa a ser la primera, i la columna j, que també passa a ser-ho. En aquest procés, la fila i i la columna j s'hauran desordenat:

$$A = \begin{pmatrix} a_1^1 & a_1^2 & \ldots & a_1^j & \ldots & a_1^n \\ a_2^1 & a_2^2 & \ldots & a_2^j & \ldots & a_2^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & a_i^j & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_n^1 & a_n^2 & \ldots & a_n^j & \ldots & a_n^n \end{pmatrix}, \qquad B = \begin{pmatrix} a_i^j & 0 & \ldots & 0 & \ldots & 0 \\ a_1^j & a_1^1 & \ldots & a_1^{j-1} & \ldots & a_1^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i-1}^j & a_{i-1}^1 & \ldots & a_{i-1}^{j-1} & \ldots & a_{i-1}^n \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_n^j & a_n^1 & \ldots & a_n^{j-1} & \ldots & a_n^n \end{pmatrix}.$$

Si representem per b_i^j els coeficients de la matriu B, podem escriure

$$\det(A) = (-1)^{i+j}\det(B) = (-1)^{i+j}\sum_{\sigma\in\mathcal{P}_n}\epsilon(\sigma)b_1^{\sigma_1}b_2^{\sigma_2}\cdots b_n^{\sigma_n}\,.$$

El primer coeficient de cada terme és zero, excepte si $\sigma_1 = 1$. Per tant,

$$\det(B) = \sum_{\substack{\sigma\in\mathcal{P}_n\\ \sigma_1=1}}\epsilon(\sigma)b_1^{\sigma_1}b_2^{\sigma_2}\cdots b_n^{\sigma_n} \;=\; b_1^1\sum_{\sigma\in\mathcal{P}_{n-1}}\epsilon(\sigma)b_2^{\sigma_2}\cdots b_n^{\sigma_n}\,,$$

ja que les permutacions $\sigma\in\mathcal{P}_n$ tals que $\sigma_1 = 1$ són les permutacions dels elements $2, 3, \ldots, n$. D'altra banda, $B_{-1}^{-1} = A_{-i}^{-j}$, per tant,

$$\det(A) = (-1)^{i+j}\det(B) = (-1)^{i+j}b_1^1B_1^1 = a_i^jA_i^j\,. \quad \square$$

Teorema 2.8 *(Regla de Laplace.) Fixada una fila i qualsevol d'una matriu A, es compleix que*

$$\det(A) = a_i^1 A_i^1 + a_i^2 A_i^2 + \cdots + a_i^n A_i^n \,.$$

DEMOSTRACIÓ. Expressem la fila i com a suma de n files:

$$A_i = (a_i^1 \; 0 \; \cdots \; 0) + (0 \; a_i^2 \; \ldots \; 0) + \cdots + (0 \; 0 \; \cdots \; a_i^n) = A_{i_1} + A_{i_2} + \cdots + A_{i_n} \,.$$

Aplicant ara la multilinealitat del determinant i el lema anterior resulta que

$$\begin{aligned} \det(A) &= \det(A_1 \; A_2 \; \cdots \; A_i \; \cdots \; A_n) \\ &= \det(A_1 \; \cdots \; A_{i_1} \; \cdots \; A_n) + \det(A_1 \; \cdots \; A_{i_2} \; \cdots \; A_n) + \\ &\qquad + \cdots + \det(A_1 \; \cdots \; A_{i_n} \; \cdots \; A_n) \\ &= a_i^1 A_i^1 + a_i^2 A_i^2 + \cdots + a_i^n A_i^n \,. \quad \square \end{aligned}$$

Observacions 2.9 (a) La regla de Laplace és l'altre procediment habitual per a calcular determinants. Tenint en compte que hi ha una regla de Laplace "anàloga" per columnes, es parla en general de *desenvolupament d'un determinant pels coeficients d'una línia*.

(b) Si es multipliquen els coeficients d'una fila (o columna) pels adjunts d'una altra el resultat és zero, ja que correspon al càlcul d'un determinant amb dues files (o columnes) iguals: si $i \neq k$,

$$a_i^1 A_k^1 + a_i^2 A_k^2 + \cdots + a_i^n A_k^n = 0,$$

i això serà d'utilitat més endavant per justificar la fórmula del càlcul de la inversa.

Exemple 2.10 En el determinant que hem calculat abans pel mètode de Gauss, una vegada obtinguts els zeros de la primera columna podem aplicar la regla de Laplace, que redueix el càlcul a aplicar la regla de Sarrus:

$$\begin{vmatrix} 2 & 1 & -3 & 2 \\ 3 & -2 & 1 & 4 \\ 2 & -3 & 1 & -2 \\ 2 & -1 & 0 & 1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 2 & 1 & -3 & 2 \\ 0 & -7 & 11 & 2 \\ 0 & -4 & 4 & -4 \\ 0 & -2 & 3 & -1 \end{vmatrix} = \frac{1}{2} 2 \begin{vmatrix} -7 & 11 & 2 \\ -4 & 4 & -4 \\ -2 & 3 & -1 \end{vmatrix} = -20 \,.$$

2.3 Matrius regulars i sistemes d'equacions

Dedicarem la primera part d'aquesta secció a caracteritzar les matrius regulars i a donar un procediment per al càlcul de la inversa que és una alternativa al mètode de Gauss-Jordan. A continuació, aplicarem els determinants i les matrius regulars a la resolució dels sistemes d'equacions lineals.

Si $A \in M_{n\times n}(\mathbb{R})$, A^{ad} representarà la *matriu adjunta* de A, que resulta de substituir cada coeficient a_i^j de A pel seu adjunt A_i^j.

Teorema 2.11 *Una matriu A és regular si, i només si,* $\det(A) \neq 0$. *En aquest cas*

$$A^{-1} = \frac{1}{\det(A)}\,(A^{ad})^t\,.$$

DEMOSTRACIÓ. ($\Longrightarrow$) Si A és regular, aplicant la fórmula del producte a la relació $AA^{-1} = I$, resulta que $\det(A)\det(A^{-1}) = 1$. Per tant, $\det(A) \neq 0$ i, a més,

$$\det(A^{-1}) = \frac{1}{\det(A)}\,.$$

($\Longleftarrow$) Aplicant la regla de Laplace i la segona de les observacions que la segueixen, tenim

$$A(A^{ad})^t = \det(A)\, I = (A^{ad})^t\, A\,;$$

per tant, si $\det(A) \neq 0$ podem dividir per $\det(A)$ i apareix, efectivament, la inversa de A que proposa l'enunciat. $\square$

Exemples 2.12 Considerem les dues matrius que hem estudiat al final del capítol anterior.

(a) La matriu

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

és regular perquè $\det(A) = -1$. Aplicant la fórmula de la inversa, tenim que

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

(b) La matriu

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 2 & 3 & 5 \end{pmatrix}$$

no és regular perquè $\det(B) = 0$, és a dir, B no té inversa.

Un sistema d'equacions lineals és un *sistema de Cramer* si té tantes equacions com incògnites i la matriu dels coeficients és regular. Tenint en compte que un sistema d'equacions lineals pot expressar-se matricialment per $AX = B$, on X és la matriu columna de les incògnites i B la dels termes independents, es dedueix que tot sistema de Cramer és compatible determinat, ja que l'única solució és

$$X = A^{-1}B\,.$$

Una forma alternativa d'obtenir aquesta solució és la següent.

Teorema 2.13 *(Regla de Cramer.) Si $AX = B$ és un sistema de Cramer i $A^1, A^2, \ldots, A^n$ són les columnes de la matriu $A \in M_{n\times n}(\mathbb{R})$, la solució del sistema ve donada per les fórmules*

$$x_i = \frac{\det(A^1 \quad A^2 \ \cdots \ B \ \cdots \ A^n)}{\det(A^1 \quad A^2 \ \cdots \ A^i \ \cdots \ A^n)} \qquad \textit{per a} \quad i = 1, 2, \ldots, n.$$

DEMOSTRACIÓ. La solució única del sistema està formada pels coeficients de la relació

$$B = x_1A^1 + x_2A^2 + \cdots + x_nA^n\,.$$

Aleshores, per a cada $i = 1, 2, \ldots, n$ tenim que

$$\begin{aligned} \det(A^1 \quad A^2 \ \cdots \ \overset{i}{\overbrace{B}} \ \cdots \ A^n) &= \sum_{k=1}^{n} \det(A^1 \quad A^2 \ \cdots \ \overset{i}{\overbrace{x_kA^k}} \ \cdots \ A^n) \\ &= x_i \det(A^1 \quad A^2 \ \cdots \ A^i \ \cdots \ A^n)\,, \end{aligned}$$

d'on es dedueix que

$$x_i = \frac{\det(A^1 \quad A^2 \ \cdots \ B \ \cdots \ A^n)}{\det(A^1 \quad A^2 \ \cdots \ A^i \ \cdots \ A^n)}\,. \quad \square$$

Exemples 2.14 (a) Estudiem novament un sistema que en el capítol anterior havíem resolt pel mètode de Gauss:

$$\left.\begin{aligned} x + 4y + 3z &= 1 \\ 2x + 5y + 4z &= 4 \\ x - 2y - 2z &= 3 \end{aligned}\right\}\,.$$

L'escriurem matricialment posant

$$\begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 4 \\ 1 & -2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}.$$

El determinant de la matriu dels coeficients és $\det(A) = 3$, per tant es tracta d'un sistema de Cramer. La seva solució és

$$x = \frac{\begin{vmatrix} 1 & 4 & 3 \\ 4 & 5 & 4 \\ 3 & -2 & -2 \end{vmatrix}}{3} = 3\,, \qquad y = \frac{\begin{vmatrix} 1 & 1 & 3 \\ 2 & 4 & 4 \\ 1 & 3 & -2 \end{vmatrix}}{3} = -2\,, \qquad z = \frac{\begin{vmatrix} 1 & 4 & 1 \\ 2 & 5 & 4 \\ 1 & -2 & 3 \end{vmatrix}}{3} = 2\,.$$

(b) (*Interpolació polinòmica.*) Aquest és un problema amb cert interès pràctic: donats n punts del pla d'abcisses diferents

$$(a_1, b_1),\ (a_2, b_2),\ \ldots,\ (a_n, b_n)\,,$$

es planteja el problema de determinar un polinomi de grau menor o igual que $n-1$ i, per tant, amb n coeficients indeterminats,

$$f(x) = A_0 + A_1 x + A_2 x^2 + \cdots + A_{n-1} x^{n-1},$$

la gràfica del qual passi pels n punts donats. Imposant les condicions obtenim un sistema de n equacions amb n incògnites

$$\left.\begin{array}{rcl} A_0 + a_1 A_1 + a_1^2 A_2 + \cdots + a_1^{n-1} A_{n-1} & = & b_1 \\ A_0 + a_2 A_1 + a_2^2 A_2 + \cdots + a_2^{n-1} A_{n-1} & = & b_2 \\ \multicolumn{3}{c}{\dotfill} \\ A_0 + a_n A_1 + a_n^2 A_2 + \cdots + a_n^{n-1} A_{n-1} & = & b_n \end{array}\right\}.$$

El determinant de la matriu dels coeficients d'aquest sistema és conegut com a *determinant de Vandermonde*, i és funció de $a_1, a_2, \ldots, a_n$:

$$V(a_1, a_2, \ldots, a_n) = \begin{vmatrix} 1 & a_1 & {a_1}^2 & \ldots & {a_1}^{n-1} \\ 1 & a_2 & {a_2}^2 & \ldots & {a_2}^{n-1} \\ \multicolumn{5}{c}{\dotfill} \\ 1 & a_n & {a_n}^2 & \ldots & {a_n}^{n-1} \end{vmatrix}.$$

Demostrem per inducció que aquest determinant és el producte de les diferències entre els a_i, és a dir,

$$V(a_1, a_2, \ldots, a_n) = \prod_{i<j} (a_j - a_i)\,.$$

El cas $n = 2$ és trivial. Considerem el polinomi

$$g(x) = V(a_1, a_2, \ldots, a_{n-1}, x)$$

que, òbviament, s'anul·la per a $x = a_1,\ x = a_2,\ \ldots,\ x = a_{n-1}$ i és de grau $n - 1$, i per tant és de la forma

$$g(x) = \lambda(x - a_1)(x - a_2) \cdots (x - a_{n-1})\,.$$

Com que λ és el coeficient de x^{n-1}, coincideix amb $V(a_1, a_2, \ldots, a_{n-1})$, que per hipòtesi d'inducció és

$$\prod_{i<j<n} (a_j - a_i)\,.$$

Per tant

$$V(a_1, a_2, \ldots, a_n) = g(a_n) = (a_n - a_1)(a_n - a_2) \cdots (a_n - a_{n-1}) \prod_{i<j<n} (a_j - a_i) = \prod_{i<j} (a_j - a_i)\,.$$

L'expressió obtinguda per al determinant de Vandermonde ens diu que $V(a_1, a_2, \ldots, a_n) \neq 0$, de manera que el sistema d'equacions és de Cramer i el problema de la interpolació polinòmica admet sempre una única solució.

Pretenem ara generalitzar la filosofia de la regla de Cramer i resoldre sistemes d'equacions lineals arbitraris. Només necessitem introduir les nocions de menor d'una matriu i inclusió de menors.

Un *menor d'ordre* r d'una matriu $A \in M_{m\times n}(\mathbb{R})$ és el determinant d'una submatriu formada per la intersecció de r files i r columnes de A escollides arbitràriament. (Per abús de llenguatge, s'anomena també menor a la pròpia submatriu.) Direm que un menor M' *conté* a un menor M si les files i columnes que determinen M formen part, respectivament, de les files i columnes que determinen M'.

Si $AX = B$ és un sistema d'equacions lineals $m \times n$ representat matricialment, direm que la matriu $A^a = A\,|\,B$, obtinguda per juxtaposició de la columna dels termes independents, és la *matriu ampliada* del sistema.

Suposem que M és un menor regular màxim de A, és a dir, que no el conté cap altre menor regular de A. En el capítol 3, veurem que si M també és un menor regular màxim de A^a aleshores el sistema és compatible. Suposem–ho així i posem que M sigui d'ordre r. Veurem també que, si $r < m$, totes les equacions que no intervenen a M són redundants i poden, per tant, suprimir–se, de manera que podem suposar $r = m$. Ara, si $r = n$ el sistema és de Cramer i sabem com resoldre'l. Si $r < n$, primer passarem a formar part dels termes independents de les equacions els termes de les incògnites que no intervenen a M, les quals faran de *paràmetres lliures*. Resolent el sistema de Cramer resultant obtindrem les r incògnites principals en funció de les $n - r$ que hem deixat com a paràmetres.

Exemple 2.15 Considerem el sistema

$$\left.\begin{array}{r} 2x+y-z+t=1 \\ x+y-z+2t=3 \\ 3x+2y-z+3t=6 \\ 6x+4y-3z+6t=10 \end{array}\right\} .$$

Preparem simultàniament la matriu dels coeficients i la matriu ampliada:

$$\left(\begin{array}{cccc|c} 2 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 2 & 3 \\ 3 & 2 & -1 & 3 & 6 \\ 6 & 4 & -3 & 6 & 10 \end{array}\right) .$$

Com que el menor

$$M = \begin{vmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \\ 3 & 2 & -1 \end{vmatrix}$$

és diferent de zero i els cinc menors 4×4 de la matriu A^a són nuls, M és un menor regular màxim de A i de A^a, i per tant el sistema és compatible indeterminat. Suprimint la darrera equació i passant a la dreta els termes de l'incògnita t, queda el sistema de Cramer (respecte x, y, z)

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \\ 3 & 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1-t \\ 3-2t \\ 6-3t \end{pmatrix} ,$$

que, resolt per la regla de Cramer o bé invertint la matriu M, ens dóna com a solució general del sistema original, en funció de la incògnita t com a paràmetre lliure,

$$x=-2+t , \qquad y=7-3t , \qquad z=2 , \qquad t=t .$$

3

Espais vectorials

Aquest capítol i el següent —i, de fet, tots els altres— tracten essencialment dels espais numèrics $\mathbb{R}^n$, prototipus dels espais vectorials de dimensió finita, i dels seus subespais. El punt central és el concepte de dimensió, del qual deriva la noció de rang que apareix constantment a la pràctica.

3.1 L'espai vectorial numèric $\mathbb{R}^n$

Recordem que les propietats d'ordenació total i completitud del conjunt $\mathbb{R}$ dels nombres reals permeten representar-los biunívocament com a punts d'una recta —una vegada escollits en ella l'origen i la unitat de longitud—, raó per la qual aquests nombres s'anomenen sovint *escalars*.

El símbol $\mathbb{R}^n$ representarà el producte cartesià $\mathbb{R} \times \mathbb{R} \times \overset{n}{\smile\cdots} \times \mathbb{R}$. Un element genèric d'aquest conjunt, des d'ara un *vector* de $\mathbb{R}^n$, és de la forma $u = (x_1, x_2, \ldots, x_n)$. Els nombres $x_1, x_2, \ldots, x_n$ són les *components* del vector u, i recordem que, si $v = (y_1, y_2, \ldots, y_n)$, es compleix que $u = v$ si, i només si, $x_i = y_i$ per a $i = 1, 2, \ldots, n$.

Es defineixen dues operacions bàsiques. La *suma* de dos vectors $u, v \in \mathbb{R}^n$ és el vector $u + v = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$. El *producte* d'un vector $u \in \mathbb{R}^n$ per un escalar $t \in \mathbb{R}$ és el vector $tu = (tx_1, tx_2, \ldots, tx_n)$. Les propietats fonamentals d'aquestes operacions —totes de fàcil comprovació— són les següents:

(a) Associativa de la suma: $(u + v) + w = u + (v + w)$.

(b) Commutativa de la suma: $u + v = v + u$.

(c) Neutre de la suma: existeix un vector 0 tal que $0 + u = u$ per a tot $u \in \mathbb{R}^n$.

(d) Oposats per la suma: per a cada vector u existeix un vector $-u$ tal que $u + (-u) = 0$.

(e) Distributiva del producte per a la suma de vectors: $t(u + v) = tu + tv$.

(f) Distributiva del producte per a la suma d'escalars: $(t + s)u = tu + su$.

(g) Associativitat dels productes: $t(su) = (ts)u$.

(h) Unitat: $1u = u$.

Es diu que $\mathbb{R}^n$, dotat amb les dues operacions descrites que compleixen aquestes propietats, és un espai vectorial, l'*espai vectorial numèric de dimensió* n. En general:

Definició 3.1 Tot conjunt provist d'una suma interna i un producte per escalars que compleixen les propietats (a)–(h) és un *espai vectorial real*.

Els conjunts de matrius d'un mateix tipus, els $M_{m\times n}(\mathbb{R})$, també són espais vectorials reals. De fet, podem identificar $\mathbb{R}^n$ amb $M_{1\times n}(\mathbb{R})$ o $M_{n\times 1}(\mathbb{R})$, segons escrivim els vectors en fila o en columna.

Si el producte està definit per a tot $t \in \mathbb{C}$ (conjunt dels nombres complexos), es parla d'*espai vectorial complex*: aquest seria, per exemple, el cas de $\mathbb{C}^n$ o $M_{m\times n}(\mathbb{C})$.

Observacions 3.2 (a) Tornant a $\mathbb{R}^n$, la *diferència* de dos vectors u i v es defineix com $u - v = u + (-v)$, i el *quocient* d'un vector u per un escalar $t \neq 0$ com

$$\frac{u}{t} = \frac{1}{t}u\,.$$

Les propietats d'aquestes dues operacions secundàries són les derivades de (a)–(h).

(b) $\mathbb{R}^2$ serà anomenat simplement *el pla*, i $\mathbb{R}^3$ *l'espai*. Aquests dos espais vectorials admeten la representació gràfica habitual, una vegada escollits l'origen, la unitat de longitud i uns eixos perpendiculars entre si. D'aquesta manera, l'espai, per exemple, coincideix amb l'*espai dels vectors lliures* usual en la Física: hom pot identificar els vectors $(1, 0, 0)$, $(0, 1, 0)$ i $(0, 0, 1)$ de $\mathbb{R}^3$ amb $\vec{\imath}$, $\vec{\jmath}$ i $\vec{k}$ respectivament, i escriure

$$u = (x_1, x_2, x_3) = x_1\vec{\imath} + x_2\vec{\jmath} + x_3\vec{k}\,.$$

La suma correspon, geomètricament, a la regla del paral·lelogram, i el producte per escalars a la prolongació proporcional dels vectors en un sentit o l'altre segons el signe del factor escalar.

Un subconjunt $S \subset \mathbb{R}^n$ és un *subespai vectorial* de $\mathbb{R}^n$ si compleix les condicions següents:

(1) $0 \in S$.

(2) $u, v \in S \Longrightarrow u + v \in S$.

(3) $u \in S \Longrightarrow tu \in S$ per a tot $t \in \mathbb{R}$.

La condició (1) expressa que $S \neq \emptyset$. Les altres dues garanteixen que la suma i la multiplicació per escalars restringides a S són operacions internes de S. Com que les propietats (a)–(h) romanen vàlides dins de S, aquest conjunt és, per ell mateix, un espai vectorial real. Per tant, les definicions i els resultats que d'ara endavant s'estableixen per a espais vectorials són aplicables no només a $\mathbb{R}^n$ sinó també als seus subespais.

Exemples 3.3 (a) $\{0\}$ i $\mathbb{R}^n$ són els anomenats *subespais trivials* de $\mathbb{R}^n$.

(b) Cada vector $e \neq 0$ de $\mathbb{R}^n$ defineix el subespai $\langle e \rangle = \{te\,;\ \forall t \in \mathbb{R}\}$, anomenat *recta engendrada* per e. En el cas del pla $\mathbb{R}^2$, aquest és l'únic tipus possible de subespai no trivial.

(c) Tota equació lineal homogènia $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$ defineix un subespai de $\mathbb{R}^n$, format pels vectors $u = (x_1, x_2, \ldots, x_n)$ que la satisfan. Més en general, és igualment fàcil veure que el conjunt de les solucions de qualsevol sistema d'equacions lineals homogènies amb n incògnites és un subespai de $\mathbb{R}^n$. Aquesta és una de les dues maneres habituals de descriure un subespai: donant–ne les *equacions implícites*.

3.2 Dependència lineal

Sigui E un espai vectorial real (per fixar idees, el lector pot suposar que E és $\mathbb{R}^n$, o qualsevol dels seus subespais).

Un vector v és *combinació lineal* dels vectors $u_1, u_2, \ldots, u_m$ si existeixen escalars $t_1, t_2, \ldots, t_m$ tals que

$$t_1u_1 + t_2u_2 + \cdots + t_mu_m = v\,.$$

El vector 0, per exemple, és combinació lineal de qualsevol conjunt de vectors.

Els vectors $u_1, u_2, \ldots, u_m$ són *linealment dependents* si existeix alguna combinació lineal no trivial (és a dir, amb algun coeficient no nul) d'aquests vectors que dóna 0. Per tant, direm que $u_1, u_2, \ldots, u_m$ són *linealment independents* si

$$t_1u_1 + t_2u_2 + \cdots + t_mu_m = 0 \quad \Longrightarrow \quad t_1 = t_2 = \cdots = t_m = 0\,.$$

Observacions 3.4 (a) Estudiar si un vector $v \in \mathbb{R}^n$ és combinació lineal de $u_1, u_2, \ldots, u_m$ condueix a discutir un sistema de n equacions lineals: les incògnites són $t_1, t_2, \ldots, t_m$, els seus coeficients són les components dels vectors $u_1, u_2, \ldots, u_m$ escrites en columna i els termes independents són les components de v. La resposta és afirmativa si, i només si, el sistema és compatible.

(b) Anàlogament, per saber si $u_1, u_2, \ldots, u_m$ són linealment independents es planteja un sistema que coincideix amb l'anterior excepte en els termes independents, que aquí són tots nuls. Ara, els vectors són linealment independents si, i només si, el sistema (homogeni i, per tant, compatible) és determinat.

Proposició 3.5 *Els vectors $u_1, u_2, \ldots, u_m$ són linealment dependents si, i només si, algun d'ells és combinació lineal dels restants.*

DEMOSTRACIÓ. ($\Longrightarrow$) Si són linealment dependents, existirà una *relació de dependència*

$$t_1u_1 + t_2u_2 + \cdots + t_mu_m = 0$$

amb algun coeficient no nul. Sense perdre generalitat podem suposar que $t_1 \neq 0$; aleshores tenim que

$$u_1 = -\frac{t_2}{t_1}\, u_2 - \cdots - \frac{t_m}{t_1}\, u_m \,,$$

i això demostra que u_1 és combinació lineal dels restants.

($\Longleftarrow$) Si, per exemple, $u_1 = s_2u_2 + \cdots + s_mu_m$, apareix la relació

$$u_1 - s_2u_2 - \cdots - s_mu_m = 0\,,$$

on almenys el coeficient de u_1 no és nul. Per tant, $u_1, u_2, \ldots, u_m$ són linealment dependents. $\square$

De manera equivalent, es pot enunciar: *els vectors $u_1, u_2, \ldots, u_m$ són linealment independents si, i només si, cap d'ells és combinació lineal dels altres.*

Proposició 3.6 *Si el vector v és combinació lineal dels vectors $u_1, u_2, \ldots, u_m$, ho és de forma única si, i només si, $u_1, u_2, \ldots, u_m$ són linealment independents.*

DEMOSTRACIÓ. ($\Longrightarrow$) Suposem que v s'expressa de manera única com a combinació lineal dels vectors $u_1, u_2, \ldots, u_m$:

$$v = s_1u_1 + s_2u_2 + \cdots + s_mu_m \,.$$

Aleshores, si

$$t_1u_1 + t_2u_2 + \cdots + t_mu_m = 0\,,$$

sumant les dues igualtats tindrem que

$$v = (t_1 + s_1)u_1 + (t_2 + s_2)u_2 + \cdots + (t_m + s_m)u_m,$$

i de la unicitat de l'expressió de v es dedueix que $t_i + s_i = s_i$ per a tot $i = 1, 2, \ldots, m$. Per tant, $t_1 = t_2 = \cdots = t_m = 0$ i els vectors $u_1, u_2, \ldots, u_m$ són linealment independents.

($\Longleftarrow$) Si

$$v = t_1u_1 + t_2u_2 + \cdots + t_mu_m$$

$$v = s_1u_1 + s_2u_2 + \cdots + s_mu_m$$

són dues expressions de v com a combinació lineal de $u_1, u_2, \ldots, u_m$, en restar-les obtenim

$$(t_1 - s_1)u_1 + (t_2 - s_2)u_2 + \cdots + (t_m - s_m)u_m = 0\,.$$

La independència lineal d'aquests vectors implica que $t_i - s_i = 0$ per a $i = 1, 2, \ldots, m$, és a dir, v s'expressa de manera única com a combinació lineal de $u_1, u_2, \ldots, u_m$. □

Proposició 3.7 *Si u_m és combinació lineal de $u_1, u_2, \ldots, u_{m-1}$, aleshores tot vector que sigui combinació lineal de $u_1, u_2, \ldots, u_m$ també ho és de $u_1, u_2, \ldots, u_{m-1}$.*

DEMOSTRACIÓ. Suposem que $u_m = s_1u_1 + s_2u_2 + \cdots + s_{m-1}u_{m-1}$. Si

$$v = t_1u_1 + t_2u_2 + \cdots + t_mu_m,$$

només cal substituir en aquesta relació l'expressió anterior de u_m per obtenir

$$v = (t_1 + s_1t_m)u_1 + (t_2 + s_2t_m)u_2 + \cdots + (t_{m-1} + s_{m-1}t_m)u_{m-1}\,. \quad \square$$

Un conjunt $\{u_1, u_2, \ldots, u_m\}$ de vectors de E és un *sistema de generadors* de E si tots els vectors d'aquest espai són combinació lineal de $u_1, u_2, \ldots, u_m$.

Una *base* de E és un sistema de generadors linealment independents de E. Si hem escollit una base $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ de E, cada vector $u \in E$ s'escriu de forma única com

$$u = x_1e_1 + x_2e_2 + \cdots + x_ne_n\,.$$

Els coeficients $x_1, x_2, \ldots, x_n$ d'aquesta combinació lineal són les *components de u en la base* $\mathcal{B}$. Habitualment, les bases —i per tant les llistes de components dels vectors— es consideren ordenades.

Exemples 3.8 (a) La *base canònica* de $\mathbb{R}^n$ és la formada pels vectors

$$e_1 = (1, 0, \ldots, 0), \;\; e_2 = (0, 1, \ldots, 0), \;\; \ldots, \;\; e_n = (0, 0, \ldots, 1),$$

i es representa per $\mathcal{B}_c$. Les components d'un vector $u = (x_1, x_2, \ldots, x_n)$ en aquesta base són, precisament, $x_1, x_2, \ldots, x_n$.

(b) Generalitzant la idea de recta engendrada per un vector, obtindrem la segona manera habitual de descriure un subespai. Donats $u_1, u_2, \ldots, u_m \in \mathbb{R}^n$, es defineix el *subespai engendrat* per aquests vectors com el conjunt de totes les seves combinacions lineals, és a dir,

$$\langle u_1, u_2, \ldots, u_m \rangle = \{t_1 u_1 + t_2 u_2 + \cdots + t_m u_m \; ; \; \forall \; t_1, t_2, \ldots, t_m \in \mathbb{R}\}$$

(el lector comprovarà que això és, efectivament, un subespai). És clar que $\{u_1, u_2, \ldots, u_m\}$ és un sistema de generadors d'aquest subespai. Si els vectors són linealment independents, formen una base del subespai; si no ho són, la proposició anterior ens dóna un procediment recursiu per eliminar vectors d'aquest sistema fins a obtenir-ne un subconjunt de generadors linealment independents que, aleshores, serà una base del subespai: només cal suprimir els vectors que siguin combinació lineal dels anteriors.

Teorema 3.9 *Suposem que $u_1, u_2, \ldots, u_k$ són vectors linealment independents de E, i que $e_1, e_2, \ldots, e_m$ formen un sistema de generadors del mateix espai. Aleshores, $k \leq m$.*

DEMOSTRACIÓ. Com que $e_1, e_2, \ldots, e_m$ generen l'espai E, els vectors $u_1, u_2, \ldots, u_k$ són combinació lineal d'ells:

$$\begin{aligned} u_1 &= a_{11}e_1 + a_{21}e_2 + \cdots + a_{m1}e_m \, , \\ u_2 &= a_{12}e_1 + a_{22}e_2 + \cdots + a_{m2}e_m \, , \\ &\vdots \\ u_k &= a_{1k}e_1 + a_{2k}e_2 + \cdots + a_{mk}e_m \, . \end{aligned}$$

En substituir aquestes expressions a la combinació lineal

$$t_1 u_1 + t_2 u_2 + \cdots + t_k u_k$$

i agrupar els coeficients de $e_1, e_2, \ldots, e_m$ obtenim

$$\begin{aligned} &(a_{11}t_1 + a_{12}t_2 + \cdots + a_{1k}t_k)e_1 + \\ &\quad +(a_{21}t_1 + a_{22}t_2 + \cdots + a_{2k}t_k)e_2 + \\ &\quad + \cdots + (a_{m1}t_1 + a_{m2}t_2 + \cdots + a_{mk}t_k)e_m \, . \end{aligned}$$

Aleshores, si $k > m$ el sistema d'equacions

$$\left.\begin{array}{rcl} a_{11}t_1 + a_{12}t_2 + \cdots + a_{1k}t_k & = & 0 \\ a_{21}t_1 + a_{22}t_2 + \cdots + a_{2k}t_k & = & 0 \\ \multicolumn{3}{c}{\dotfill} \\ a_{m1}t_1 + a_{m2}t_2 + \cdots + a_{mk}t_k & = & 0 \end{array}\right\} ,$$

que té m equacions i k incògnites, és indeterminat, és a dir, té solucions diferents de la trivial, les quals compleixen que

$$t_1u_1 + t_2u_2 + \cdots + t_ku_k = 0\,.$$

Això és una contradicció, ja que els vectors $u_1, u_2, \ldots, u_k$ són linealment independents. Per tant, ha de ser $k \leq m$. □

En particular, si $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ és una base de E, k vectors de E amb $k > n$ no poden ser linealment independents, i k vectors de E amb $k < n$ no poden formar sistema de generadors de E. Una consequència fonamental és el resultat següent.

Teorema 3.10 *(Teorema de les bases.) Totes les bases d'un espai vectorial E tenen el mateix nombre de vectors.*

DEMOSTRACIÓ. Siguin $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ i $\mathcal{B}' = \{e'_1, e'_2, \ldots, e'_m\}$ dues bases de E. Si $m < n$, $\mathcal{B}'$ no és un sistema de generadors, i si $m > n$, els vectors de $\mathcal{B}'$ no són linealment independents. Per tant, $m = n$. □

Observacions 3.11 Suposem que $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ és una base de E.

(a) (*Teorema de Steinitz.*) Si $u_1, u_2, \ldots, u_k$ són vectors linealment independents de E, es poden escollir $n - k$ vectors de la base $\mathcal{B}$ que juntament amb $u_1, u_2, \ldots, u_k$ formin una base de E.

El raonament és recurrent, començant pel fet que algun vector de la base $\mathcal{B}$ no és combinació lineal de $u_1, u_2, \ldots, u_k$ i, per tant, es pot afegir a ells sense perdre's la independència lineal.

(b) Si $\{u_1, u_2, \ldots, u_n\}$ són vectors linealment independents de E, aquests vectors formen base de E.

Observem que si no fossin generadors de E existiria algun vector u_{n+1} de E que no seria combinació lineal d'ells, és a dir, tindríem $n + 1$ vectors $u_1, u_2, \ldots, u_n, u_{n+1}$ linealment independents, en contradicció amb l'existència del sistema de generadors $\mathcal{B}$.

(c) Si $\{u_1, u_2, \ldots, u_n\}$ és un sistema de generadors de E, aleshores aquests vectors també formen base de E.

En efecte: si els vectors $u_1, u_2, \ldots, u_n$ no fossin linealment independents, algun d'ells seria combinació lineal dels altres, i eliminant-lo obtindríem un sistema de generadors de E amb $n-1$ vectors. Una altra contradicció amb l'existència de $\mathcal{B}$.

3.3 Dimensió i rang

En aquesta secció retrobarem idees que estaven latents en els dos primers capítols i que ara podem establir amb rigor.

Definició 3.12 La *dimensió* d'un espai vectorial és el nombre de vectors de qualsevol de les seves bases. Evidentment, aquesta definició queda justificada pel teorema de les bases.

El *rang d'un conjunt de vectors* és la dimensió del subespai que generen: simbòlicament

$$\text{rang}\,\{u_1, u_2, \ldots, u_m\} = \dim \langle u_1, u_2, \ldots, u_m \rangle \,.$$

Exemples 3.13 (a) Òbviament, $\dim(\mathbb{R}^n) = n$. I, per conveni, es pren $\dim\{0\} = 0$.

(b) Si $e \neq 0$, $\dim\langle e\rangle = 1$. Si S és el subespai de $\mathbb{R}^3$ definit per l'equació $3x - 2y - z = 0$, tenim que $\dim(S) = 2$, perquè els vectors de S es poden escriure en la forma

$$(x, y, z) = (x, y, 3x - 2y) = x(1, 0, 3) + y(0, 1, -2).$$

En general, una *recta* és un subespai de dimensió 1, un *pla* és un subespai de dimensió 2, i un *hiperpla* de $\mathbb{R}^n$ és un subespai de dimensió $n-1$.

(c) Algunes vegades cal fer servir que si S és un subespai de E, i $\dim(S) = \dim(E)$, aleshores $S = E$.

(d) I sempre és bo tenir present que la dimensió és el nombre de paràmetres lliures que cal donar per determinar un vector de l'espai considerat (*nombre de graus de llibertat*). Així veiem, per exemple, sense necessitat d'explicitar una base, que

$$\dim\, M_{m\times n}(\mathbb{R}) = mn.$$

El *rang d'una matriu* $A \in M_{m\times n}(\mathbb{R})$ és el rang del conjunt de les seves columnes, és a dir, la dimensió del subespai generat per les n columnes de A interpretades com a vectors de $\mathbb{R}^m$.

De vegades es distingeix entre el rang que acabem de definir, que s'anomena *rang per columnes* de A, escrit $\text{rang}_C(A)$, i el concepte anàleg de *rang per files* de A, escrit $\text{rang}_F(A)$, és a dir, la dimensió del subespai generat per les m files de A interpretades com a vectors de $\mathbb{R}^n$. Demostrarem en aquesta mateixa secció que $\text{rang}_C(A) = \text{rang}_F(A)$ per a tota matriu A.

Sigui A una matriu quadrada $n \times n$, i $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ els vectors que, escrits en la base canònica, formen les columnes de A. Recollim en el resultat següent les diverses caracteritzacions de les matrius regulars.

Teorema 3.14 *Són equivalents:*

(a) A *és regular.*

(b) $\det(A) \neq 0$.

(c) La triangulació de A *dóna lloc a* n *pivots.*

(d) $u_1, u_2, \ldots, u_n$ *són linealment independents (és a dir, formen base de* $\mathbb{R}^n$*).*

(e) $\text{rang}(A) = n$.

DEMOSTRACIÓ. (a) i (b) són equivalents segons hem vist en el capítol anterior, i (d) i (e) ho són per la definició de rang d'una matriu. La demostració quedarà completa veient que

$$\text{(b)} \implies \text{(d)} \implies \text{(c)} \implies \text{(b)}\,.$$

(b) $\Longrightarrow$ (d) Si considerem el sistema homogeni que permet estudiar la independència lineal de $u_1, u_2, \ldots, u_n$, la condició $\det(A) \neq 0$ ens diu que es tracta d'un sistema de Cramer i, per tant, compatible determinat: això significa que $u_1, u_2, \ldots, u_n$ són linealment independents.

(d) $\Longrightarrow$ (c) Si els vectors són linealment independents, la triangulació del sistema anterior ha de donar lloc a n pivots, per tal que sigui determinat.

(c) $\Longrightarrow$ (b) La triangulació de la matriu A només modifica $\det(A)$ segons un factor de proporcionalitat no nul; si han aparegut n pivots, el determinant serà diferent de zero. $\square$

A continuació, estudiarem dos procediments per calcular el rang d'una matriu arbitrària $A \in M_{m\times n}(\mathbb{R})$, l'un pel mètode de Gauss i l'altre per determinants.

Proposició 3.15 *El nombre de pivots que apareixen en triangular una matriu* A *coincideix amb el seu rang i, per tant, és independent del camí seguit per triangular-la.*

DEMOSTRACIÓ. Considerem el sistema homogeni que permet estudiar la dependència lineal entre les columnes $u_1, u_2, \ldots, u_n$ de la matriu A. Una vegada triangulat, hauran aparegut, diguem, r pivots. Això significa que les r columnes on figuren aquests pivots són linealment independents, perquè el sistema homogeni corresponent és compatible determinat. I que cada una de les restants és combinació lineal d'elles, perquè el sistema que tradueix aquesta qüestió és compatible. En definitiva, $r = \text{rang}(A)$. $\square$

Lema 3.16 *El rang d'una matriu A és r si, i només si, existeix un menor regular màxim d'ordre r.*

DEMOSTRACIÓ. Posem $r' = \text{rang}(A)$. En primer lloc $r \leq r'$, perquè si hi ha un menor regular d'ordre r les seves subcolumnes, i per tant les columnes senceres que hi intervenen, són linealment independents.

D'altra banda, $r' \leq r$, perquè si hi ha r' columnes linealment independents en podem extreure, aplicant la proposició anterior, un menor regular d'ordre r'. $\square$

Corol·lari 3.17 *Per a qualsevol matriu A es compleix que* $\text{rang}(A) = \text{rang}(A^t)$, *és a dir,* $\text{rang}_C(A) = \text{rang}_F(A)$.

DEMOSTRACIÓ. És immediata a partir de la caracterització del rang donada pel lema anterior i del fet que el determinant de tota matriu coincideix amb el de la seva transposada. $\square$

Proposició 3.18 *El rang d'una matriu A és r si, i només si, existeix algun menor regular d'ordre r i tots els menors d'ordre $r+1$ que el contenen són nuls.*

DEMOSTRACIÓ. ($\Longrightarrow$) Només cal aplicar el lema anterior.

($\Longleftarrow$) Suposem que $u_1, u_2, \ldots, u_r$ són les columnes que defineixen el menor regular. Només cal veure que qualsevol altra columna u_j és combinació lineal d'aquestes. Segons una de les propietats característiques de les matrius regulars, el rang per files de la matriu definida per $u_1, u_2, \ldots, u_r, u_j$ és r. Pel corol.lari anterior, el rang (per columnes) d'aquesta matriu també és r, i això vol dir que u_j és combinació lineal de $u_1, u_2, \ldots, u_r$. $\square$

Teorema 3.19 *(Teorema de Rouché-Frobenius.) (a) El sistema d'equacions lineals* $AX = B$ *és compatible si, i només si,* $\mathrm{rang}(A) = \mathrm{rang}(A^a)$.
(b) Si és compatible, el sistema és determinat si, i només si, $\mathrm{rang}(A) = n$, *on* n *és el nombre d'incògnites del sistema.*

DEMOSTRACIÓ. (a) Si $A^1, A^2, \ldots, A^n$ són les columnes de A, la condició de compatibilitat és equivalent a que

$$B \in \langle A^1, A^2, \ldots, A^n \rangle ,$$

és a dir,

$$\langle A^1, A^2, \ldots, A^n \rangle = \langle A^1, A^2, \ldots, A^n, B \rangle ,$$

i això és equivalent a que $\mathrm{rang}(A) = \mathrm{rang}(A^a)$.

(b) Suposant satisfeta la condició anterior, el fet que el sistema sigui determinat equival a que B sigui combinació lineal única de $A^1, A^2, \ldots, A^n$, i això passa si, i només si, aquests vectors són linealment independents, és a dir, si, i només si, $\mathrm{rang}(A) = n$. $\square$

Observacions 3.20 (a) Cal destacar que l'estudi del rang d'una matriu A ens permet escollir un conjunt màxim de columnes (o files) linealment independents.

(1) Si apliquem el mètode de Gauss, seran aquelles on hagin aparegut els pivots.

(2) Si fem servir determinants, seran aquelles que intervenen en el menor regular màxim. Això justifica el procediment de resolució de sistemes que hem exposat al final del capítol 2.

(b) Quan es tracta de discutir el tipus d'un sistema d'equacions lineals amb coeficients que depenen d'un o més paràmetres, es convenient fer ús del teorema de Rouché–Frobenius i seguir la regla pràctica següent: igualar a 0 un menor d'ordre màxim de la matriu ampliada A^a.

3.4 Subespais vectorials de $\mathbb{R}^n$

Hi ha dues formes bàsiques de descriure un subespai S de $\mathbb{R}^n$: donant un sistema de generadors o bé donant un sistema d'equacions lineals homogènies que caracteritzen els vectors de S, les anomenades *equacions implícites* del subespai. En primer lloc ens ocuparem de veure com es passa d'una forma a l'altra i com es troba una base i la dimensió del subespai.

(a) Donat $S = \langle u_1, u_2, \ldots, u_m \rangle$, el càlcul d'una base de S és equivalent al càlcul del rang de la matriu A formada per les components d'aquests vectors escrites en columna, el qual serà la dimensió del subespai.

Si apliquem el mètode de Gauss triangulant la matriu A, la dimensió de S és el nombre de pivots que en resulten, i els vectors corresponents a les columnes on apareixen aquests pivots formen una base de S.

Si operem amb determinants, la dimensió de S és el tamany del menor regular màxim de A, i els vectors corresponents a les columnes involucrades en aquest menor formen una base de S.

(b) Per obtenir les equacions implícites d'un subespai S a partir d'una base $\{u_1, u_2, \ldots, u_r\}$ de S, hem d'imposar a un vector genèric $v = (x_1, x_2, \ldots, x_n)$ que sigui combinació lineal dels anteriors. A la pràctica, preparem la matriu ampliada A^a que resulta d'afegir a A la columna de les components de v i imposem la condició $\text{rang}(A^a) = r$.

Si apliquem el mètode de Gauss, després de triangular A anul.larem els coeficients de la columna nova que donarien un pivot addicional; si fem servir determinants, una vegada localitzat el menor que indica el rang de A anul.larem els menors d'ordre $r + 1$ que el contenen i es formen afegint part de la columna nova.

(c) Finalment, donades les equacions implícites d'un subespai S, cal resoldre el sistema (pel mètode de Gauss, per determinants o per qualsevol altre procediment) per obtenir una base del subespai. Notem que, en aquest cas, si el rang de la matriu del sistema és r, la dimensió del subespai S és $n - r$, és a dir, el nombre de graus de llibertat del sistema.

Exemples 3.21 (a) Càlcul de la dimensió, una base i les equacions implícites del subespai de $\mathbb{R}^4$, $S = \langle(1, 2, -1, 0), (1, 3, 2, 1), (1, 4, 5, 2), (1, 2, 0, 4)\rangle$.

Triangulem la matriu ampliada

$$\left(\begin{array}{rrrr|c} 1 & 1 & 1 & 1 & x \\ 2 & 3 & 4 & 2 & y \\ -1 & 2 & 5 & 0 & z \\ 0 & 1 & 2 & 4 & t \end{array}\right).$$

Això ens permetrà seleccionar una base d'entre els generadors de S, conèixer la dimensió del subespai i, finalment, obtenir–ne les equacions implícites.

$$\left(\begin{array}{rrrr|c} \boxed{1} & 1 & 1 & 1 & x \\ 2 & 3 & 4 & 2 & y \\ -1 & 2 & 5 & 0 & z \\ 0 & 1 & 2 & 4 & t \end{array}\right) \simeq \left(\begin{array}{rrrr|c} \boxed{1} & 1 & 1 & 1 & x \\ 0 & \boxed{1} & 2 & 0 & -2x + y \\ 0 & 3 & 6 & 1 & x + z \\ 0 & 1 & 2 & 4 & t \end{array}\right)$$

$$\simeq \left(\begin{array}{cccc|c} \boxed{1} & 1 & 1 & 1 & x \\ 0 & \boxed{1} & 2 & 0 & -2x+y \\ 0 & 0 & 0 & \boxed{1} & 7x-3y+z \\ 0 & 0 & 0 & 4 & 2x-y+t \end{array}\right) \simeq \left(\begin{array}{cccc|c} \boxed{1} & 1 & 1 & 1 & x \\ 0 & \boxed{1} & 2 & 0 & -2x+y \\ 0 & 0 & 0 & \boxed{1} & 7x-3y+z \\ 0 & 0 & 0 & 0 & -26x+11y-4z+t \end{array}\right).$$

D'aquí resulta que $\dim(S) = 3$, que els vectors $(1, 2, -1, 0)$, $(1, 3, 2, 1)$ i $(1, 2, 0, 4)$ formen una base de S i que l'equació implícita d'aquest subespai és

$$-26x + 11y - 4z + t = 0\,.$$

(b) Càlcul d'una base del subespai S de $\mathbb{R}^4$ definit per les equacions

$$\left.\begin{array}{r} x+y+z-t=0 \\ x+2y-z+t=0 \\ 3x+5y-z+t=0 \end{array}\right\}.$$

Operarem amb determinants. La matriu del sistema és

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -1 & 0 \\ 1 & 2 & -1 & 1 & 0 \\ 3 & 5 & -1 & 1 & 0 \end{array}\right).$$

Un menor regular màxim és

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix},$$

i això ens diu que el rang del sistema és 2 i, per tant, $\dim(S) = 4 - 2 = 2$. A més, podem suprimir la tercera equació i guardar x i y com a incògnites principals. Queda, doncs, un sistema de Cramer respecte d'aquestes dues incògnites,

$$\left.\begin{array}{l} x+y=-z+t \\ x+2y=z-t \end{array}\right\},$$

la solució general del qual és

$$x = -3z + 3t, \quad y = 2z - 2t, \quad z = z, \quad t = t.$$

El vector

$$u = (-3z + 3t, 2z - 2t, z, t) = z(-3, 2, 1, 0) + t(3, -2, 0, 1)\,,$$

on z i t són els paràmetres lliures, és l'element genèric del subespai S. Els vectors $(-3, 2, 1, 0)$ i $(3, -2, 0, 1)$ formen una base de S.

Recomanem al lector que torni a resoldre aquests dos exemples, operant en el primer cas per determinants i en el segon pel mètode de Gauss.

La resta d'aquesta secció es dedica a l'estudi de les operacions habituals amb subespais de $\mathbb{R}^n$.

La *intersecció* de dos subespais S i T és també un subespai, que es representa per $S \cap T$. Naturalment, la manera més directa de descriure'l és reunir en un sol sistema les equacions implícites de S i les de T.

La *suma* de dos subespais és el subespai representat per $S + T$ i format per totes les sumes de la forma $u + v$, on $u \in S$ i $v \in T$. La manera més directa de descriure'l és reunir en un sol conjunt un sistema de generadors de S i un de T, la qual cosa ens dóna un sistema de generadors de $S + T$. En general, tot i prenent bases de cadascun dels subespais, es pot perdre la independència lineal quan les ajuntem. Això motiva la definició següent.

Es diu que la suma de dos subespais S i T és *directa*, i s'escriu $S \oplus T$, si cada vector de la suma s'escriu de manera única com a suma d'un vector de S i un vector de T. Es fàcil veure que això equival a que

$$u + v = 0,\ u \in S,\ v \in T \quad \Longrightarrow \quad u = v = 0\,.$$

Donades dues bases $\mathcal{B}$ de S i $\mathcal{B}'$ de T, $S + T$ és directa si, i només si, $\mathcal{B} \cup \mathcal{B}'$ és un conjunt linealment independent, i aquesta és, potser, la caracterització més útil a la pràctica. Finalment, la condició de suma directa també és equivalent a que $S \cap T = \{0\}$.

Totes les definicions i resultats anteriors (amb l'única excepció del criteri de la intersecció que acabem d'exposar per a la suma directa) es generalitzen immediatament al cas de més de dos subespais.

Es diu que dos subespais S i T són *suplementaris* si $S \oplus T = \mathbb{R}^n$. Els subespais $\{0\}$ i $\mathbb{R}^n$ són suplementaris. Per a qualsevol altre subespai S és fàcil trobar un suplementari: ampliant una base $\{u_1, u_2, \ldots, u_r\}$ de S fins a obtenir una base $\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ de $\mathbb{R}^n$, és fàcil veure que el subespai $T = \langle u_{r+1}, \ldots, u_n \rangle$ és un suplementari de S.

Teorema 3.22 *(Fórmula de Grassmann.) Siguin S i T subespais de $\mathbb{R}^n$.*

(a) Quan $S \cap T = \{0\}$, $\dim(S \oplus T) = \dim(S) + \dim(T)$.

(b) En general, $\dim(S + T) = \dim(S) + \dim(T) - \dim(S \cap T)$.

DEMOSTRACIÓ. (a) Es dedueix immediatament de la caracterització de suma directa per reunió de bases.

(b) Escollim un suplementari S_1 de $S \cap T$ respecte a S, de manera que $S = S_1 \oplus (S \cap T)$. Anàlogament, escollim T_1 tal que $T = T_1 \oplus (S \cap T)$. Ara no és difícil comprovar que S_1, $S \cap T$

i T_1 formen suma directa i que

$$S_1 \oplus (S \cap T) \oplus T_1 = S + T.$$

Aplicant la fórmula de l'apartat (a) resulta que

$$\begin{aligned}\dim(S + T) &= \dim(S_1) + \dim(S \cap T) + \dim(T_1)\\ &= \dim(S) + \dim(T) - \dim(S \cap T)\,. \quad \square\end{aligned}$$

3.5 Canvis de base

Dedicarem aquesta darrera secció a plantejar i resoldre el problema del canvi de base, i serà suficient il·lustrar–lo a $\mathbb{R}^n$. Considerem dues bases,

$$\mathcal{B} = \{e_1, e_2, \ldots, e_n\} \qquad \text{i} \qquad \mathcal{B}' = \{e_1', e_2', \ldots, e_n'\}\,.$$

Per fixar idees, direm que $\mathcal{B}$ és la *base vella* i $\mathcal{B}'$ és la *base nova*. Cada vector $u \in \mathbb{R}^n$ queda determinat tant per les seves components $(x_1, x_2, \ldots, x_n)$ en la base $\mathcal{B}$, que són els coeficients de la combinació lineal

$$u = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n\,,$$

com per les seves components $(x_1', x_2', \ldots, x_n')$ en la base $\mathcal{B}'$, que són els coeficients de

$$u = x_1' e_1' + x_2' e_2' + \cdots + x_n' e_n'\,.$$

Es tracta de trobar la relació entre unes components i les altres. Per a això, necessitem conèixer com s'expressen els vectors nous en funció dels vells, és a dir, les relacions

$$e_j' = \sum_{i=1}^{n} a_{ij} e_i \quad \text{per a} \quad j = 1, 2, \ldots, n\,.$$

Substituint aquestes relacions i calculant, obtenim

$$\begin{aligned}\sum_{i=1}^{n} x_i e_i = u = \sum_{j=1}^{n} x_j' e_j' = \sum_{j=1}^{n} x_j' \Big(\sum_{i=1}^{n} a_{ij} e_i \Big) =\\ = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_j' e_i = \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} a_{ij} x_j' \Big) e_i\,,\end{aligned}$$

i, comparant els extrems d'aquestes igualtats,

$$x_i = \sum_{j=1}^{n} a_{ij} x_j' \quad \text{per a} \quad i = 1, 2, \ldots, n\,.$$

Aquestes relacions resolen la qüestió, però és convenient expressar–les en forma matricial.

Denotarem el canvi de base que estem considerant per

$$(\mathbb{R}^n, \mathcal{B}') \xrightarrow{\quad C \quad} (\mathbb{R}^n, \mathcal{B}) ,$$

i anomenarem *matriu del canvi de base de $\mathcal{B}'$ a $\mathcal{B}$* a la matriu

$$C = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} ,$$

que ens dóna, per columnes, les components de cadascun dels vectors nous en la base dels vectors vells.

Si X i X' representen, respectivament, les matrius columna formades amb les components de u en les bases $\mathcal{B}$ i $\mathcal{B}'$, tenim, com a expressió matricial de les relacions entre components que hem deduit anteriorment,

$$X = CX' ,$$

que és la *fórmula matricial del canvi de base* i expressa les components velles d'un vector arbitrari en funció de les noves.

Pel teorema de caracterització de les matrius regulars vist a la secció 3 d'aquest capítol, sabem que la matriu C és regular. Per tant, si volem calcular les components noves d'un vector coneixent les velles, farem servir la relació inversa,

$$X' = C^{-1}X ,$$

observant que C^{-1} és precisament la matriu del canvi de base de $\mathcal{B}$ a $\mathcal{B}'$. Resumim tota la informació obtinguda en el resultat següent.

Teorema 3.23 *(a) Si C és la matriu del canvi de base de $\mathcal{B}'$ a $\mathcal{B}$, que descriu per columnes els vectors de la base $\mathcal{B}'$ en funció de la base $\mathcal{B}$, la relació*

$$X = CX'$$

expressa les components d'un vector en la base $\mathcal{B}$ en funció de les seves components en la base $\mathcal{B}'$.

(b) C^{-1} és la matriu del canvi de base invers:

$$(\mathbb{R}^n, \mathcal{B}) \xrightarrow{\quad C^{-1} \quad} (\mathbb{R}^n, \mathcal{B}') , \qquad X' = C^{-1}X . \quad \square$$

Exemples 3.24 (a) Els vectors $e'_1 = (1, 1, 1)$, $e'_2 = (1, 1, 2)$ i $e'_3 = (0, 1, 1)$ formen una base $\mathcal{B}'$ de $\mathbb{R}^3$. Si $\mathcal{B}_c$ denota la base canònica, el canvi de base natural i la seva matriu són

$$(\mathbb{R}^3, \mathcal{B}') \xrightarrow{\quad C \quad} (\mathbb{R}^3, \mathcal{B}_c)\,, \qquad C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

La relació entre les components canòniques (x, y, z) d'un vector arbitrari u de $\mathbb{R}^3$ i les components (x', y', z') del mateix vector u en la base $\mathcal{B}'$ són

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \qquad \text{és a dir,} \qquad \begin{cases} x = x' + y' \\ y = x' + y' + z' \\ z = x' + 2y' + z'\,. \end{cases}$$

La matriu inversa de C és

$$C^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix},$$

expressa els vectors e_1, e_2, e_3 de la base canònica en funció dels de $\mathcal{B}'$,

$$e_1 = e'_1 - e'_3, \qquad e_2 = e'_1 - e'_2 + e'_3, \qquad e_3 = -e'_1 + e'_2\,,$$

i permet invertir la relació entre les components dels vectors:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \text{és a dir,} \qquad \begin{cases} x' = x + y - z \\ y' = -y + z \\ z' = -x + y\,. \end{cases}$$

En el cas que es vulgui transformar l'equació d'un subespai, i conèixer, per exemple, l'equació que tindria en la base nova el pla que en la base canònica es descriu per $2x - y + z = 0$, només cal substituir a l'equació donada les expressions de x, y i z que ens dóna el canvi de base original. Així doncs,

$$2(x' + y') - (x' + y' + z') + (x' + 2y' + z') = 0\,,$$

és a dir, $2x' + 3y' = 0$, és l'equació del pla en la base $\mathcal{B}'$.

(b) En algunes situacions no es coneix d'entrada la relació entre les dues bases, però disposem de la seva relació amb la base canònica. Suposem que $\mathcal{B}'$ i $\mathcal{B}''$ són les bases, i que sabem les matrius de canvi a la base canònica per a cada una:

$$(\mathbb{R}^n, \mathcal{B}') \xrightarrow{\quad P \quad} (\mathbb{R}^n, \mathcal{B}_c) \qquad \text{i} \qquad (\mathbb{R}^n, \mathcal{B}'') \xrightarrow{\quad Q \quad} (\mathbb{R}^n, \mathcal{B}_c)\,.$$

Si volem determinar la matriu del canvi denotat per

$$(\mathbb{R}^n, \mathcal{B}'') \xrightarrow{\quad C \quad} (\mathbb{R}^n, \mathcal{B}')\,,$$

només cal observar que de les relacions $PX' = X = QX''$ es dedueix que

$$X' = P^{-1}QX'' ,$$

per tant la matriu buscada és

$$C = P^{-1}Q .$$

Per exemple, si tenim la base $\mathcal{B}'$ de l'exemple anterior, i la base $\mathcal{B}''$ formada pels vectors

$$e_1'' = (1, 1, 2), \quad e_2'' = (0, 0, 1) \quad \text{i} \quad e_3'' = (-1, 1, 2) ,$$

resulta

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix} , \quad Q = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 2 & 1 & 2 \end{pmatrix} , \quad C = P^{-1}Q = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} .$$

4

L'estructura euclidiana de $\mathbb{R}^n$

El producte escalar és la tercera operació bàsica amb vectors de $\mathbb{R}^n$. D'ell deriven conceptes mètrics com l'ortogonalitat, la norma i l'angle, que obren el camí a múltiples aplicacions geomètriques i físiques de l'àlgebra lineal. Així, el mètode dels mínims quadrats que presentem a la tercera secció no és més que l'aplicació de la idea de projecció ortogonal a l'estudi dels sistemes sobredeterminats.

4.1 El producte escalar

El *producte escalar* de dos vectors $u, v \in \mathbb{R}^n$ és el nombre real

$$u \cdot v = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \, ,$$

suposant que $u = (x_1, x_2, \ldots, x_n)$ i $v = (y_1, y_2, \ldots, y_n)$.

És immediat comprovar que aquesta operació és

(a) definida positiva: $u \cdot u > 0$ per a tot $u \neq 0$;

(b) simètrica (commutativa): $u \cdot v = v \cdot u$;

(c) compatible amb el producte per escalars: $(tu) \cdot v = t(u \cdot v) = u \cdot (tv)$;

(d) distributiva respecte de la suma: $u \cdot (v + w) = u \cdot v + u \cdot w$.

Sempre que, a més de les operacions lineals, es considera el producte escalar, és habitual denominar a $\mathbb{R}^n$ l'*espai euclidià de dimensió* n.

Es diu que dos vectors u, v són *ortogonals* (o *perpendiculars*), i s'escriu $u \perp v$, si $u \cdot v = 0$. De manera similar, es diu que dos subespais S i T són *ortogonals* si $u \perp v = 0$ per a tot $u \in S$ i tot $v \in T$.

Proposició 4.1 *(a) Vectors no nuls ortogonals dos a dos són linealment independents.*
(b) Subespais ortogonals dos a dos formen suma directa.

DEMOSTRACIÓ. La idea és bàsicament la mateixa, per tant provarem només l'apartat (a). Si tenim una relació

$$t_1 u_1 + t_2 u_2 + \cdots + t_k u_k = 0$$

entre vectors no nuls ortogonals dos a dos, fent el producte escalar per qualsevol u_i resulta que

$$t_i (u_i \cdot u_i) = 0$$

i per tant $t_i = 0$, i això ens diu que u_1 , u_2 ... , u_k són linealment independents. □

Teorema 4.2 *(Mètode d'ortogonalització de Gram-Schmidt.) Si* $u_1 , u_2 , \ldots , u_m$ *són vectors linealment independents, existeixen vectors* $v_1 , v_2 , \ldots , v_m$ *ortogonals dos a dos tals que*

$$\langle u_1 , u_2 , \ldots , u_j \rangle = \langle v_1 , v_2 , \ldots , v_j \rangle \qquad \textit{per a} \qquad j = 1, 2, \ldots , m \, .$$

DEMOSTRACIÓ. Prenem $v_1 = u_1$. El segon vector serà de la forma $v_2 = u_2 - t_1 v_1$, i això assegura que $\langle u_1, u_2 \rangle = \langle v_1, v_2 \rangle$; cal, però, imposar que $v_2 \perp v_1$, i aquesta condició determina t_1:

$$v_1 \cdot v_2 = 0 \, , \qquad v_1 \cdot u_2 - t_1 v_1 \cdot v_1 = 0 \, , \qquad t_1 = \frac{v_1 \cdot u_2}{v_1 \cdot v_1} \, .$$

Suposem obtinguts v_1, v_2,..., v_k que satisfan les condicions; prenem ara

$$v_{k+1} = u_{k+1} - (t_1 v_1 + t_2 v_2 + \cdots + t_k v_k) \, ,$$

expressió que garanteix la segona de les igualtats següents:

$$\langle u_1, \ldots, u_k, u_{k+1} \rangle = \langle v_1, \ldots, v_k, u_{k+1} \rangle = \langle v_1, \ldots, v_k, v_{k+1} \rangle \, .$$

L'ortogonalitat de v_{k+1} amb v_i per a $i = 1, 2, \dots, k$ dóna

$$v_i \cdot v_{k+1} = 0\,, \qquad v_i \cdot u_{k+1} - t_i v_i \cdot v_i = 0\,, \qquad t_i = \frac{v_i \cdot u_{k+1}}{v_i \cdot v_i}\,,$$

i aquests coeficients $t_1, t_2, \dots, t_k$ determinen v_{k+1}. □

Corol·lari 4.3 *(a) Tot subespai $S \neq \{0\}$ de $\mathbb{R}^n$ admet bases ortogonals.*
(b) Els vectors ortogonals a (tots els vectors de) S formen un subespai $S^\perp$, anomenat suplementari ortogonal de S perquè compleix que

$$S \oplus S^\perp = \mathbb{R}^n \qquad i \qquad (S^\perp)^\perp = S\,.$$

DEMOSTRACIÓ. Ampliem una base qualsevol $\{ u_1\,, u_2\,, \dots,\, u_r \}$ de S fins a obtenir una base $\{ u_1\,, u_2\,, \dots,\, u_r,\, u_{r+1}, \dots,\, u_n \}$ de $\mathbb{R}^n$. Aplicant el mètode de Gram-Schmidt a aquest conjunt resulta una base ortogonal de $\mathbb{R}^n$: $\{ v_1\,, v_2\,, \dots,\, v_r,\, v_{r+1}, \dots,\, v_n \}$. Com que, segons el teorema,

$$\langle u_1\,, u_2\,, \dots\,, u_r \rangle = \langle v_1\,, v_2\,, \dots\,, v_r \rangle\,,$$

els vectors $v_1\,, v_2\,, \dots\,, v_r$ formen una base ortogonal de S, i això demostra l'apartat (a).

D'altra banda, és fàcil veure que $\langle v_{r+1}, \dots,\, v_n \rangle = S^\perp$, de manera que $S \oplus S^\perp = \mathbb{R}^n$. Finalment, $S \subset (S^\perp)^\perp$, però per dimensions veiem que aquesta inclusió és una igualtat, i així queda provat l'apartat (b). □

Exemples 4.4 (a) Càlcul d'una base ortogonal del subespai $S = \langle u_1, u_2, u_3 \rangle$ de $\mathbb{R}^4$, on

$$u_1 = (1, 1, 0, 1), \quad u_2 = (0, 0, 1, 1) \quad \text{i} \quad u_3 = (2, 0, -1, 0)\,.$$

Aplicant el mètode Gram-Schmidt a $u_1\,, u_2\,, u_3$ tenim:

$$v_1 = u_1 = (1, 1, 0, 1)\,;$$

$$v_2 = u_2 - t v_1\,, \qquad \text{on} \qquad t = \frac{v_1 \cdot u_2}{v_1 \cdot v_1} = \frac{1}{3}\,,$$

i per tant podem prendre $v_2 = u_2 - {}^1\!/\!_3\, v_1 \cong (1, 1, -3, -2)$, ja que els factors lineals no afecten la independència lineal ni l'ortogonalitat; finalment,

$$v_3 = u_3 - (t v_1 + s v_2)\,, \qquad \text{on} \qquad t = \frac{v_1 \cdot u_3}{v_1 \cdot v_1} = \frac{2}{3}\,, \qquad s = \frac{v_2 \cdot u_3}{v_2 \cdot v_2} = \frac{1}{3}\,,$$

i queda $v_3 = u_3 - {}^2\!/\!_3\, v_1 - {}^1\!/\!_3\, v_2 \cong (1, -1, 0, 0)$. En definitiva, els vectors $(1, 1, 0, 1)$, $(1, 1, -3, 2)$ i $(1, -1, 0, 0)$ formen una base ortogonal de S.

(b) Si S és un pla de $\mathbb{R}^3$ d'equació $Ax + By + Cz = 0$, posem $w = (A, B, C)$ i $u = (x, y, z)$. L'equació del pla diu, simplement, que $w \cdot u = 0$, per tant $u \in S$ si, i només si, $u \perp w$, i resulta, doncs, que $S^\perp = \langle w \rangle$.

4.2 Norma, angle i projecció ortogonal

La *norma* (*mòdul* o *longitud*) d'un vector $u \in \mathbb{R}^n$ es defineix com

$$\| u \| = \sqrt{u \cdot u} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}\,,$$

on $u = (x_1, x_2, \ldots, x_n)$.

Proposició 4.5 *La norma compleix les propietats següents:*

(a) $\| u \| > 0$ *si* $u \neq 0$.

(b) $\| tu \| = | t | \| u \|$ *per a tot* $t \in \mathbb{R}$.

(c) $\| u + v \|^2 = \| u \|^2 + \| v \|^2$ *si, i només si,* $u \perp v$ *(teorema de Pitàgores).*

(d) $| u \cdot v | \leq \| u \| \| v \|$ *(desigualtat de Schwarz).*

(e) $\| u + v \| \leq \| u \| + \| v \|$ *(desigualtat triangular).*

DEMOSTRACIÓ. (a) Per definició.

(b) $\| tu \| = \sqrt{(tu) \cdot (tu)} = \sqrt{t^2 u \cdot u} = | t | \sqrt{u \cdot u} = | t | \| u \|$.

(c) En general es compleix que

$$\| u + v \|^2 = (u + v) \cdot (u + v) = \| u \|^2 + \| v \|^2 + 2u \cdot v\,,$$

per tant el teorema de Pitàgores resulta evident.

(d) Si $v = 0$ es produeix la igualtat; suposem que $v \neq 0$ i considerem els vectors

$$u' = tv \quad \text{i} \quad u'' = u - u'\,, \qquad \text{on} \quad t = \frac{u \cdot v}{v \cdot v}\,.$$

Evidentment $u = u' + u''$ i, a més, $u' \perp u''$, ja que

$$u' \cdot u'' = (tv) \cdot (u - tv) = \frac{u \cdot v}{v \cdot v}(u \cdot v) - \frac{(u \cdot v)^2}{(v \cdot v)^2}(v \cdot v) = 0\,;$$

aleshores, aplicant el teorema de Pitàgores tindrem que $\| u \| \geq \| u' \|$, és a dir,

$$\| u \| \geq \frac{| u \cdot v |}{v \cdot v} \| v \| = \frac{| u \cdot v |}{\| v \|} \quad \text{o bé} \quad |u \cdot v| \leq \| u \| \| v \|\,.$$

(e) Aplicant la igualtat que hem fet servir a (c) i la desigualtat de Schwarz tenim que

$$\begin{aligned}\| u+v \|^2 &= \| u \|^2 + 2u \cdot v + \| v \|^2 \\ &\leq \| u \|^2 + 2\| u \| \| v \| + \| v \|^2 = (\| u \| + \| v \|)^2 ,\end{aligned}$$

per tant $\| u+v \| \leq \| u \| + \| v \|$. □

Es diu que un vector u és *unitari* si $\| u \| = 1$. *Normalitzar* un vector $u \neq 0$ és substituir-lo per

$$u' = \frac{u}{\| u \|} ,$$

vector que òbviament és unitari i conserva les relacions d'independència i ortogonalitat de u amb altres vectors. Una base formada per vectors unitaris i ortogonals dos a dos es diu que és una *base ortonormal*. La base canònica de $\mathbb{R}^n$ n'és l'exemple més senzill. De fet, qualsevol subespai de $\mathbb{R}^n$ admet bases ortonormals: només cal fer unitaris els vectors d'una base ortogonal del subespai.

Proposició 4.6 *Si $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ és una base de $\mathbb{R}^n$, i C és la matriu de canvi de base de $\mathcal{B}$ a la base canònica,*

$$\mathcal{B} \text{ és base ortonormal} \iff C \text{ és ortogonal} .$$

DEMOSTRACIÓ. Només cal observar que els coeficients p_{ij} de la matriu C^tC són els productes escalars dels vectors de $\mathcal{B}$ entre ells,

$$p_{ij} = u_i \cdot u_j ,$$

i recordar les nocions de vector unitari i vectors ortogonals. □

Observació 4.7 La comprovació de que una matriu C és ortogonal equival, segons aquest resultat, a veure que

$$\begin{aligned} \sum_{i=1}^{n} {c_{ij}}^2 = 1 \qquad & \text{per a tot} \quad j \qquad \text{(columnes unitàries)} \\ \sum_{i=1}^{n} c_{ij} c_{ik} = 0 \qquad & \text{quan} \quad j \neq k \qquad \text{(columnes ortogonals dos a dos)} . \end{aligned}$$

El lector pot comprovar d'aquesta manera que les matrius següents són ortogonals:

$$\frac{1}{5}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix} , \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} , \qquad \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix} .$$

L'*angle* (no orientat) entre dos vectors no nuls u i v de $\mathbb{R}^n$ es defineix com

$$\alpha(u, v) = \arccos \frac{u \cdot v}{\| u \| \| v \|} \, .$$

Per justificar aquesta definició, observem que, segons la desigualtat de Schwarz,

$$-1 \leq \frac{u \cdot v}{\| u \| \| v \|} \leq 1 \, ,$$

i, com que la funció $\cos : [0, \pi] \longrightarrow [-1, 1]$ és bijectiva per ser decreixent, existeix un únic angle $\alpha \in [0, \pi]$ tal que

$$\cos \alpha = \frac{u \cdot v}{\| u \| \| v \|} \, .$$

Aquest angle α és el que hem pres com a $\alpha(u, v)$. De la definició resulta, doncs, que $\alpha(u, v) = \alpha(v, u)$. A més, és immediat comprovar les propietats següents:

(a) $0 \leq \alpha(u, v) \leq \pi$.

(b) u i v són linealment dependents si, i només si, $\alpha(u, v) = 0$ ó π .

(c) $u \perp v$ si, i només si, $\alpha(u, v) = {}^{\pi}/_{2}$.

(d) Si $t, s \in \mathbb{R}$,

$$\alpha(tu, sv) = \begin{cases} \alpha(u, v) & \text{si} \quad ts > 0 \\ \pi - \alpha(u, v) & \text{si} \quad ts < 0 \, . \end{cases}$$

També recuperem, d'aquesta manera, la *fórmula clàssica del producte escalar*, generalitzada ara a $\mathbb{R}^n$:

$$u \cdot v = \| u \| \| v \| \cos \alpha(u, v)$$

(si $u = 0$ o $v = 0$, $\cos \alpha(u, v)$ no està definit, però es pren $u \cdot v = 0$).

Exemples 4.8 (a) Si $\mathcal{B} = \{ e_1, \ e_2, \ldots, \ e_n \}$ és una base ortonormal de $\mathbb{R}^n$ i

$$u = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n \, ,$$

es compleix que $x_i = u \cdot e_i = \| u \| \cos \alpha_i$, on $\alpha_i = \alpha(u, e_i)$ per a $i = 1, 2, \ldots, n$; els coeficients $\cos \alpha_i$ són els *cosinus directors* de u en la base $\mathcal{B}$, i es pot escriure

$$u = \| u \| \sum_{i=1}^{n} (\cos \alpha_i) e_i \, , \quad \text{amb} \quad \sum_{i=1}^{n} \cos^2 \alpha_i = 1 \, .$$

(b) No sempre és el millor camí fer servir una base ortonormal com a referència. Suposem que volem determinar les longituds de les diagonals d'un paral·lelogram, del qual coneixem les longituds dels costats, 2 i 8, i l'angle agut que formen, 60°. Prenem com a base del pla dos vectors u i v que representen els costats no paral·lels. Aleshores les diagonals corresponen als vectors $u + v$ i $u - v$, i tenint en compte que

$$u \cdot u = \| u \|^2 = 4\,, \quad v \cdot v = \| v \|^2 = 64\,,$$

$$u \cdot v = \| u \| \| v \| \cos \alpha(u, v) = 2 \cdot 8 \cdot \cos 60^\circ = 8\,,$$

resulta

$$\| u + v \|^2 = (u + v) \cdot (u + v) = \| u \|^2 + 2u \cdot v + \| v \|^2 = 84 \qquad \text{i}$$

$$\| u - v \|^2 = (u - v) \cdot (u - v) = \| u \|^2 - 2u \cdot v + \| v \|^2 = 52\,,$$

per tant les longituds de les diagonals són $\sqrt{84}$ i $\sqrt{52}$.

Teorema 4.9 *(Teorema de la projecció ortogonal.) Donat un subespai H de $\mathbb{R}^n$, per a cada vector $u \in \mathbb{R}^n$ existeix un únic vector $u' \in H$ tal que $u - u' \perp H$. La norma $\| u - u' \|$ és la mínima distància entre u i els vectors de H. Es diu que u' és la projecció ortogonal de u sobre H.*

DEMOSTRACIÓ. Escollida una base ortogonal $\mathcal{B} = \{ e_1, e_2, \dots, e_r \}$ de H, i donat el vector $u \in \mathbb{R}^n$, es tracta de determinar un vector $u' = t_1 e_1 + t_2 e_2 + \cdots + t_r e_r$. La condició $u - u' \perp H$ es tradueix en que $(u - u') \cdot e_i = 0$ per a $i = 1, 2, \dots, r$, i d'aquí obtenim que

$$t_i = \frac{u \cdot e_i}{e_i \cdot e_i} \qquad \text{per a} \quad i = 1, 2, \dots, r\,.$$

Això demostra l'existència i unicitat del vector u'. Ara, si $v' \in H$, i $v' \neq u'$, tenim que

$$u - v' = (u - u') + (u' - v')\,;$$

aplicant el teorema de Pitàgores,

$$\| u - v' \|^2 = \| u - u' \|^2 + \| u' - v' \|^2 > \| u - u' \|^2\,,$$

i per tant $\| u - v' \| > \| u - u' \|$. □

Observacions 4.10 (a) El fet que $\| u - u' \|$ sigui mínima s'expressa dient que u' és *la millor aproximació a u dins de H*, i aquesta és la propietat més útil de la projecció ortogonal.

(b) El vector $u'' = u - u'$ és la projecció ortogonal de u sobre el subespai $H^\perp$. En algun cas és més fàcil calcular primer u'' i obtenir després $u' = u - u''$. La relació $u = u' + u''$ és la descomposició única de u segons la suma directa $H \oplus H^\perp = \mathbb{R}^n$.

(c) El *simètric de u respecte a H* és el vector

$$u^* = u' - u'' = u - 2u'' = 2u' - u\,.$$

És immediat comprovar que $\| u^* \| = \| u \|$, i que $u + u^* = 2u' \in H$.

Exemples 4.11 (a) La projecció ortogonal d'un vector $u \in \mathbb{R}^n$ sobre una recta $H = \langle w \rangle$ és, segons el teorema,

$$u' = \frac{u \cdot w}{w \cdot w} w .$$

(b) Si H és un pla de $\mathbb{R}^3$, definit per una equació $Ax + By + Cz = 0$ en la base canònica —o en qualsevol base ortonormal—, sabem que $H^\perp = \langle w \rangle$, on $w = (A, B, C)$; per tant, donat $u \in \mathbb{R}^3$ tenim

$$u'' = \frac{u \cdot w}{w \cdot w} w , \qquad u' = u - u'' = u - \frac{u \cdot w}{w \cdot w} w .$$

(c) El simètric d'un vector $u \in \mathbb{R}^n$ respecte al subespai $\{0\}$ és $u^* = -u$. El simètric de u respecte a una recta $H = \langle w \rangle$ és

$$u^* = 2 \frac{u \cdot w}{w \cdot w} w - u .$$

Finalment, el simètric d'un vector $u \in \mathbb{R}^3$ respecte a un pla $H = \langle w \rangle^\perp$ és

$$u^* = u - 2 \frac{u \cdot w}{w \cdot w} w .$$

4.3 El mètode dels mínims quadrats

El *criteri dels mínims quadrats* és un dels més habituals a nivell científic i tècnic. Com a exemple elemental d'aplicació d'aquest principi, revisarem en primer lloc la noció de mitjana aritmètica. Considerem una sèrie de dades $x_1, x_2, \ldots, x_n$ i interpretem–les com els resultats de n mesuraments d'una magnitud X. Qualsevol quantitat x pot ser un valor de centralització de la sèrie de dades, sempre que satisfaci alguna condició de "proximitat" a totes elles com a conjunt. Si prenem com a mesura de desviació l'expressió

$$\delta(x) = (x - x_1)^2 + (x - x_2)^2 + \cdots + (x - x_n)^2 ,$$

obtindrem el valor x que minimitza aquesta funció derivant i igualant a zero:

$$\delta'(x) = 2(x - x_1) + 2(x - x_2) + \cdots + 2(x - x_n) = 0 .$$

La solució única d'aquesta equació és

$$x = \overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} ,$$

expressió que coneixem com a *mitjana aritmètica* de les dades $x_1, x_2, \ldots, x_n$. La funció δ presenta efectivament un mínim absolut en aquest punt, perquè la seva gràfica és una paràbola i el coeficient principal és positiu.

Per tant, la mitjana aritmètica és el valor de centralització que minimitza la suma dels quadrats de les desviacions de les dades. Es pot arribar al mateix resultat aplicant la teoria de la projecció ortogonal i projectant el vector $x = (x_1, x_2, \ldots, x_n)$ sobre la recta engendrada pel vector $w = (1, 1, \ldots, 1)$: el resultat és el vector $\overline{x}w$, on $\overline{x}$ és la mitjana aritmètica.

Generalitzant aquesta idea, considerem ara un sistema de m equacions lineals amb n incògnites

$$\left.\begin{aligned} c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n &= d_1 \\ c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n &= d_2 \\ &\vdots \\ c_{m1}x_1 + c_{m2}x_2 + \cdots + c_{mn}x_n &= d_m \end{aligned}\right\} , \quad \text{o matricialment} \quad Cx = d\,.$$

El sistema és compatible si, i només si, existeix algun vector $x \in \mathbb{R}^n$ tal que $\| Cx - d \| = 0$. Si és incompatible (de vegades és diu que és un *sistema sobredeterminat*), el criteri dels mínims quadrats indica que cal considerar com a *solució aproximada* qualsevol vector x que faci mínim l'*error quadràtic*, és a dir, $\varepsilon = \| Cx - d \|$.

Si $c_1, c_2, \ldots, c_n$ són les columnes de la matriu C, el sistema és equivalent a la relació

$$x_1c_1 + x_2c_2 + \cdots + x_nc_n = d\,.$$

La incompatibilitat equival a dir que $d \notin H = \langle c_1, c_2, \ldots, c_n \rangle$, i per tant cal trobar la projecció ortogonal $d\,'$ de d sobre H i després resoldre

$$x_1c_1 + x_2c_2 + \cdots + x_nc_n = d\,'$$

(aquest enfocament inclou, i per tant generalitza, el cas de compatibilitat, en el qual $d\,' = d$). Això resol conceptualment el problema, però és útil i senzill sistematitzar els càlculs necessaris.

Teorema 4.12 *En les condicions anteriors:*

(a) $Cx = d\,'$ *si, i només si,* $Ax = b$*, on* $A = C^tC$ *i* $b = C^td$.

(b) x *és únic si, i només si,* $\mathrm{rang}(C) = n$*; en aquest cas,* A *és regular.*

DEMOSTRACIÓ. (a) $Cx = d\,'$ significa que $Cx - d \perp H$. Per tant, el producte escalar de cada generador de H (columna c_i de la matriu C) per $Cx - d$ ha de ser zero. Matricialment, l'anul·lació d'aquests productes escalars s'expressa per

$$C^t(Cx - d) = 0\,, \qquad C^tCx = C^td\,, \qquad Ax = b\,.$$

(b) Segons el teorema de Rouché-Frobenius, $\mathrm{rang}(C) = n$ és equivalent a que el sistema $Cx = d\,'$, que sabem que és compatible, sigui determinat. I això és equivalent a dir que la matriu quadrada A sigui regular. $\square$

Exemples 4.13 (a) (*Regressió lineal*.) El problema pràctic potser més freqüent és el de trobar l'equació de la recta que s'ajusta millor a una distribució de punts del pla (*recta de regressió*). Sense obtenir fórmules sistemàtiques (tasca més pròpia d'un curs d'estadística), resoldrem un exemple concret aplicant el teorema anterior.

Suposem que es tracta d'ajustar una recta als punts de coordenades $(0,0)$, $(2,1)$ i $(4,1)$. La recta serà $y = ax + b$, i les condicions

$$\left.\begin{aligned} b &= 0 \\ 2a + b &= 1 \\ 4a + b &= 1 \end{aligned}\right\}, \quad \text{o bé} \quad Cx = d\,, \quad \text{on} \quad C = \begin{pmatrix} 0 & 1 \\ 2 & 1 \\ 4 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{i} \quad d = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

El sistema és òbviament incompatible, el rang és 2 i, multiplicant per la matriu C^t, obtenim

$$\begin{pmatrix} 20 & 6 \\ 6 & 3 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}, \qquad \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{24}\begin{pmatrix} 3 & -6 \\ -6 & 20 \end{pmatrix}\begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/6 \end{pmatrix}.$$

La recta buscada és

$$y = \frac{x}{4} + \frac{1}{6}\,,$$

que no passa per cap dels tres punts però minimitza l'error quadràtic ε: concretament, $x = (a,b) = (1/4,\ 1/6)$, $Cx = (1/6,\ 2/3,\ 7/6)$ i $d = (0,1,1)$, per tant

$$\varepsilon = \|\, Cx - d \,\| = \frac{\sqrt{6}}{6}\,.$$

(Si el lector ha dibuixat els tres punts i cau en la temptació de creure més adequada la recta horitzontal $y = 1$, observarà que l'error quadràtic corresponent a aquesta recta és $1 > \sqrt{6}/6$.)

(b) (*Regressió quadràtica*.) Un problema anàleg a l'anterior és el de trobar l'equació de la paràbola que s'ajusta millor a una distribució de punts del pla (l'elecció del tipus de funció en un problema de regressió depèn de diversos factors, un dels quals pot ser la disposició geomètrica del punts). Com a il·lustració, determinarem el polinomi de regressió quadràtica dels punts $(-1,2)$, $(0,1)$, $(1,-1)$ i $(2,0)$. L'equació serà $y = ax^2 + bx + c$, i les condicions

$$\left.\begin{aligned} a + b - c &= 2 \\ c &= 1 \\ a + b + c &= -1 \\ 4a + 2b + c &= 0 \end{aligned}\right\}, \qquad \text{o bé} \quad Cx = d\,,$$

on

$$C = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad \text{i} \quad d = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

El sistema és incompatible, el rang de C és 3 i, aplicant novament el procediment indicat en el teorema anterior, obtenim el sistema

$$\begin{pmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} ,$$

la solució del qual és $a = 1/2$, $b = -13/10$ i $c = 2/5$. Per tant, el polinomi de regressió quadràtica dels quatre punts donats és

$$y = \frac{1}{10}(5x^2 - 13x + 4) ,$$

i l'error quadràtic comès $\varepsilon = \dfrac{2\sqrt{5}}{5}$.

(c) Considerem finalment el sistema següent, en el qual $\operatorname{rang}(C) = 2$ i $n = 3$:

$$\left.\begin{aligned} x + 2y + z &= 3 \\ x + y + 4z &= 2 \\ 2x + 3y + 5z &= 5 \\ 4x + 6y + 10z &= 0 \end{aligned}\right\} , \qquad \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 4 \\ 2 & 3 & 5 \\ 4 & 6 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 0 \end{pmatrix} .$$

Multiplicant per la transposada resulta el sistema

$$\begin{pmatrix} 22 & 33 & 55 \\ 33 & 50 & 81 \\ 55 & 81 & 142 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 15 \\ 23 \\ 36 \end{pmatrix} ,$$

que és compatible indeterminat i té per solució general

$$\left.\begin{aligned} x &= -7z - \frac{9}{11} \\ y &= 3z + 1 \end{aligned}\right\} .$$

Val la pena observar que en el cas d'un sistema indeterminat, com ho és aquest, *totes les solucions donen el mateix error quadràtic*, ja que aquest error és $\varepsilon = \| d' - d \|$ i aquesta expressió no depèn de quin vector x ens dóna $Cx = d'$. Aquí, fent, per exemple, $z = 0$, surt $d' = 1/11(13, 2, 15, 30)$ i un error quadràtic

$$\varepsilon = \| d' - d \| = \frac{10\sqrt{33}}{11} .$$

4.4 Orientació i producte vectorial

Es defineix una relació d'equivalència entre les bases ordenades de $\mathbb{R}^n$ dient que

$$\mathcal{B} \simeq \mathcal{B}' \quad \text{si, i només si,} \quad \det(C) > 0 ,$$

on C és la matriu del canvi de base de $\mathcal{B}'$ a $\mathcal{B}$. Aquesta relació és d'equivalència i classifica les bases de $\mathbb{R}^n$ en dos tipus. *Orientar* $\mathbb{R}^n$ és escollir un dels dos tipus: es diu que les seves bases son *positives*, i que les bases de l'altre tipus són *negatives*. En algunes qüestions (angle orientat, producte vectorial, girs, rotacions) és indispensable haver fixat una orientació, i el criteri habitual és prendre com a positives la base canònica de $\mathbb{R}^n$ i les del seu tipus.

La interpretació física de la relació d'equivalència és la següent. En el pla: si anem del primer vector de la base al segon pel camí més curt, o seguim el sentit horari o l'antihorari, i aquesta és la classificació. Es habitual prendre com a positives les bases de sentit *antihorari*. A l'espai: girant un tornavís segons el camí més curt del primer vector al segon, l'estri es mourà en el sentit del tercer vector o en sentit contrari. És habitual prendre com a positives les bases ortogonals en les quals el tornavís segueix el sentit del tercer vector (*regla del tornavís*) i les equivalents a elles.

Exemples 4.14 (a) Si $\mathcal{B}$ i $\mathcal{B}'$ són bases ortonormals de $\mathbb{R}^n$ i C la matriu (ortogonal) de canvi de base, $\mathcal{B}$ i $\mathcal{B}'$ són del mateix tipus si, i només si, $\det(C) = 1$.

(b) Si $\mathcal{B} = \{e_1, e_2\}$ és una base positiva de $\mathbb{R}^2$, $\{e_2, e_1\}$ és negativa. A $\mathbb{R}^3$, l'efecte de permutar els vectors d'una base depèn del signe de la permutació. Així, si $\mathcal{B} = \{e_1, e_2, e_3\}$ és positiva, $\{e_3, e_1, e_2\}$ també ho és, però $\{e_1, e_3, e_2\}$ és negativa.

(c) (*Angle orientat en el pla.*) Recordem que, donats dos vectors no nuls $u, v \in \mathbb{R}^2$, $\alpha(u, v) \in [0, \pi]$ representa l'angle no orientat entre u i v. Suposem ara orientat $\mathbb{R}^2$. Donats dos vectors linealment independents u i v, definim l'*angle orientat de u a v* com

$$\alpha^+(u, v) = \begin{cases} \alpha(u, v) & \text{si } \{u, v\} \text{ és base positiva} \\ -\alpha(u, v) \equiv 2\pi - \alpha(u, v) & \text{si } \{u, v\} \text{ és base negativa}. \end{cases}$$

Si u i v són diferents de zero però linealment dependents, les dues opcions ens donen $\alpha^+(u, v) = \alpha(u, v)$, ja que $\alpha(u, v)$ és 0 ó π. Notem que l'interval de variació de $\alpha^+(u, v)$ és d'amplitud 2π. Es immediat comprovar que $\alpha^+(v, u) = -\alpha^+(u, v)$, i que, si $\{e_1, e_2\}$ és la base canònica, $\alpha^+(e_1, e_2) = \pi/2$.

(d) (*Coordenades polars.*) Sigui $u = (x, y) \neq 0$ un vector arbitrari del pla orientat. Com ja hem vist, si posem $\alpha = \alpha(u, e_1)$, $\beta = \alpha(u, e_2)$, i $r = \| u \|$, tenim $x = r \cos\alpha$, $y = r \cos\beta$, amb $\cos^2\alpha + \cos^2\beta = 1$ i, per tant, $\cos\beta = \pm\sin\alpha$. Definim l'*argument* del vector u com $\varphi = \alpha^+(e_1, u)$. Com que el determinant de la matriu del canvi de $\{e_1, u\}$ a $\{e_1, e_2\}$ val y, i $\sin\alpha \geq 0$ perquè $\alpha \in [0, \pi]$, tenim: en els quadrants primer i segon, on $y \geq 0$, $\varphi = \alpha$ i $\cos\beta = \sin\alpha = \sin\varphi$; en els quadrants tercer i quart, on $y \leq 0$, $\varphi = -\alpha$ i $\cos\beta = -\sin\alpha = \sin\varphi$. Per tant,

$$x = r\cos\varphi\,, \qquad y = r\sin\varphi \qquad \forall u = (x, y) \neq 0 \text{ de } \mathbb{R}^2\,.$$

El mòdul r i l'argument φ són les *coordenades polars* del vector u.

(e) (*Orientació de plans a l'espai.*) La noció d'orientació per a subespais de $\mathbb{R}^n$ és anàloga a la que hem descrit per a $\mathbb{R}^n$. Limitant–nos ara a $\mathbb{R}^3$, cal observar que l'orientació de l'espai no determina una orientació en els seus plans, i això reflecteix el fet que un observador pot situar–se a qualsevol dels dos costats d'un pla. El procediment per orientar un pla H de l'espai és el següent: escollim un dels dos sentits possibles dels vectors ortogonals al pla (recordem que $\dim H^{\perp} = 1$), seleccionant un vector concret $w \perp H$; llavors orientem el pla H prenent com a positives les bases $\{u, v\}$ de H tals que $\{u, v, w\}$ és positiva a $\mathbb{R}^3$.

Considerem, com a il·lustració, el pla d'equació $x - 2y + z = 0$. L'orientem escollint el vector ortogonal $w = (1, -2, 1)$. Donada una base del pla, per exemple $\{(2, 1, 0), (0, 1, 2)\}$, podem classificar–la com a positiva perquè ho és a $\mathbb{R}^3$ la base $\{(2, 1, 0), (0, 1, 2), (1, -2, 1)\}$, ja que la matriu del canvi a la base canònica i el seu determinant són

$$C = \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & -2 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{i} \quad \det(C) = 10 > 0\,.$$

Considerem l'espai $\mathbb{R}^3$ orientat de la forma habitual. Introduirem el producte vectorial d'una manera constructiva que posarà de manifest les seves propietats característiques.

Teorema 4.15 *Donats dos vectors linealment independents $u, v \in \mathbb{R}^3$, existeix un únic vector $w \in \mathbb{R}^3$ tal que:*

(a) $w \perp u$ *i* $w \perp v$.

(b) $\| w \| = \| u \| \| v \| \sin\alpha$, *on* $\alpha = \alpha(u, v)$.

(c) $\{u, v, w\}$ *és una base positiva de* $\mathbb{R}^3$.

DEMOSTRACIÓ. Escrivim $u = (x_1, x_2, x_3)$, $v = (y_1, y_2, y_3)$ i $w = (z_1, z_2, z_3)$. La condició (a) dóna lloc al sistema següent:

$$\left.\begin{aligned} x_1z_1 + x_2z_2 + x_3z_3 = 0 \\ y_1z_1 + y_2z_2 + y_3z_3 = 0 \end{aligned}\right\} .$$

La solució general d'aquest sistema és

$$\left.\begin{aligned} z_1 &= A_1t \\ z_2 &= A_2t \\ z_3 &= A_3t \end{aligned}\right\} , \quad \text{on} \quad A_1 = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} , \quad A_2 = \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix} \quad \text{i} \quad A_3 = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} .$$

De la condició (b) obtenim

$$\| w \|^2 = \| u \|^2 \| v \|^2 \sin^2 \alpha = \| u \|^2 \| v \|^2 (1 - \cos^2 \alpha) = \| u \|^2 \| v \|^2 - (u \cdot v)^2 .$$

Substituint les components de u, v i w resulta

$$\begin{aligned} t^2(A_1^2 + A_2^2 + A_3^2) &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1 y_1 + x_2 y_2 + x_3 y_3)^2 \\ &= (x_1 y_2 - x_2 y_1)^2 + (x_2 y_3 - x_3 y_2)^2 + (x_3 y_1 - x_1 y_3)^2 \\ &= A_1^2 + A_2^2 + A_3^2 , \end{aligned}$$

i d'aquí $t = \pm 1$. Aplicant, finalment, la condició (c), tenim que

$$0 < \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ tA_1 & tA_2 & tA_3 \end{vmatrix} = t(A_1^2 + A_2^2 + A_3^2) ,$$

i es dedueix que $t = 1$. Per tant, queda determinat un únic vector $w = (A_1, A_2, A_3)$ que compleix les tres condicions de l'enunciat. $\square$

El vector w determinat pel teorema anterior és el *producte vectorial* de u i v, escrit des d'ara $u \wedge v$. Una regla mnemotècnica per recordar la seva expressió explícita és la següent:

$$u \wedge v = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

on $\{e_1, e_2, e_3\}$ és la base canònica de $\mathbb{R}^3$. Això permet estendre la definició al cas en que u i v són linealment dependents, posant $u \wedge v = 0$ en concordància amb el determinant formal.

És immediat comprovar que el producte vectorial és una operació distributiva respecte a la suma, compatible amb el producte per escalars i anticommutativa: $v \wedge u = -u \wedge v$. A més, si $\{u, v\}$ és un conjunt ortonormal, $\{u, v, u \wedge v\}$ és una base ortonormal positiva de $\mathbb{R}^3$. La propietat següent, útil en alguns contextos, és més complexa.

Proposició 4.16 *(Doble producte vectorial.) Si* $u, v, w \in \mathbb{R}^3$,

$$u \wedge (v \wedge w) = (u \cdot w)v - (u \cdot v)w .$$

DEMOSTRACIÓ. Per tal de simplificar al màxim els càlculs, escollim (sempre és possible) una base ortonormal positiva $\{e_1, e_2, e_3\}$ tal que

$$u = ae_1 , \qquad v = be_1 + ce_2 \qquad \text{i} \qquad w = pe_1 + qe_2 + re_3 .$$

Aleshores tenim, fent les operacions,

$$\begin{aligned} u \wedge (v \wedge w) &= -a(bq - cp)e_2 - abre_3 \\ (u \cdot w)v - (u \cdot v)w &= ap(be_1 + ce_2) - ab(pe_1 + qe_2 + re_3)\,, \end{aligned}$$

i és evident que els dos resultats coincideixen. □

Una primera aplicació del producte vectorial és la següent. Donada una base $\{u, v\}$ d'un pla, $u \wedge v$ és un generador de la recta ortogonal al pla. Dit d'una altra manera igualment útil: si una recta està definida com a intersecció de dos plans, el producte vectorial dels vectors formats amb els coeficients de les equacions dels plans és un generador de la recta.

Finalment, el *producte mixt* de tres vectors $u, v, w \in \mathbb{R}^3$ es defineix com $u \cdot (v \wedge w)$, i per tant coincideix amb el determinant de la matriu de les components dels tres vectors:

$$u \cdot (v \wedge w) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}\,.$$

En particular, $u \cdot (v \wedge w) = 0$ si, i només si, u, v, w són linealment dependents, i $u \cdot (v \wedge w) = (u \wedge v) \cdot w$.

5

Transformacions lineals

Molts processos matemàtics, geomètrics, físics, químics i econòmics són representables, si més no en primera aproximació, per relacions lineals entre les variables que hi intervenen. Aquesta evidència, juntament amb l'íntima connexió entre les transformacions lineals i el càlcul matricial, concedeix un paper molt especial a aquest capítol i el següent.

5.1 Definicions i exemples

Observem en primer lloc que, de la mateixa manera que és habitual descriure una funció $f : \mathbb{R} \longrightarrow \mathbb{R}$ per una fórmula com $f(x) = x^2 + 3x$, o per l'equació equivalent $y = x^2 + 3x$, una transformació entre dos espais euclidians, per exemple $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, es pot definir per una fórmula com

$$T(x, y, z) = (xy - 2, x + y + 3z^2)\,,$$

o per les equacions corresponents

$$\begin{cases} x' = xy - 2 \\ y' = x + y + 3z^2\,. \end{cases}$$

Recordem també la nomenclatura relativa a les aplicacions en general. Sigui $T : A \longrightarrow B$ una aplicació entre dos conjunts arbitraris. Si $T(a) = b$, direm que b és la *imatge* de a per la transformació T i que a és una *antiimatge* de b. D'altra banda, es diu que T és *injectiva* si $a \neq a'$ implica $T(a) \neq T(a')$. Es diu que T és *exhaustiva* si per a qualsevol $b \in B$ existeix algun $a \in A$ tal que $T(a) = b$. Es diu que T és *bijectiva* si és injectiva i exhaustiva. Si T és bijectiva, cada element $b \in B$ té una sola antiimatge $a \in A$, i escrivint $T^{-1}(b) = a$ definim una nova aplicació $T^{-1} : B \longrightarrow A$, anomenada *inversa* de T. Més en general, el símbol $T^{-1}(b)$ denota el conjunt de totes les antiimatges de l'element $b \in B$, encara que T no sigui bijectiva. Finalment, donades dues aplicacions $T : A \longrightarrow B$ i $S : B \longrightarrow C$, podem formar la *composició* de S i T, que és l'aplicació $S \circ T : A \longrightarrow C$ definida per $(S \circ T)(a) = S(T(a))$.

Definició 5.1 Una transformació $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ és *lineal* si compleix les condicions

(a) $T(u + u') = T(u) + T(u')$

(b) $T(tu) = tT(u)$

per a qualssevol $t \in \mathbb{R}$, $u, u' \in \mathbb{R}^n$. De les propietats (a) i (b) es dedueix immediatament que $T(0) = 0$, que $T(u - u') = T(u) - T(u')$ i que, en general,

$$T(t_1u_1 + t_2u_2 + \cdots + t_ku_k) = t_1T(u_1) + t_2T(u_2) + \cdots + t_kT(u_k)\,.$$

Una transformació lineal $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ s'anomena un *endomorfisme* de $\mathbb{R}^n$, i aquest és el cas que tractarem amb més freqüència.

Teorema 5.2 *(Caracterització matricial.) Una transformació $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ és lineal si, i només si, es pot representar matricialment per*

$$Y = AX\,.$$

DEMOSTRACIÓ. ($\Rightarrow$) Donades les bases $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ de $\mathbb{R}^n$ i $\mathcal{B}' = \{v_1, v_2, \ldots, v_m\}$ de $\mathbb{R}^m$, siguin

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{pmatrix} \qquad \text{i} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

les columnes de components d'un vector arbitrari $u \in \mathbb{R}^n$ i un vector arbitrari $v \in \mathbb{R}^m$ en les bases respectives i la matriu obtinguda de les relacions

$$T(u_j) = \sum_{i=1}^{m} a_{ij} v_i \quad \text{per a} \quad j = 1, 2, \ldots, n\,,$$

que expressen les imatges dels vectors de la primera base com a combinacions lineals dels vectors de la segona. Aplicant la linealitat de T al càlcul de la imatge del vector

$$u = \sum_{j=1}^{n} x_j u_j$$

obtenim que

$$\begin{aligned} T(u) &= T\left(\sum_{j=1}^{n} x_j u_j\right) = \sum_{j=1}^{n} x_j T(u_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} v_i\right) \\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} x_j v_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right) v_i \end{aligned}$$

i, comparant amb $v = \sum_{i=1}^{m} y_i v_i$, resulta que la relació intrínseca

$$v = T(u)$$

es tradueix en components per

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \quad \text{per a} \quad i = 1, 2, \ldots, m ,$$

o, matricialment,

$$Y = AX .$$

($\Leftarrow$) Si T està definida respecte a dues bases per $Y = AX$, on A és una matriu arbitrària $m \times n$, és evident que, si $Y_1 = AX_1$, $Y_2 = AX_2$ i $t \in \mathbb{R}$, aleshores

$$A(X_1 + X_2) = AX_1 + AX_2 = Y_1 + Y_2 \qquad \text{i} \qquad A(tX_1) = tAX_1 = tY_1 ,$$

i per tant T és lineal. $\square$

Observacions 5.3 Podem extreure més informació d'aquest teorema.

(a) Les components de les imatges $T(u_1), T(u_2), \ldots, T(u_n)$ en la base $\{v_1, v_2, \ldots, v_m\}$ formen les columnes de la matriu A, que rep el nom de *matriu de T relativa a les dues bases*:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} .$$

Simbolitzarem aquesta relació escrivint $A = M(T, \mathcal{B}, \mathcal{B}')$.

(b) Aquestes imatges determinen T i, a més, poden ser escollides arbitràriament. És a dir, una vegada fixades bases dels dos espais, *tota matriu A del tipus $m \times n$ descriu una transformació lineal T*.

(c) A efectes pràctics, convé notar que és característic d'una transformació lineal el fet que les seves equacions vinguin donades per *relacions lineals homogènies* entre els dos conjunts de variables. Amb aquest criteri, el lector pot observar que l'exemple de transformació proposat a l'inici d'aquest capítol no és, en absolut, lineal.

(d) En el cas d'un endomorfisme, emprarem només una base de referència. La matriu serà quadrada, i escriurem simplement $A = M(T, \mathcal{B})$. Les columnes de A ens donaran les components de les imatges $T(u_1), T(u_2), \ldots, T(u_n)$ en funció dels vectors $u_1, u_2, \ldots, u_n$.

Exemples 5.4 (a) És fàcil comprovar que les funcions lineals $f : \mathbb{R} \longrightarrow \mathbb{R}$ són les de la forma $y = f(x) = ax$ (notem l'analogia formal amb el cas general, on tenim la relació matricial $Y = T(X) = AX$). Com és ben sabut, amb funcions lineals descrivim, entre molts altres processos, costos, interessos, descomptes, recàrrecs...

(b) (*Equacions d'un gir en el pla orientat.*) Un *gir d'angle* $\alpha \in [0, 2\pi)$ al voltant de l'origen del pla orientat $\mathbb{R}^2$ és la transformació $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ que conserva el mòdul dels vectors i incrementa els arguments en un angle α. Per tant, $G(0,0) = (0,0)$ i, si $u = (x, y) = (r\cos\varphi, r\sin\varphi)$, i $v = G(u) = (x', y')$, tenim les relacions

$$x' = r\cos(\varphi + \alpha)\,, \qquad y' = r\sin(\varphi + \alpha)\,,$$

que, introduint les variables x, y, ens condueixen a unes equacions típicament lineals segons el teorema de caracterització:

$$\left.\begin{aligned} x' &= x\cos\alpha - y\sin\alpha \\ y' &= x\sin\alpha + y\cos\alpha \end{aligned}\right\} \qquad \text{o bé} \qquad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Així doncs, tot gir G del pla orientat és un endomorfisme, i la seva matriu en la base canònica és

$$A = M(G, \mathcal{B}_c) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

(c) (*Equacions d'una simetria axial en el pla.*) Considerem la simetria $S : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecte a la recta $x = 3y$ (*eix de simetria*). Suposant ara comprovades —per exemple geomètricament— les condicions de linealitat, només caldrà trobar les imatges dels dos vectors de la base canònica per determinar la matriu. Aplicant la fórmula del capítol anterior,

$$S(u) = 2\frac{u \cdot w}{w \cdot w}\, w \, - \, u\,,$$

i tenint en compte que un vector director de l'eix de simetria és, per exemple, $w = (3, 1)$, obtenim $S(1, 0) = (4/5, 3/5)$ i $S(0, 1) = (3/5, -4/5)$. Per tant, l'equació matricial d'aquesta simetria axial en la base canònica és

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} .$$

En aquest exemple hem fet servir el teorema de caracterització en el sentit recíproc del de l'exemple anterior: suposant justificada a priori la linealitat, l'aprofitem per trobar amb el mínim esforç les equacions generals de la transformació.

5.2 Canvis de base

Tot i que és habitual —i sovint convenient— treballar amb la matriu d'una transformació lineal relativa a les bases canòniques, és important conèixer els efectes dels canvis de base. Per fixar idees, considerarem amb detall només el cas dels endomorfismes.

Teorema 5.5 *(Fórmula del canvi de base.) Considerem un endomorfisme $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ i dues bases, $\mathcal{B}$ i $\mathcal{B}'$. Si $A = M(T, \mathcal{B})$, $A' = M(T, \mathcal{B}')$, i C és la matriu del canvi de base de $\mathcal{B}'$ a $\mathcal{B}$, aleshores*

$$A' = C^{-1}AC .$$

DEMOSTRACIÓ. La relació $u = T(v)$ es tradueix per $Y = AX$ en la base $\mathcal{B}$, i per $Y' = A'X'$ en la base $\mathcal{B}'$. El pas de X' a Y' es pot obtenir també en tres fases: (a) passem el vector u de $\mathcal{B}'$ a $\mathcal{B}$; (b) transformem el vector u en $v = T(u)$; (c) passem el vector v de $\mathcal{B}$ a $\mathcal{B}'$. Matricialment,

(a) $X' \mapsto X = CX'$

(b) $X \mapsto Y = AX = ACX'$

(c) $Y \mapsto Y' = C^{-1}Y = C^{-1}ACX'$.

Com que aquesta era la transformació que duia a terme la matriu A', deduim que

$$A' = C^{-1}AC . \quad \square$$

Observació 5.6 El lector no hauria de tenir dificultat en generalitzar el resultat anterior per a una transformació lineal qualsevol $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ aplicant un canvi de base a cada espai: Si $\mathcal{B}_1$ i $\mathcal{B}'_1$ són bases de $\mathbb{R}^n$, i $\mathcal{B}_2$ i $\mathcal{B}'_2$ són bases de $\mathbb{R}^m$, la relació entre les matrius

$$A = M(T, \mathcal{B}_1, \mathcal{B}_2) \qquad \text{i} \qquad A' = M(T, \mathcal{B}'_1, \mathcal{B}'_2)$$

és
$$A' = D^{-1}AC\,,$$
on C és la matriu del canvi de base en el primer espai, de $\mathcal{B}'_1$ a $\mathcal{B}_1$, i D és la matriu del canvi de base en el segon, de $\mathcal{B}'_2$ a $\mathcal{B}_2$.

Exemples 5.7 (a) Si $\mathcal{B}'$ és una base ortogonal positiva del pla orientat, formada per dos vectors de la mateixa longitud, $e'_1 = (a, b)$ i $e'_2 = (-b, a)$, la matriu del gir d'angle α en la base $\mathcal{B}'$ coincideix amb la matriu del gir en la base canònica, ja que tenim
$$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$
i
$$A' = C^{-1}AC = \frac{1}{a^2+b^2}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

(b) Considerem la simetria axial estudiada a la secció anterior. Amb el vector director de l'eix $e'_1 = (3, 1)$ i el vector ortogonal $e'_2 = (-1, 3)$ podem formar una base adaptada a la simetria, $\mathcal{B}' = \{e'_1, e'_2\}$. Es evident que $S(e'_1) = e'_1$, $S(e'_2) = -e'_2$ i, per tant, la matriu de la simetria axial en aquesta base és ben simple:
$$A' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
Si ara volem calcular la matriu de S en la base canònica, només cal desfer el canvi $A' = C^{-1}AC$, obtenint
$$A = CA'C^{-1}\,.$$
Aplicant aquesta nova relació resulta
$$C = \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}, \qquad A' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{i} \qquad A = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}.$$

(c) *El determinant de la matriu d'un endomorfisme no varia en fer un canvi de base.* En efecte: aplicant determinants a la fórmula del canvi resulta
$$\det(A') = \det(C^{-1}AC) = \det(C^{-1})\det(A)\det(C) = \frac{1}{\det(C)}\det(A)\det(C) = \det(A)\,.$$

(d) Suposem que un endomorfisme $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ queda definit de la manera següent:
$$T(1, 2) = (5, 1), \qquad T(-2, 1) = (4, 3)\,.$$

Per trobar $A = M(T, \mathcal{B}_c)$ només cal observar que tenim una nova base $\mathcal{B}'$, formada pels vectors $e_1' = (1, 2)$ i $e_2' = (-2, 1)$, i considerar les matrius següents:

$$C = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \text{ (canvi de } \mathcal{B}' \text{ a } \mathcal{B}_c) \qquad \text{i} \qquad B = M(T, \mathcal{B}', \mathcal{B}_c) = \begin{pmatrix} 5 & 4 \\ 1 & 3 \end{pmatrix}.$$

Aplicant un raonament similar al del teorema del canvi de base resulta

$$A = BC^{-1} = \frac{1}{5}\begin{pmatrix} -3 & 14 \\ -5 & 5 \end{pmatrix}.$$

5.3 *Nucli, imatge i rang*

Considerem una transformació lineal arbitrària $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Es defineix el *nucli* de T com el subespai format per les antiimatges del vector 0: simbòlicament

$$\mathrm{Nuc}(T) = T^{-1}(0) = \{u \in \mathbb{R}^n : T(u) = 0\}.$$

Es defineix la *imatge* de T com el subespai format per totes les imatges dels vectors de $\mathbb{R}^n$:

$$\mathrm{Im}(T) = \{v \in \mathbb{R}^m : v = T(u),\ u \in \mathbb{R}^n\}.$$

(El lector hauria de comprovar que, efectivament, $\mathrm{Nuc}(T)$ i $\mathrm{Im}(T)$ són subespais.) Finalment, el *rang* de T es defineix per

$$\mathrm{rang}(T) = \dim \mathrm{Im}(T).$$

Teorema 5.8 *(a) Si $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ és una base de $\mathbb{R}^n$, $\{T(u_1), T(u_2), \ldots, T(u_n)\}$ és un sistema de generadors del subespai $\mathrm{Im}(T)$. Per tant, $\mathrm{rang}(T)$ és el rang de qualsevol matriu A que representi a T.*

(b) $\dim\ \mathrm{Nuc}(T) + \dim\ \mathrm{Im}(T) = n$.

DEMOSTRACIÓ. (a) Si $v \in \mathrm{Im}(T)$, sabem que $v = T(u)$ per a algun $u \in \mathbb{R}^n$. Com que $u = t_1u_1 + t_2u_2 + \cdots + t_nu_n$, resulta

$$v = T(u) = T(t_1u_1 + t_2u_2 + \cdots + t_nu_n) = t_1T(u_1) + t_2T(u_2) + \cdots + t_nT(u_n).$$

Com que les columnes de la matriu A representen, doncs, un sistema de generadors de $\mathrm{Im}(T)$, queda establerta també la identitat entre els rangs.

(b) El sistema d'equacions descrit matricialment per $AX = 0$ defineix el nucli. Per tant,

$$\dim\ \mathrm{Nuc}(T) = n - \mathrm{rang}(A) = n - \dim\ \mathrm{Im}(T). \quad \square$$

Corol·lari 5.9 *(a) T és una transformació exhaustiva si, i només si,* $\mathrm{rang}(T) = m$.

(b) T és injectiva si, i només si, $\mathrm{Nuc}(T) = \{0\}$, *és a dir, si, i només si,* $\mathrm{rang}(T) = n$.

(c) Un endomorfisme T de $\mathbb{R}^n$ és bijectiu si, i només si, $\mathrm{rang}(T) = n$, *és a dir,* $\det(A) \neq 0$, *on A és qualsevol matriu que representi a T.*

DEMOSTRACIÓ. (a) T és exhaustiva si, i només si, $\mathrm{Im}(T) = \mathbb{R}^m$, i això és equivalent a que $\mathrm{rang}(T) = m$.

(b) Si T és injectiva, òbviament ha de ser $\mathrm{Nuc}(T) = \{0\}$. Recíprocament, si $\mathrm{Nuc}(T) = \{0\}$, de la igualtat $T(u) = T(u')$ resulta $T(u - u') = 0$, és a dir, $u - u' \in \mathrm{Nuc}(T)$ i, per tant, $u = u'$: això significa que T és injectiva. La segona equivalència és una conseqüència immediata de la fórmula de les dimensions del teorema anterior.

(c) Es desprèn dels dos apartats anteriors. $\square$

Exemples 5.10 (a) Considerem la projecció ortogonal de $\mathbb{R}^n$ sobre un subespai H, és a dir, la transformació $P : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tal que $P(u) = u'$, projecció ortogonal de u sobre H. Es fàcil veure que P és lineal, i per tant un endomorfisme de $\mathbb{R}^n$, i també que

$$\mathrm{Nuc}(P) = H^{\perp}, \qquad \mathrm{Im}(P) = H \qquad \text{i} \qquad \mathrm{rang}(P) = \dim(H)\,.$$

(b) Aquest exemple també permet il·lustrar una fórmula general per al càlcul d'antiimatges que és útil en alguns contextos i el lector pot justificar com a exercici: si $v = T(u)$, aleshores

$$T^{-1}(v) = u + \mathrm{Nuc}(T)\,,$$

és a dir, coneguda una antiimatge concreta u de v, totes les antiimatges d'aquest vector s'obtenen sumant a u un vector arbitrari del nucli de la transformació.

(c) Considerem ara, també a $\mathbb{R}^n$, la simetria respecte a un subespai H, és a dir, la transformació $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tal que $S(u) = u^*$, simètric de u respecte a H. Si $H = \{0\}$ és diu que S és la *simetria central*, si H és una recta S és una *simetria axial*, i si H és un pla S és una *simetria plana*. També és fàcil veure que S és lineal, i per tant un endomorfisme de $\mathbb{R}^n$, però en aquest cas resulta

$$\mathrm{Nuc}(S) = \{0\}, \qquad \mathrm{Im}(S) = \mathbb{R}^n \qquad \text{i} \qquad \mathrm{rang}(S) = n\,,$$

de manera que S és una transformació bijectiva.

5.4 Operacions

Com en el cas dels canvis de base, en aquesta secció ens limitarem a considerar operacions amb endomorfismes, deixant al lector la senzilla tasca de generalitzar la major part de les idees i els resultats a transformacions lineals entre espais diferents.

Si $t \in \mathbb{R}$, i T i S són endomorfismes de $\mathbb{R}^n$, és immediat comprovar que les transformacions de $\mathbb{R}^n$ en si mateix definides per

$$(T + S)(u) = T(u) + S(u) \qquad \text{i} \qquad (tT)(u) = tT(u) \qquad \forall u \in \mathbb{R}^n$$

són lineals. I establir les corresponents relacions matricials:

$$M(T + S, \mathcal{B}) = M(T, \mathcal{B}) + M(S, \mathcal{B}) \qquad \text{i} \qquad M(tT, \mathcal{B}) = tM(T, \mathcal{B})\,.$$

Anàlogament, és fàcil veure que la transformació composta $S \circ T$ és lineal, i correspon al producte de matrius:

$$M(S \circ T, \mathcal{B}) = M(S, \mathcal{B})M(T, \mathcal{B})\,.$$

Finalment, recordem que, si $A = M(T, \mathcal{B})$, T és bijectiu si, i només si, A és regular. Seguint la mateixa tècnica que en els casos anteriors, es pot provar que, en aquest supòsit, T^{-1} també és lineal i

$$M(T^{-1}, \mathcal{B}) = M(T, \mathcal{B})^{-1}\,.$$

Es important esmentar l'existència, per a cada n, de l'endomorfisme anomenat *identitat*, denotat per $I : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ i definit per $I(u) = u$ per a tot $u \in \mathbb{R}^n$. La matriu d'aquest endomorfisme en qualsevol base és la matriu unitat I, i d'aquí l'abús de notació que fem. Les relacions de I amb les dues darreres operacions són les següents:

$$T \circ I = T = I \circ T$$

i, si T és bijectiu,

$$T \circ T^{-1} = I = T^{-1} \circ T\,.$$

5.5 Endomorfismes i producte escalar

Estudiarem ara diversos tipus d'endomorfismes de l'espai $\mathbb{R}^n$ que són interessants per la seva connexió amb l'estructura euclidiana, fent especial insistència en els casos del pla i de l'espai. Aquests tipus de transformacions són les projeccions ortogonals i les denominades isometries (transformacions rígides), entre les quals cal destacar les simetries i, sobre tot, els girs del pla i les rotacions de l'espai.

Recordem que, donat un subespai H de $\mathbb{R}^n$, denotem per P i S, respectivament, la projecció ortogonal sobre H i la simetria respecte a H. Les relacions següents són clares:

$$P^2 = P, \qquad S^2 = I \qquad \text{i} \qquad S = 2P - I\,.$$

Les matrius de P i S relatives a una base $\mathcal{B}'$ adaptada a la suma directa $H \oplus H^\perp$ són, respectivament,

$$A' = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \qquad \text{i} \qquad B' = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}.$$

Aquestes matrius A' i B' són simètriques, però també ho són les matrius respectives A i B en la base canònica $\mathcal{B}_c$. En efecte, suposant ortonormal la base $\mathcal{B}'$, la matriu C del canvi de base de $\mathcal{B}'$ a $\mathcal{B}_c$ és ortogonal, per tant $C^{-1} = C^t$, $A = CA'C^t$ i

$$A^t = (CA'C^t)^t = CA'C^t = A\,,$$

i un argument anàleg serveix per a la matriu B. A més, la matriu B és ortogonal, perquè $B^t = B$ i $B^2 = I$, de manera que $B^tB = I$.

Abans de discutir la noció general d'isometria, passem als girs del pla i les rotacions de l'espai. La descripció dels girs en el pla orientat s'ha dut a terme amb tota generalitat en la primera secció d'aquest capítol, i serà útil per a l'estudi de les rotacions.

Donats un angle $\alpha \in [0, 2\pi)$ i un vector unitari w de l'espai $\mathbb{R}^3$ orientat, considerarem el pla $H = \langle w\rangle^\perp$ orientat per w, i definirem la *rotació d'angle α al voltant del vector w* com l'endomorfisme $R : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ descrit de la forma següent en termes de la suma directa $\mathbb{R}^3 = \langle w\rangle \oplus H$:

$$R(u) = \begin{cases} u & \text{si } u \in \langle w\rangle \\ G(u) & \text{si } u \in H \end{cases}$$

on G és el gir d'angle α en el pla H orientat per w. Com que qualsevol vector $u \in \mathbb{R}^3$ s'escriu de forma única com $u = u' + u''$, amb $u' \in H$ i $u'' \in \langle w\rangle$, tenim, per linealitat,

$$R(u) = R(u') + R(u'') = G(u') + u''\,.$$

Evidentment, si $\{u, v\}$ és una base ortonormal positiva de H, la matriu de R en la base ortonormal positiva $\{w, u, v\}$ de $\mathbb{R}^3$ és

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

El resultat següent dóna una fórmula de càlcul directe d'imatges per a les rotacions.

Proposició 5.11 *Si R és la rotació d'angle α al voltant del vector unitari $w \in \mathbb{R}^3$,*

$$R(u) = u\cos\alpha + (1-\cos\alpha)(u\cdot w)w + (w\wedge u)\sin\alpha \qquad \forall u \in \mathbb{R}^3 .$$

DEMOSTRACIÓ. Escrivim $u = u' + u''$. Si $u' = 0$ aleshores $u = u''$, i la fórmula dóna, efectivament, $R(u) = R(u'') = u''$. Si $u' \neq 0$, $\{u', w\wedge u'\}$ és una base ortogonal positiva de H, amb $\|u'\| = \|w\wedge u'\|$. Per tant,

$$R(u') = u'\cos\alpha + (w\wedge u')\sin\alpha .$$

A més, $R(u'') = u'' = (u\cdot w)w$, ja que $\|w\| = 1$. Finalment, $w\wedge u' = w\wedge u$. Així doncs,

$$\begin{aligned} R(u) = R(u') + R(u'') &= u'\cos\alpha + (w\wedge u)\sin\alpha + (u\cdot w)w = \\ &= u\cos\alpha + (1-\cos\alpha)(u\cdot w)w + (w\wedge u)\sin\alpha . \quad \square \end{aligned}$$

Exemple 5.12 Calculem la imatge del vector $u = (1, 0, 1)$ segons la rotació de 90° al voltant del vector $w = (1, 1, 1)$. La rotació està ben definida perquè el vector w precisa el *sentit* de l'eix de rotació, però a la fórmula caldrà fer servir el vector normalitzat, és a dir, $w = {}^1\!/\!\sqrt{3}(1, 1, 1)$. Els càlculs parcials principals i el resultat són:

$$u\cdot w = \frac{2}{\sqrt{3}}, \quad w\wedge u = \frac{1}{\sqrt{3}}(1, 0, -1) \quad \text{i} \quad R(u) = \frac{1}{3}(2+\sqrt{3}, 2, 2-\sqrt{3}) .$$

Les isometries formen una classe important d'endomorfismes de l'espai euclidià. Són les transformacions rígides, és a dir, aquelles que no deformen les figures de l'espai. Intuitivament, la rigidesa equival a la conservació de les longituds dels vectors i dels angles entre ells; com que aquests dos conceptes estan estretament lligats al producte escalar, és natural donar la definició següent.

Un endomorfisme $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ és una *isometria* de $\mathbb{R}^n$ si conserva el producte escalar dels vectors, és a dir, si

$$T(u)\cdot T(v) = u\cdot v \qquad \forall u, v \in \mathbb{R}^n .$$

(Aquesta condició és tan forta que es pot demostrar que tota transformació $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ que conserva el producte escalar ha de ser lineal.) A partir de la definició, és immediat comprovar que tota isometria T

(a) conserva l'ortogonalitat: si $u \perp v$, $T(u) \perp T(v)$;

(b) conserva la norma: $\|T(u)\| = \|u\| \quad \forall u \in \mathbb{R}^n$;

(c) conserva els angles no orientats: si $u, v \neq 0$, $\alpha(T(u), T(v)) = \alpha(u, v)$;

(d) és bijectiva .

Proposició 5.13 *Si A és la matriu d'un endomorfisme $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ relativa a una base ortonormal (per exemple, la base canònica),*

$$T \text{ és isometria} \iff A \text{ és ortogonal}.$$

DEMOSTRACIÓ. ($\Rightarrow$) Si T és una isometria ha de conservar els productes escalars entre els vectors de la base ortonormal de referència, condició que es tradueix en la relació $A^t A = I$.

($\Leftarrow$) Recíprocament, la condició $A^t A = I$ expressa que es conserven els productes escalars entre els vectors de la base i, per tant, també es conserva el producte escalar de qualsevol parell de vectors. □

Exemples 5.14 (a) Aplicant aquesta caracterització matricial de les isometries és fàcil comprovar, a partir de les matrius corresponents, que totes les simetries, els girs del pla i les rotacions de l'espai són isometries, mentre que les projeccions ortogonals no ho són.

(b) De la mateixa manera, o directament per la definició, es demostra que la identitat I, la inversa de qualsevol isometria i la composició d'isometries també són isometries.

(c) Tenint en compte que el determinant de la matriu d'un endomorfisme no depèn de la base a la qual està referida, les isometries es poden classificar en dos tipus. Les isometries *directes* són les que compleixen que $\det(A) = 1$, i evidentment conserven l'orientació de l'espai vectorial i, per tant, són realitzables físicament: els girs del pla i les rotacions de l'espai, així com la identitat (gir o rotació de 0°), la simetria central del pla (gir de 180°) i les simetries axials de l'espai (rotacions de 180°) són isometries directes. Les isometries *inverses* són les que compleixen que $\det(A) = -1$ i, per tant, inverteixen l'orientació de l'espai vectorial i no són realitzables físicament: les simetries axials del pla i la simetria central i les simetries planes de l'espai són isometries inverses.

Una qüestió òbvia en aquest punt és la següent: hem esgotat, amb els exemples precedents, el catàleg d'isometries? La resposta es donarà en el pròxim capítol, aplicant parcialment la teoria de la diagonalització de matrius, i proveïrà un criteri per classificar una isometria a partir de la seva matriu en una base ortonormal (la base canònica, per fixar idees). Limitarem l'estudi als casos del pla i l'espai, els més interessants des del punt de vista físic.

6

Diagonalització de matrius

D'entre els procediments de reducció de matrius a formes senzilles, la diagonalització és un dels més útils de cara a les aplicacions pràctiques del càlcul matricial. A diferència d'altres situacions, en aquesta qüestió s'obtenen millors resultats en el cas complex que en el cas real, ja que un punt crucial és la descomposició del polinomi característic i això fa intervenir el teorema fonamental de l'àlgebra (vegeu l'apèndix B). Exposarem la teoria en el cas real, tot remarcant els punts on l'ús dels nombres complexos suposa variacions. I encara que el marc natural de la diagonalització és el dels endomorfismes, serà convenient traduir les nocions i els resultats al llenguatge de les matrius quadrades, que és l'habitual de la tècnica.

6.1 Endomorfismes diagonalitzables

Hem vist en el capítol anterior que les matrius de la simetria axial $S : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ respecte a la recta $x = 3y$, en la base canònica i en la base adaptada $\mathcal{B}' = \{(3, 1), (-1, 3)\}$, són respectivament

$$\frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} \qquad \text{i} \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Observem que la segona matriu permet fer una concisa descripció de l'actuació de S: (a) deixa invariants els vectors u de l'eix de simetria; (b) canvia de signe els vectors v perpendiculars a l'eix; (c) converteix qualsevol altre vector del pla, que serà de la forma $u + v$, en $u - v$.

Més en general, també hem comprovat la simplicitat de les matrius de les simetries i les projeccions ortogonals a $\mathbb{R}^n$ quan sabem escollir una base adequada.

Recordem que una matriu quadrada $A = (a_{ij})$ és *diagonal* si $a_{ij} = 0$ quan $i \neq j$.

Un endomorfisme $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ és *diagonalitzable* si la seva matriu en alguna base $\mathcal{B}'$ és diagonal:

$$A' = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix} .$$

Per diagonalitzar un endomorfisme cal, doncs, trobar n vectors $e'_1, e'_2, \ldots, e'_n$ linealment independents tals que $T(e'_i) = \alpha_i e'_i$ per a $i = 1, 2, \ldots, n$. Això suggereix les definicions següents.

Un vector $e \in \mathbb{R}^n$ és un *vector propi* d'un endomorfisme T si

$$T(e) = \alpha e$$

per a algun $\alpha \in \mathbb{R}$. En particular, el vector 0 és un vector propi amb qualsevol $\alpha \in \mathbb{R}$. Per tant, només direm que un nombre $\alpha \in \mathbb{R}$ és un *valor propi* de T si existeix algun vector $e \neq 0$ tal que $T(e) = \alpha e$. Cada vector propi $e \neq 0$ està associat a un únic valor propi, perquè

$$\beta e = T(e) = \alpha e \quad \Longrightarrow \quad (\beta - \alpha)e = 0 \quad \Longrightarrow \quad \alpha = \beta .$$

La condició de vector propi es pot escriure en la forma $(T - \alpha I)(e) = 0$, i això és equivalent a dir que

$$e \in \mathrm{Nuc}(T - \alpha I) .$$

Ara podem tornar a la definició inicial, i dir que T és diagonalitzable si, i només si, existeix una base de $\mathbb{R}^n$ formada per vectors propis de T. I com que la condició $e \in \mathrm{Nuc}(T - \alpha I)$ dóna una descripció operativa dels vectors propis de valor propi α, ens ocuparem en primer lloc de veure com podem obtenir tots els valors propis de T.

Proposició 6.1 *α és un valor propi de T si, i només si,* $\det(\alpha I - T) = 0$.

DEMOSTRACIÓ. Recordem que un endomorfisme té nucli no nul si, i només si, el seu determinant és zero; per tant, existeixen vectors $e \neq 0$ dins de $\mathrm{Nuc}(T - \alpha I)$ si, i només si, $\det(T - \alpha I) = 0$, i això és equivalent a dir que $\det(\alpha I - T) = 0$. $\square$

Introduirem, doncs, una variable x i calcularem $\det(xI - T)$, és a dir, $\det(xI - A)$, on A és qualsevol matriu que representi a T (vegeu la primera de les observacions següents). El resultat és un polinomi de la forma

$$\det(xI - T) = x^n - A_1 x^{n-1} + A_2 x^{n-2} - \cdots + (-1)^n A_n \,,$$

que s'anomena *polinomi característic* de T. Les seves arrels, cada una amb la multiplicitat corresponent, són els valors propis de T. Per a cada arrel α, el subespai $\mathrm{Nuc}(T - \alpha I)$ està format per tots els vectors propis de T amb valor propi α.

Observacions 6.2 (a) El polinomi característic no depèn, com a determinant d'un endomorfisme, de la matriu a partir de la qual es calcula. En efecte: si A i A' són dues matrius de T, i C és la matriu de canvi de base que ens dóna $A' = C^{-1}AC$, tenim

$$\det(xI - A') = \det(C^{-1}xIC - C^{-1}AC) = \det(C^{-1}\{xI - A\}C)$$

$$= \det(C^{-1})\det(xI - A)\det(C) = \det(xI - A)\,.$$

(b) A més del càlcul directe del polinomi característic, és interessant conèixer un mètode alternatiu. Cada coeficient A_i de la fórmula de $\det(xI - T)$ és la suma dels *menors diagonals* d'ordre i de la matriu A, és a dir, menors d'ordre i que tenen la seva diagonal principal situada en la diagonal principal de A. Per a cada i hi ha $\binom{n}{i}$ menors diagonals d'ordre i. En particular, A_1 s'anomena la *traça* de A (suma dels coeficients de la diagonal principal), mentre que A_n és el determinant de A. Tot i que una demostració general d'aquesta fórmula és complicada de notació, comprovar–la per a $n = 2$ i $n = 3$ és un bon exercici d'aplicació del càlcul de determinants.

6.2 Condicions de diagonalització

Sigui T un endomorfisme de $\mathbb{R}^n$.

Lema 6.3 *Si $\alpha_1, \alpha_2, \ldots, \alpha_k$ són valors propis diferents de T, la suma dels subespais de vectors propis corresponents és directa.*

DEMOSTRACIÓ. Aplicarem inducció sobre el nombre de valors propis, k.

Si k=1, no hi ha res a demostrar.

Suposem que $k \geq 2$ i que l'enunciat és cert per a $k - 1$, és a dir, que la suma dels $k - 1$ primers subespais de vectors propis és directa:

$$\mathrm{Nuc}(T - \alpha_1 I) \oplus \mathrm{Nuc}(T - \alpha_2 I) \oplus \cdots \oplus \mathrm{Nuc}(T - \alpha_{k-1} I)\,.$$

Per demostrar que la suma dels k subespais és directa, hem de veure que si $e_i \in \mathrm{Nuc}(T - \alpha_i I)$ per a $i = 1, 2, \ldots, k$, només és possible la igualtat

$$e_1 + e_2 + \cdots + e_k = 0$$

quan cada $e_i = 0$ (que és una de les condicions de suma directa).

Aplicant T a aquesta igualtat tenim que

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_k e_k = 0\,,$$

mentre que multiplicant-la per α_k obtenim

$$\alpha_k e_1 + \alpha_k e_2 + \cdots + \alpha_k e_k = 0\,.$$

Restant aquestes dues igualtats resulta que

$$(\alpha_k - \alpha_1)e_1 + (\alpha_k - \alpha_2)e_2 + \cdots + (\alpha_k - \alpha_{k-1})e_{k-1} = 0\,.$$

Com que els coeficients d'aquesta suma són diferents de zero, i la suma dels $k-1$ primers subespais és directa, podem assegurar que $e_1 = e_2 = \cdots = e_{k-1} = 0$ i, en conseqüència, $e_k = 0$. □

De vegades, l'enunciat d'aquest lema s'expressa —sense massa precisió— dient que "vectors propis de valors propis diferents són linealment independents".

Lema 6.4 *Si α és un valor propi de T de multiplicitat m, i $d = \dim \mathrm{Nuc}(T - \alpha I)$, aleshores $1 \le d \le m$.*

DEMOSTRACIÓ. És evident que $d \ge 1$, ja que la definició de valor propi garanteix l'existència d'algun $e \ne 0$ dins de $\mathrm{Nuc}(T - \alpha I)$.

Prenem ara una base $\{e_1, e_2, \ldots, e_d\}$ de $\mathrm{Nuc}(T - \alpha I)$ i l'ampliem fins a obtenir una base $\{e_1, e_2, \ldots, e_d, e_{d+1}, \ldots, e_n\}$ de $\mathbb{R}^n$. L'aspecte de la matriu de T en aquesta base és

$$\begin{pmatrix} \begin{pmatrix} \alpha & & & \\ & \alpha & & \\ & & \overset{d}{\ddots} & \\ & & & \alpha \end{pmatrix} & \begin{pmatrix} & & \\ & C & \\ & & \end{pmatrix} \\ & \\ 0 & \begin{pmatrix} & & \\ & B & \\ & & \end{pmatrix} \end{pmatrix},$$

de manera que desenvolupant $\det(xI - T)$ per les primeres d columnes queda

$$\det(xI - T) = (x - \alpha)^d \det(xI - B)\,.$$

Això assegura que la multiplicitat de l'arrel α és $m \geq d$. $\square$

Teorema 6.5 *(Condició necessària i suficient de diagonalització.) Un endomorfisme T de $\mathbb{R}^n$ és diagonalitzable si, i només si, compleix les dues condicions següents:*

(a) El polinomi característic es descompon totalment en factors de primer grau:

$$\det(xI - T) = (x - \alpha_1)^{m_1}(x - \alpha_2)^{m_2} \cdots (x - \alpha_k)^{m_k}\,.$$

(b) La dimensió de cada nucli coincideix amb la multiplicitat del valor propi corresponent:

$$d_i = m_i \quad \textit{per a} \quad i = 1, \ldots, k\,.$$

DEMOSTRACIÓ. ($\Longrightarrow$) Si T és diagonalitzable, la suma directa dels nuclis coincideix amb tot l'espai $\mathbb{R}^n$, perquè els vectors propis que formen la base en la que la matriu de T és diagonal surten necessàriament d'aquests nuclis. Així doncs, la matriu diagonal de T

$$\begin{pmatrix} \alpha_1 &&&&&&&& \\ & \ddots^{m_1} &&&&&&& \\ && \alpha_1 &&&&&& \\ &&& \alpha_2 &&&&& \\ &&&& \ddots^{m_2} &&&& \\ &&&&& \alpha_2 &&& \\ &&&&&& \alpha_k && \\ &&&&&&& \ddots^{m_k} & \\ &&&&&&&& \alpha_k \end{pmatrix}$$

dóna directament $\det(xI - T) = (x - \alpha_1)^{m_1}(x - \alpha_2)^{m_2} \cdots (x - \alpha_k)^{m_k}$. D'altra banda, tenim que $d_i \leq m_i$ per a tot i segons el lema precedent; a més,

$$d_1 + d_2 + \cdots + d_k = n$$

perquè la suma directa dels nuclis és $\mathbb{R}^n$; finalment,

$$m_1 + m_2 + \cdots + m_k = n$$

segons resulta de calcular el grau del polinomi característic. De tot això es dedueix que

$$d_i = m_i \quad \text{per a} \quad i = 1, \ldots, k\,.$$

($\Longleftarrow$) Si T compleix les condicions (a) i (b), la suma directa dels nuclis és $\mathbb{R}^n$, ja que

$$\begin{aligned}
&\dim\left[\operatorname{Nuc}(T-\alpha_1 I)\oplus\cdots\oplus\operatorname{Nuc}(T-\alpha_k I)\right]=\\
&=\sum_{i=1}^{k}\dim\operatorname{Nuc}(T-\alpha_i I)=\sum_{i=1}^{k}d_i=\sum_{i=1}^{k}m_i=n\,.
\end{aligned}$$

Per tant, prenent una base adaptada a la suma directa tindrem una base de $\mathbb{R}^n$ formada per vectors propis de T, el qual serà en conseqüència un endomorfisme diagonalitzable. $\square$

Corol·lari 6.6 *(Una condició suficient.) Si un endomorfisme T de $\mathbb{R}^n$ té n valors propis diferents, T és diagonalitzable.*

DEMOSTRACIÓ. En aquest cas, $\det(xI-T)$ es descompon totalment en producte de n factors de primer grau:

$$\det(xI-T)=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)\,.$$

Com que cada m_i és igual a 1, el segon dels lemes previs assegura que $d_i=1=m_i$ per a cada $i=1,\ldots,n$, per tant T és diagonalitzable. $\square$

Observacions 6.7 (a) Si considerem endomorfismes $T:\mathbb{C}^n\longrightarrow\mathbb{C}^n$ (espai vectorial numèric complex de dimensió n), la condició necessària i suficient de diagonalització es redueix a l'apartat (b) del teorema, ja que, segons el teorema fonamental de l'àlgebra, la condició (a) sempre es complirà.

(b) Tant el teorema com el corol·lari permeten determinar si un endomorfisme és diagonalitzable *sense trobar els vectors propis*, ja que, una vegada descompost el polinomi característic, o bé té n arrels diferents, o bé només cal calcular els rangs de $T-\alpha_i I$ i comprovar si es compleix la igualtat

$$m_i=n-\operatorname{rang}(T-\alpha_i I)\quad\text{per a}\quad i=1,\ldots,k\,.$$

(De fet, només cal comprovar-la quan $m_i>1$.)

Exemples 6.8 (a) Ni els girs en el pla ni les rotacions a l'espai admeten vectors propis si l'angle és $\alpha\neq 0,\pi$; per tant no són endomorfismes diagonalitzables.

(b) Si la matriu d'un endomorfisme T de $\mathbb{R}^2$ en la base canònica és simètrica, T és diagonalitzable. El lector pot comprovar directament aquesta propietat, que és un cas particular del teorema espectral que demostrarem a continuació.

Un endomorfisme T de $\mathbb{R}^n$ és *simètric* si compleix que

$$T(u) \cdot v = u \cdot T(v) \qquad \text{per a qualssevol} \quad u, v \in \mathbb{R}^n .$$

Si A és la matriu de T en la base canònica (o en qualsevol base ortonormal), T és simètric si, i només si, A és simètrica. En efecte: si X i Y són les columnes de les components de u i v en aquesta base, la condició de simetria és equivalent a que

$$(AX)^t Y = X^t AY , \quad \text{és a dir,} \quad X^t A^t Y = X^t AY ,$$

i això és cert per a qualssevol X i Y si, i només si, $A^t = A$.

Lema 6.9 *Vectors propis d'un endomorfisme simètric corresponents a valors propis diferents són ortogonals.*

DEMOSTRACIÓ. Si $T(e) = \alpha e$ i $T(e') = \beta e'$, amb $\alpha \neq \beta$, aplicant la condició de simetria tenim que

$$\alpha(e \cdot e') = T(e) \cdot e' = e \cdot T(e') = \beta(e \cdot e') ,$$

i d'aquí es dedueix que $(\alpha - \beta)(e \cdot e') = 0$, i per tant $e \cdot e' = 0$ ja que $\alpha - \beta \neq 0$. □

Teorema 6.10 *(Teorema espectral.) Tot endomorfisme simètric de $\mathbb{R}^n$ és diagonalitzable, i admet una base ortonormal de vectors propis.*

DEMOSTRACIÓ. Sigui T l'endomorfisme i A la seva matriu (simètrica) en la base canònica. En primer lloc veurem que $\det(xI - T)$ es descompon totalment sobre $\mathbb{R}$. Considerem l'endomorfisme $T^* : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ definit per la matriu A en la base canònica de $\mathbb{C}^n$. Es obvi que $\det(xI - T) = \det(xI - T^*)$, i sabem que si $\alpha \in \mathbb{C}$ és una arrel amb multiplicitat m d'aquest polinomi amb coeficients reals també ho és el seu conjugat $\overline{\alpha}$.

Per veure que α ha de tenir la part imaginària nul·la, escollim un vector propi $u \in \mathbb{C}^n$ de l'endomorfisme T^* amb valor propi α. Com que A té tots els coeficients reals, $\overline{u}$ també és un vector propi de T^* amb valor propi $\overline{\alpha}$. Aleshores, si U és la matriu columna de les components de u, tenim que $\overline{U}^t U$ és un nombre real i positiu, per tant $\overline{U}^t U = (\overline{U}^t U)^t = U^t \overline{U}$. Pel mateix motiu, $\overline{U}^t AU = (\overline{U}^t AU)^t = U^t A\overline{U}$ i, en conseqüència,

$$\alpha\overline{U}^t U = \overline{U}^t \alpha U = \overline{U}^t AU = U^t A\overline{U} = U^t \overline{\alpha}\overline{U} = \overline{\alpha} U^t \overline{U} = \overline{\alpha}\overline{U}^t U ,$$

d'on es desprèn que $\alpha = \overline{\alpha}$ i per tant $\alpha \in \mathbb{R}$.

Finalment, comprovarem per inducció sobre n que existeix una base ortonormal de vectors propis de T. Si $n = 1$, no hi ha res a dir. Si l'enunciat és cert per al cas $n-1$, seleccionem un vector propi unitari e_n de T amb valor propi α i escrivim

$$\mathbb{R}^n = H \oplus \langle e_n \rangle ,$$

on $H = \langle e_n \rangle^{\perp}$. Si $u \in H$ es compleix que $T(u) \in H$, perquè

$$T(u) \cdot e_n = u \cdot T(e_n) = \alpha u \cdot e_n = 0 .$$

Per tant, T es restringeix a un endomorfisme, encara simètric, de H. Com que H és un espai vectorial de dimensió $n-1$, per hipòtesi d'inducció existeix una base ortonormal de H, diguem $\{e_1, e_2, \ldots, e_{n-1}\}$, formada per vectors propis de T.

Ara és evident que $\{e_1, e_2, \ldots, e_{n-1}, e_n\}$ és una base ortonormal de $\mathbb{R}^n$ formada per vectors propis de T, que per tant és diagonalitzable. $\square$

Exemple 6.11 Estudiem l'endomorfisme T de $\mathbb{R}^3$ definit en la base canònica per la matriu

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

Pel teorema espectral, T és diagonalitzable. El polinomi característic és

$$\det(xI - A) = x^3 - 6x^2 + 9x - 4 = (x-1)^2(x-4) ,$$

i els valors propis són 1 (doble) i 4 (simple). Els subespais de vectors propis són

$$\mathrm{Nuc}(T - I) = \langle\, (1,0,-1),\ (0,1,-1)\, \rangle \quad \text{i} \quad \mathrm{Nuc}(T - 4I) = \langle\, (1,1,1)\, \rangle .$$

El vector $u_3 = (1,1,1)$ és ortogonal als vectors $u_1 = (1,0,-1)$ i $u_2 = (0,1,-1)$, com assegura el lema previ al teorema espectral, però aquests dos últims no ho són entre ells i cal aplicar el mètode de Gram-Schmidt, prenent $v_1 = u_1$ i

$$v_2 = u_2 - t v_1 , \quad \text{amb} \quad t = \frac{u_1 \cdot u_2}{u_1 \cdot u_1} = \frac{1}{2} ,$$

és a dir, $v_2 = 2u_2 - u_1 = (-1, 2-1)$ per evitar fraccions. Afegint $v_3 = u_3$, els vectors

$$v_1 = (1,0,-1) , \quad v_2 = (-1,2,-1) \quad \text{i} \quad v_3 = (1,1,1)$$

formen una base ortogonal de vectors propis de T. Normalitzant-los obtenim els vectors

$$w_1 = \frac{1}{\sqrt{2}}(1,0,-1) , \quad w_2 = \frac{1}{\sqrt{6}}(-1,2,-1) \quad \text{i} \quad w_3 = \frac{1}{\sqrt{3}}(1,1,1) ,$$

que formen una base ortonormal $\mathcal{B}'$ de vectors propis de T. La matriu del canvi és

$$C = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{3} & -1 & \sqrt{2} \\ 0 & 2 & \sqrt{2} \\ -\sqrt{3} & -1 & \sqrt{2} \end{pmatrix}$$

(ortogonal: $C^{-1} = C^t$), i la matriu de l'endomorfisme T en la base $\mathcal{B}'$ és

$$A' = C^{-1}AC = C^tAC = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 4 \end{pmatrix}.$$

6.3 *Diagonalització de matrius*

Traduirem ara la teoria de diagonalització al llenguatge matricial. Es diu que una matriu quadrada A és *diagonalitzable* si existeix una matriu regular C tal que $A' = C^{-1}AC$ sigui diagonal.

De cara a la pràctica, cal notar que les columnes de C seran els vectors propis de l'endomorfisme representat per A, i la diagonal de A' estarà ocupada pels valors propis disposats en el mateix ordre que s'hagi assignat als vectors propis.

Si la matriu A és de coeficients reals, serà convenient distingir dos casos. (1) Direm que A és *diagonalitzable sobre els nombres reals* si podem escollir com a matriu de canvi C una matriu de coeficients reals (i per tant A' també ho serà). (2) Direm, en canvi, que A és *diagonalitzable sobre els nombres complexos* si hem de prendre C (i per tant A') amb coeficients en general complexos. En les aplicacions, aquesta segona opció potser només serà admissible en determinats contextos.

Finalment, observem que si B és una matriu *conjugada* de A, és a dir, si $B = D^{-1}AD$ amb D regular, i A és diagonalitzable, aleshores B també ho serà, ja que A i B representen un mateix endomorfisme en bases diferents.

Exemples 6.12 (a) El polinomi característic de la matriu

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

és $\det(xI - A) = (x-1)^2(x-2)$. Tot i que aquest polinomi es descompon totalment, la matriu no és diagonalitzable, perquè

$$A - I = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \operatorname{rang}(A - I) = 2$$

i, per tant, $m_1 = 2$ però $d_1 = 3 - \text{rang}(A - I) = 1$.

(b) El polinomi característic de la matriu

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

és $\det(xI - B) = (x-1)(x+1)(x-2)^2$. Per saber si la matriu és diagonalitzable, només cal trobar $\text{rang}(B - 2I)$: com que

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

tenim que $\text{rang}(B - 2I) = 2$ i $d_2 = 4 - 2 = 2 = m_2$. Per tant, B és diagonalitzable.

(c) En el cas de la matriu

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

apliquem el corol·lari perquè

$$\det(xI - C) = x(x-1)(x-3),$$

i per tant hi ha tres valors propis diferents, condició suficient per afirmar que C és diagonalitzable.

(d) El polinomi característic de la matriu

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

és $\det(xI - D) = x^3 + x = x(x - i)(x + i)$. Aquesta matriu no és diagonalitzable sobre $\mathbb{R}$, perquè el polinomi característic no es descompon totalment en factors de primer grau, mentre que és diagonalitzable sobre $\mathbb{C}$ perquè presenta tres valors propis diferents.

(e) (*Teorema espectral generalitzat.*) *Tota matriu simètrica és diagonalitzable amb un canvi ortogonal.*

En efecte: podem suposar que la matriu representa un endomorfisme de $\mathbb{R}^n$ en la base canònica, i aplicar el teorema espectral original. Com que la base de vectors propis es pot escollir ortonormal, la matriu C del canvi de base serà ortogonal, és a dir, complirà que $C^t C = I$ (fins i tot podem fer que $\det(C) = 1$ si volem que el canvi conservi l'orientació). Per tant tenim

$$A' = C^{-1} AC = C^t AC\,.$$

Aquest resultat és fonamental per a l'estudi de tensors, formes quadràtiques, còniques i quàdriques.

Exemple 6.13 (*Càlcul de les potències d'una matriu diagonalitzable.*) Si A és una matriu quadrada, les seves potències naturals A^m seguiran una llei de recurrència que, en general, és molt difícil de trobar directament. En el cas que A sigui diagonalitzable, disposem d'un mètode alternatiu passant per la matriu diagonal conjugada $A' = C^{-1}AC$. En efecte: si

$$A' = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}$$

(amb $\alpha_1, \alpha_2, \ldots, \alpha_n$ no necessàriament diferents), és immediat comprovar que

$$(A')^m = \begin{pmatrix} {\alpha_1}^m & & & \\ & {\alpha_2}^m & & \\ & & \ddots & \\ & & & {\alpha_n}^m \end{pmatrix} \qquad \text{per a tot } m \in \mathbb{N}.$$

D'altra banda, de la relació $A' = C^{-1}AC$ obtenim

$$A = CA'C^{-1}\,,$$

$$A^2 = CA'C^{-1}CA'C^{-1} = C(A')^2C^{-1}$$

i en general, per inducció,

$$A^m = C(A')^m C^{-1}\,,$$

fórmula que resoldrà la qüestió. Com a il·lustració, considerem la matriu

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

El polinomi característic és $\det(xI - A) = (x-1)(x-4)$, i per tant A és diagonalitzable. Calculant una base de vectors propis, obtenim

$$A' = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \qquad \text{i} \qquad C = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix},$$

i aplicant la fórmula de retorn tenim que

$$\begin{aligned} A^m &= \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4^m \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{3} \begin{pmatrix} 2+4^m & -2+2\cdot 4^m \\ -1+4^m & 1+2\cdot 4^m \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix} + \frac{4^m}{3} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \qquad \forall\, m \in \mathbb{N}\,. \end{aligned}$$

6.4 *Classificació de les isometries*

Al final del capítol anterior hem deixat pendent l'estudi general de les isometries, que ara podem dur a terme aplicant, en part, la teoria de la diagonalització. A la primera part de la secció classificarem les isometries del pla, i a la segona les de l'espai.

Isometries del pla

En el pla, les úniques isometries possibles són els girs i les simetries axials (la identitat i la simetria central són, de fet, girs d'angles respectius 0 i π). A més, el determinant les classifica: una isometria T és un gir si $\det(T) = 1$, i és una simetria axial si $\det(T) = -1$. Suposarem el pla orientat per tal de descriure els girs unívocament.

Sigui T una isometria del pla orientat $\mathbb{R}^2$ definida en la base canònica per la matriu

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Aquesta matriu compleix que $A^t = A^{-1}$, on

$$A^t = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

i $\det(A) = \pm 1$. Estudiem aquests dos casos per separat.

Cas 1. Si $\det(A) = 1$, l'anterior igualtat és

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \qquad \text{d'on} \qquad \begin{cases} d = a \\ c = -b\,. \end{cases}$$

Aleshores la matriu és de la forma

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{amb} \quad a^2 + b^2 = 1$$

i, en conseqüència, existeix un únic $\alpha \in [0, 2\pi)$ tal que $a = \cos\alpha$ i $b = \sin\alpha$. En definitiva, la matriu de T en la base canònica és

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

i T és un *gir d'angle* α.

Com a casos particulars tenim la *identitat* $T = I$ quan $\alpha = 0$ i la *simetria central* $T = -I$ quan $\alpha = \pi$.

Cas 2. Si $\det(A) = -1$, tenim que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -d & c \\ b & -a \end{pmatrix}, \qquad \text{d'on} \qquad \begin{cases} d = -a \\ c = b\,. \end{cases}$$

Aleshores

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \quad \text{amb} \quad a^2 + b^2 = 1\,.$$

Aquesta matriu és simètrica i, per tant, diagonalitzable en una base ortonormal $\mathcal{B}' = \{e_1', e_2'\}$. El seu polinomi característic és $\det(xI - A) = x^2 - 1 = (x-1)(x+1)$, de manera que la matriu de T en la base $\mathcal{B}'$ és

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

T és, doncs, la *simetria axial* d'eix

$$\langle e_1' \rangle = \mathrm{Nuc}(T - I)\,.$$

Observacions 6.14 (a) També podem distingir els girs de les simetries a partir del següent fet: si la matriu A és simètrica, T és la identitat, una simetria axial o la simetria central; si A no és simètrica, T és un gir.

(b) Els girs tals que $\alpha \neq 0, \pi$ són les úniques isometries del pla que no tenen vectors propis.

Isometries de l'espai

Estudiem ara les isometries de l'espai orientat $\mathbb{R}^3$. A diferència de les isometries del pla, les de l'espai sempre tenen vectors propis, ja que el seu polinomi característic, en ser de grau 3, té com a mínim una arrel real. El resultat següent ens dóna més informació.

Lema 6.15 *Sigui T és una isometria de $\mathbb{R}^3$.*

(a) Si $\det(T) = 1$, T *admet el valor propi 1.*

(b) Si $\det(T) = -1$, T *admet el valor propi* -1.

DEMOSTRACIÓ. Sigui A la matriu de T en la base canònica.

(a) La matriu $A - I$ compleix que

$$A - I = A - A^t A = (I - A^t)A \,,$$

i per tant

$$\det(A - I) = \det(I - A^t)\det(A) = \det(I - A) = -\det(A - I)\,,$$

ja que aquests determinants són d'ordre 3. Així doncs, $\det(A - I) = 0$ i T admet el valor propi 1.

(b) La matriu $A + I$ compleix que

$$A + I = A + A^t A = (I + A^t)A \,,$$

de forma que

$$\det(A + I) = \det(I + A^t)\det(A) = -\det(I + A)\,.$$

Per tant, $\det(A + I) = 0$ i T admet el valor propi -1. □

Tal com hem fet en el pla, distingirem dos casos segons el determinant de la isometria.

Cas 1. Si T és una isometria de $\mathbb{R}^3$ amb determinant 1, sigui w un vector no nul amb valor propi 1, i considerem la descomposició de l'espai

$$\mathbb{R}^3 = \langle\, w \,\rangle \oplus H \,, \quad \text{on} \quad H = \langle\, w \,\rangle^{\perp} \,.$$

Si $u \in H$, tenim que $T(u) \cdot w = T(u) \cdot T(w) = u \cdot w = 0$, és a dir, $T(u) \in H$, i la restricció de T a H és una isometria plana. D'altra banda, el vector w determina una orientació sobre el pla H i, com que aquesta restricció no pot ser una simetria perquè el determinant de T és 1, ha de ser un gir d'angle α. Aleshores, si $\{u, v\}$ és una base ortonormal positiva de H, la matriu de T en la base $\mathcal{B} = \{w, u, v\}$ de $\mathbb{R}^3$ és

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} ,$$

i T és una *rotació d'angle* α *al voltant del vector* w.

Observacions 6.16 (a) L'angle α depèn de T, de l'orientació de l'espai i del sentit del vector propi w escollit. En particular, la rotació d'angle α al voltant de w coincideix amb la rotació d'angle $-\alpha$ al voltant de $-w$.

(b) Com a casos particulars tenim la *identitat* $T = I$ quan $\alpha = 0$ i la *simetria axial* respecte a la recta $\langle w \rangle = \mathrm{Nuc}(T - I)$ quan $\alpha = \pi$.

Cas 2. Si T és una isometria de $\mathbb{R}^3$ amb determinant -1, sigui w un vector no nul amb valor propi -1, i considerem la descomposició de l'espai

$$\mathbb{R}^3 = \langle\, w \,\rangle \oplus H\,, \quad \text{on} \quad H = \langle\, w \,\rangle^{\perp}\,.$$

El mateix raonament d'abans ens permet assegurar que la restricció de T a H també és un gir d'angle α. Aleshores, si $\{u, v\}$ és una base ortonormal positiva de H, la matriu de T en la base $\mathcal{B} = \{w, u, v\}$ és

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix},$$

i T és la *composició de la rotació d'angle α al voltant del vector w amb la simetria plana respecte a H*. A més, aquesta composició és *commutativa*, i cal notar que l'eix de rotació i el pla de simetria són ortogonals.

Observació 6.17 Com a casos particulars tenim la *simetria central* $T = -I$ quan $\alpha = \pi$ i la *simetria plana* respecte al subespai $H = \mathrm{Nuc}(T - I)$ quan $\alpha = 0$.

Podem resumir els resultats obtinguts en aquest apartat amb l'esquema de classificació següent. Si la matriu A representa la isometria T en la base canònica o en qualsevol base ortonormal positiva de $\mathbb{R}^3$, aleshores:

Exemples 6.18 (a) Considerem la isometria T de $\mathbb{R}^3$ que en la base canònica ve definida per la matriu

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

És immediat comprovar que aquesta matriu és ortogonal, no simètrica, que $\det(A) = 1$ i que $\mathrm{Nuc}(T - I) = \langle\, (1, 1, 1) \,\rangle$. Per tant, es tracta d'una *rotació* al voltant del vector $(1, 1, 1)$.

Per obtenir l'angle α de rotació hem de calcular $\sin\alpha$ i $\cos\alpha$. Apliquem la fórmula

$$T(u) = u\cos\alpha + (1 - \cos\alpha)(u \cdot w)w + (w \wedge u)\sin\alpha$$

$\det(A)$	A simètrica	T
1	Si	Identitat Simetria axial respecte al subespai $\mathrm{Nuc}(T-I)$
1	No	Rotació al voltant del vector $w \in \mathrm{Nuc}(T-I)$
-1	Si	$-$Identitat o simetria central Simetria plana respecte al subespai $\mathrm{Nuc}(T-I)$
-1	No	Composició d'una rotació al voltant de $w \in \mathrm{Nuc}(T+I)$ i una simetria plana respecte al subespai $\langle\, w\,\rangle^{\perp}$

Taula 6.1: Classificació de les isometries a l'espai

amb $w = \frac{1}{\sqrt{3}}(1,1,1)$ i $u = (1,0,0)$:

$$(0,1,0) = (1,0,0)\cos\alpha + (1-\cos\alpha)\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}(1,1,1) + \frac{1}{\sqrt{3}}(0,1,-1)\sin\alpha\,.$$

Igualant les components tenim les relacions

$$\left.\begin{aligned} \cos\alpha + \frac{1-\cos\alpha}{3} &= 0 \\ \frac{1-\cos\alpha}{3} + \frac{\sin\alpha}{\sqrt{3}} &= 1 \\ \frac{1-\cos\alpha}{3} - \frac{\sin\alpha}{\sqrt{3}} &= 0 \end{aligned}\right\}\,,$$

amb les quals determinem les raons trigonomètriques de l'angle α buscat:

$$\cos\alpha = -\frac{1}{2} \qquad \text{i} \qquad \sin\alpha = \frac{\sqrt{3}}{2}\,.$$

Per tant, T és la rotació de 120° al voltant del vector $(1,1,1)$: aquests són els dos elements geomètrics de la isometria.

Hi ha un mètode alternatiu per determinar l'angle α d'una rotació. Es basa en el fet que la traça de la matriu d'un endomorfisme és invariant pels canvis de base, perquè és un coeficient del polinomi característic. Comparant la matriu A que tenim amb la matriu de la rotació en una base adaptada a l'eix resulta la relació

$$1 + 2\cos\alpha = \mathrm{tr}(A)\,.$$

Quan $\cos\alpha \neq \pm 1$, aquesta equació admet dues solucions $\alpha_1, \alpha_2 \in [0, 2\pi)$, tals que

$$0 < \alpha_1 < \pi \qquad \text{i} \qquad \pi < \alpha_2 < 2\pi\,.$$

Si w és el vector que defineix el sentit de l'eix de rotació, escollim un vector no nul $u \perp w$, calculem la seva imatge $v = T(u)$ i considerem la base $\{w, u, v\}$: si és positiva, $\alpha = \alpha_1$; si és negativa, $\alpha = \alpha_2$. (Si $\cos\alpha = \pm 1$, $\alpha = 0$ ó π sense ambigüitat.)

A l'exemple que acabem d'estudiar tenim

$$1 + 2\cos\alpha = 0\,,$$

$\cos\alpha = -1/2$ i, per tant, $\alpha_1 = 120^\circ$ i $\alpha_2 = 240^\circ$. Prenent $w = (1, 1, 1)$ i $u = (1, -1, 0)$ resulta $v = T(u) = (0, 1, -1)$. Com que

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{vmatrix} = 3 > 0$$

deduim que $\alpha = \alpha_1 = 120^\circ$.

(b) Considerem ara l'endomorfisme S de $\mathbb{R}^3$ definit per

$$B = M(S, \mathcal{B}_c) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,.$$

B és ortogonal i simètrica però $\det(B) = -1$, per tant S és una simetria plana. El pla de simetria és el subespai $\mathrm{Nuc}(S - I)$, definit pel sistema d'equacions

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\,,$$

que es redueix a l'equació $x - y = 0$. Aquest és el pla de simetria, i l'únic element geomètric que cal determinar en aquest cas.

6.5 Sistemes d'equacions diferencials lineals

Una aplicació interessant de la diagonalització de matrius és la resolució de sistemes d'equacions diferencials lineals de primer ordre amb coeficients constants.

Una *equació diferencial de primer ordre en forma normal* és una relació

$$y' = f(x, y)$$

entre una variable independent x, una variable y que és funció de x i la derivada corresponent

$$y' = \frac{dy}{dx}\,.$$

El problema consisteix a trobar les *solucions*, és a dir, les funcions $y = y(x)$ que satisfan la relació anterior per a tot x d'un cert interval.

Exigint condicions a la funció f apareixen diversos tipus d'equacions diferencials. Si es demana que f sigui *lineal en* y, ens trobem amb l'*equació diferencial lineal*

$$y' = P(x)y + Q(x)\,.$$

Si $P(x) = \alpha$ és una constant tenim l'*equació diferencial lineal de primer ordre amb coeficient constant*

$$y' = \alpha y + Q(x)\,,$$

on suposarem que Q és una funció almenys contínua en un interval $I \subset \mathbb{R}$. Si, a més, $Q(x) = 0$, resulta l'anomenada equació *homogènia*

$$y' = \alpha y\,,$$

la més senzilla de totes però també la més freqüent en les primeres aplicacions pràctiques.

Teorema 6.19 *(a) Les solucions de l'equació homogènia de primer ordre*

$$y' = \alpha y$$

són totes les funcions de la forma

$$y(x) = Ce^{\alpha x} \quad \text{amb} \quad C \in \mathbb{R}\,.$$

(b) Una solució particular de l'equació completa

$$y' = \alpha y + Q(x)$$

és la funció

$$y_p(x) = e^{\alpha x} \int Q(x) e^{-\alpha x} \, dx \, .$$

(c) La solució general de l'equació completa és

$$y = y_p(x) + C e^{\alpha x} \quad \textit{amb} \quad C \in \mathbb{R} \, .$$

DEMOSTRACIÓ. (a) Representem per D l'acció de derivar. La condició $D(y) = \alpha y$ implica que

$$D(ye^{-\alpha x}) = D(y) e^{-\alpha x} - \alpha y e^{-\alpha x} = 0 \, ,$$

de manera que la funció $ye^{-\alpha x}$ és constant sobre $\mathbb{R}$ i resulta

$$ye^{-\alpha x} = C \, , \qquad \text{o bé} \quad y(x) = C e^{\alpha x} \quad \text{amb} \quad C \in \mathbb{R} \, .$$

Recíprocament, tota funció d'aquesta forma és solució de l'equació homogènia.

(b) Provem una funció de la forma

$$y_p(x) = C(x) e^{\alpha x}$$

com a solució de l'equació completa, suposant $C(x)$ derivable (*mètode de variació de les constants*). Derivant,

$$y_p'(x) = C'(x) e^{\alpha x} + \alpha C(x) e^{\alpha x} \, ,$$

i substituint a l'equació completa i simplificant resulta

$$C'(x) e^{\alpha x} = Q(x) \, , \qquad \text{és a dir,} \quad C(x) = \int Q(x) e^{-\alpha x} \, dx \, ,$$

expressió que ens dóna la solució particular $y_p(x)$ indicada a l'enunciat.

(c) Si $y = y(x)$ és una solució de l'equació completa, $y - y_p$ ho és de l'homogènia, per tant

$$y - y_p = C e^{\alpha x} \qquad \text{i} \qquad y(x) = y_p(x) + C e^{\alpha x} \quad \text{amb} \quad C \in \mathbb{R} \, .$$

Recíprocament, tota funció d'aquesta forma és solució de l'equació completa. □

Exemple 6.20 (*Condicions inicials.*) La constant arbitrària C que apareix en la solució general d'una equació de primer ordre es determina imposant a la funció solució una condició addicional, que acostuma a ser del tipus

$$y = y_0 \qquad \text{quan} \qquad x = x_0 \, ,$$

on x_0 i y_0 són dades (és freqüent prendre $x_0 = 0$, i d'aquí ve el nom de "condicions inicials").

Com a il·lustració, resoldrem l'equació diferencial

$$y' = 2y + x$$

amb les condicions inicials $y = 1/2$ quan $x = 0$.

(a) L'equació homogènia associada és $y' = 2y$, i la seva solució general $y = Ce^{2x}$, amb $C \in \mathbb{R}$.

(b) Una solució particular de l'equació completa serà de la forma $y_p = C(x)e^{2x}$, on

$$C(x) = \int xe^{-2x}\,dx = -\left(\frac{x}{2} + \frac{1}{4}\right)e^{-2x}$$

i per tant

$$y_p = -\frac{x}{2} - \frac{1}{4}\,.$$

(c) La solució general de l'equació proposada és la família

$$y = Ce^{2x} - \frac{x}{2} - \frac{1}{4} \quad \text{amb} \quad C \in \mathbb{R}\,.$$

(d) Imposant les condicions inicials exigides,

$$\frac{1}{2} = C - \frac{1}{4}\,, \qquad C = \frac{3}{4}$$

i queda com a solució única del problema la funció

$$y = \frac{3e^{2x} - 2x - 1}{4}\,.$$

Un *sistema d'equacions diferencials de primer ordre* per a dues variables x, y que depenen d'una variable independent t és un parell d'equacions de la forma

$$\left.\begin{aligned} \dot{x} &= \frac{dx}{dt} = F(x, y, t) \\ \dot{y} &= \frac{dy}{dt} = G(x, y, t) \end{aligned}\right\}\,.$$

Les *solucions* seran parelles de funcions $x = x(t)$, $y = y(t)$ que amb les seves derivades satisfacin les dues equacions per a tot t d'un cert interval.

Es pot donar a la variable t el significat de temps, i entendre x i y com les coordenades d'un punt mòbil del pla: en aquesta interpretació cinemàtica del sistema, el vector velocitat $v = (\dot{x}, \dot{y})$ està relacionat amb la posició (x, y) i el temps t.

El sistema és *lineal* si les funcions F i G ho són respecte x i y conjuntament, i amb *coeficients constants* si ho són els d'aquestes variables, quedant de la forma

$$\left.\begin{array}{l} \dot{x} = ax + cy + f(t) \\ \dot{y} = bx + dy + g(t) \end{array}\right\} ,$$

on suposem que f i g són almenys contínues en un interval $I \subset \mathbb{R}$. Si les funcions f i g són nul·les, el sistema és *homogeni*.

Matricialment, un sistema d'equacions diferencials lineals amb coeficients constants es pot representar per

$$\dot{X} = AX + H(t) ,$$

on

$$X = \begin{pmatrix} x \\ y \end{pmatrix} , \qquad \dot{X} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} , \qquad A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad \text{i} \qquad H(t) = \begin{pmatrix} f(t) \\ g(t) \end{pmatrix} .$$

Teorema 6.21 *Si la matriu A és diagonalitzable (sobre $\mathbb{R}$ o sobre $\mathbb{C}$), el sistema d'equacions diferencials*

$$\dot{X} = AX + H(t)$$

es pot resoldre fent el canvi de variables $X = CU$ que diagonalitza A, ja que condueix a un sistema format per dues equacions lineals de primer ordre resolubles separadament.

DEMOSTRACIÓ. Com que A és diagonalitzable (sobre $\mathbb{C}$ si és necessari), existirà una matriu C tal que $D = C^{-1}AC$ serà diagonal. Introduint el vector de noves variables U definit per $X = CU$, tenim que

$$U = C^{-1}X , \qquad \dot{U} = C^{-1}\dot{X} \qquad \text{i} \qquad C^{-1}A = DC^{-1} .$$

Aleshores

$$\dot{U} = C^{-1}\dot{X} = C^{-1}[AX + H(t)] = DC^{-1}X + C^{-1}H(t) = DU + C^{-1}H(t) .$$

Si posem

$$U = \begin{pmatrix} u \\ v \end{pmatrix} , \qquad D = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \qquad \text{i} \qquad C^{-1}H(t) = \begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} ,$$

aquest sistema s'escriu

$$\left.\begin{array}{l} \dot{u} = \alpha u + \varphi(t) \\ \dot{v} = \beta v + \psi(t) \end{array}\right\} .$$

Una vegada resoltes separadament aquestes dues equacions i obtingudes les solucions $u = u(t)$ i $v = v(t)$, desfent el canvi per $X = CU$ tindrem la solució general del sistema original, donada per $x = x(t)$ i $y = y(t)$. □

Exemples 6.22 (a) Considerem el sistema següent i la matriu que el defineix:

$$\left.\begin{aligned} \dot{x} &= x + 2y - 4t - 3 \\ \dot{y} &= 2x + y - 2t - 1 \end{aligned}\right\} \quad \text{i} \quad A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

El polinomi característic és $\det(xI - A) = x^2 - 2x - 3$, i els valors propis 3 i -1. Calculant els vectors propis obtenim la matriu del canvi que diagonalitza A i el sistema en les noves variables:

$$C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad \text{i} \qquad \left.\begin{aligned} \dot{u} &= 3u - 3t - 2 \\ \dot{v} &= -v - t - 1 \end{aligned}\right\}.$$

Resolent per separat les equacions del nou sistema i desfent el canvi trobem la solució general del sistema original:

$$\left.\begin{aligned} u &= Ae^{3t} + t + 1 \\ v &= Be^{-t} - t \end{aligned}\right\}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} Ae^{3t} + t + 1 \\ Be^{-t} - t \end{pmatrix}$$

i, finalment,

$$\left.\begin{aligned} x(t) &= Ae^{3t} + Be^{-t} + 1 \\ y(t) &= Ae^{3t} - Be^{-t} + 2t + 1 \end{aligned}\right\}.$$

(b) (*Condicions inicials*.) Les condicions inicials habituals per a un sistema d'equacions diferencials de primer ordre són de la forma

$$x = x_0 \quad \text{i} \quad y = y_0 \qquad \text{quan} \qquad t = t_0\,,$$

on x_0, y_0 i t_0 (sovint és $t_0 = 0$) són dades. Això permet determinar les constants que apareixen en la solució general i escollir la solució adequada.

Com a il·lustració, considerem un mòbil del pla que a l'instant $t = 0$ es troba en el punt $(2, 0)$ i en tot moment es regeix per les equacions diferencials de l'exemple (a).

Imposant les condicions inicials, tenim que $A = 0$ i $B = 1$, constants que defineixen la solució única o trajectòria del móbil

$$x = e^{-t} + 1\,, \qquad y = -e^{-t} + 2t + 1\,.$$

(c) Un fluid que es mou en pla s'anomena *estacionari* quan la velocitat de cada partícula no depèn més que de la posició (i no del temps): altrament dit, si en passar per cada punt totes les partícules porten la mateixa velocitat, donada pel *camp vectorial* que assigna a cada punt un vector. Les solucions d'un tal sistema

$$\left.\begin{aligned} \dot{x} &= F(x, y) \\ \dot{y} &= G(x, y) \end{aligned}\right\}$$

són trajectòries $x = x(t)$, $y = y(t)$ que en cada punt tenen un vector tangent prefixat.

Considerem per exemple el sistema

$$\left.\begin{array}{l} \dot{x} = y \\ \dot{y} = -x \end{array}\right\} .$$

A cada punt, el camp vectorial és ortogonal al vector de posició i de la mateixa longitud: la descripció gràfica suggereix que les partícules del fluid seguiran órbites tancades al voltant de l'origen en sentit negatiu (el de les agulles del rellotge). La resolució del sistema ho confirmarà.

La matriu que defineix el sistema i el seu polinomi característic són

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \text{i} \qquad \det(xI - A) = x^2 + 1 = (x - i)(x + i) \,,$$

per tant és necessari diagonalitzar sobre $\mathbb{C}$: la matriu del canvi i el nou sistema són

$$C = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \qquad \text{i} \qquad \left.\begin{array}{l} \dot{u} = iu \\ \dot{v} = -iv \end{array}\right\} .$$

Resolent aquest nou sistema i desfent el canvi surt la solució general complexa del sistema original:

$$\left.\begin{array}{l} u = Ae^{it} \\ v = Be^{-it} \end{array}\right\} \qquad \text{i} \qquad \left.\begin{array}{l} x = Ae^{it} + Be^{-it} \\ y = iAe^{it} - iBe^{-it} \end{array}\right\} \qquad \text{amb } A\,, B \in \mathbb{C}\,.$$

Imposant que x i y tinguin valors reals per a tot t surt $B = \overline{A}$, i escrivint $A = C + Di$ queda, incloent els signes en les constants,

$$\left.\begin{array}{l} x = C\cos t + D\sin t \\ y = D\cos t - C\sin t \end{array}\right\} \qquad \text{amb } C\,, D \in \mathbb{R}\,.$$

Afegint ara com a condicions inicials $x = r$ i $y = 0$ quan $t = 0$, resulta que $C = r$ i $D = 0$. Aleshores, les equacions de la trajectòria del fluid, que confirmen la predicció feta al principi a partir del camp vectorial, són

$$x = r\cos t\,, \qquad y = -r\sin t\,,$$

i representen circumferències de centre l'origen i radi r recorregudes en el sentit de les agulles del rellotge.

Molt probablement, el lector ja haurà notat que el concepte de sistema lineal de primer ordre amb coeficients constants i el teorema de resolució són òbviament generalitzables a més de dues variables dependents de t.

7

Tensors i formes quadràtiques

Pensant en les aplicacions a la física i al càlcul infinitesimal, dediquem aquest capítol a l'exemple més interessant de l'àlgebra tensorial: els tensors covariants simètrics de segon ordre —que aquí anomenarem simplement tensors— i les formes quadràtiques associades.

7.1 Introducció

Abans de donar la definició formal de les dues nocions bàsiques que volem considerar, proposarem per a cada una d'elles una situació on es presenta de manera natural.

Exemple 7.1 (*La forma hessiana d'una funció.*) Considerem una funció real de dues variables $z = f(x, y)$ i un punt $p = (x_0, y_0)$ del seu domini. Denotem les derivades parcials segones de f en el punt p per

$$a = \left(\frac{\partial^2 f}{\partial x^2}\right)_p , \qquad b = \left(\frac{\partial^2 f}{\partial y \partial x}\right)_p = \left(\frac{\partial^2 f}{\partial x \partial y}\right)_p \qquad \text{i} \qquad c = \left(\frac{\partial^2 f}{\partial y^2}\right)_p .$$

L'aplicació $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ que a cada vector $u = (x, y)$ —interpretat com un increment a partir del punt p— li assigna

$$q(u) = q(x, y) = ax^2 + 2bxy + cy^2$$

és l'*hessiana* de la funció f en el punt p. El seu estudi condueix a un criteri per seleccionar els extrems locals de la funció f d'entre els seus punts estacionaris (vegeu la darrera secció d'aquest capítol).

En general, una *forma quadràtica* és una aplicació $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ definida per una expressió polinòmica homogènia de segon grau en les variables $x_1, x_2, \ldots, x_n$:

$$q(u) = q(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

(on sempre es pot suposar, i així ho farem, que $a_{ij} = a_{ji}$). Observem que si $A = (a_{ij})$ i $X = (x_1\ x_2\ \cdots\ x_n)^t$ tenim que

$$q(u) = q(x_1, x_2, \ldots, x_n) = X^t A X .$$

És immediat de comprovar que $q(tu) = t^2 q(u)$.

Exemple 7.2 (*El tensor d'inèrcia*.) Suposem que una partícula de massa m es mou a l'espai fent una rotació al voltant de l'eix $\langle w \rangle$ amb velocitat angular $\| w \|$. Si r és el vector de posició de la partícula, sabem que la seva velocitat és $v = w \wedge r$.

Es defineix el *moment cinètic* de la partícula a cada posició r per

$$M(w) = mr \wedge v = mr \wedge (w \wedge r) .$$

La transformació M és un endomorfisme simètric de $\mathbb{R}^3$. Definim el *tensor d'inèrcia* de la partícula a la posició r per

$$T(u, v) = u \cdot M(v) .$$

T dóna la mateixa informació que M, però és més útil en determinades qüestions. Per exemple, l'energia cinètica de la partícula es pot escriure com

$$E = \frac{1}{2} m \| v \|^2 = \frac{1}{2} T(w, w) .$$

(En el cas d'un sòlid rígid compost per un continu de partícules, cal introduir a tot arreu el símbol d'integració.)

En general, un *tensor* (un tensor covariant simètric de segon ordre, per ser precisos) és una aplicació bilineal simètrica $T : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$. *Bilineal* vol dir lineal respecte a cada una de les dues variables vectorials, i això significa que tenim

$$T(u + u', v) = T(u, v) + T(u', v) \qquad \text{i} \qquad T(tu, v) = tT(u, v)$$

i dues relacions anàlogues quan fixem u en comptes de v. *Simètrica* vol dir que $T(u, v) = T(v, u)$.

Teorema 7.3 *(Caracterització dels tensors.) Una aplicació $T : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ és un tensor si, i només si, es pot representar matricialment per*

$$T(u, v) = X^t AY \,,$$

on A és una matriu simètrica.

DEMOSTRACIÓ. El raonament és molt semblant al del teorema de caracterització de les transformacions lineals del capítol 5, així que el deixarem com a exercici per al lector interessat. □

Observacions 7.4 (a) A la demostració anterior es comprova que, si $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ és una base de $\mathbb{R}^n$, els escalars

$$a_{ij} = T(e_i, e_j)$$

determinen totalment T, poden ser escollits arbitràriament sempre que $a_{ij} = a_{ji}$ (simetria) i són els coeficients de la matriu A, que és la *matriu de T en la base $\mathcal{B}$*. Simbolitzarem aquesta última relació per $A = M(T, \mathcal{B})$. Així doncs, si X i Y són les columnes respectives de components de u i v en la base $\mathcal{B}$, tenim que

$$T(u, v) = X^t AY = \sum_{i,j=1}^{n} a_{ij} x_i y_j \,.$$

En general, els coeficients diagonals a_{ii} s'anomenen *components principals* del tensor, i els coeficients no diagonals a_{ij} *components secundàries*. En el cas del tensor d'elasticitat (origen del mot *tensor*) es parla, respectivament, de "tensions normals" i "tensions tangencials".

(b) (*Forma quadràtica associada*.) Donat el tensor T, definim la *forma quadràtica associada* a T com l'aplicació $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ donada per

$$q(u) = T(u, u) \,.$$

Amb les notacions de l'observació anterior tenim que

$$q(u) = X^t AX \,,$$

i és immediat comprovar la propietat següent, dita *igualtat de polarització*, que permet afirmar que tota forma quadràtica q està associada a un únic tensor T:

$$q(u + v) = q(u) + q(v) + 2T(u, v) \,.$$

(c) (*Fórmula del canvi de base*.) Si $\mathcal{B}'$ és una altra base de $\mathbb{R}^n$, $A' = M(T, \mathcal{B}')$ i C és la matriu del canvi de base de $\mathcal{B}'$ a $\mathcal{B}$, aleshores es compleix la relació

$$A' = C^t AC \,.$$

(Novament, deixem la comprovació a les mans del lector, que pot inspirar–se en el cas dels endomorfismes tractat en el capítol 5.) Des del punt de vista físic, és habitual pensar en un tensor com un ens descrit a cada base per una matriu simètrica A que, en fer un canvi de base, varia seguint la fórmula anterior. Escrita en components, la fórmula s'expressa per

$$a'_{ij} = \sum_{k,\ell=1}^{n} c_{\ell i} a_{\ell k} c_{kj} \quad \text{per a} \quad i, j = 1, 2, \ldots, n\,.$$

(d) (*Restricció d'un tensor a un subespai*.) Suposem que T és un tensor definit sobre $\mathbb{R}^n$, $\mathcal{B}$ és una base de $\mathbb{R}^n$ i $A = M(T, \mathcal{B})$. Suposem també que H és un subespai del qual hem escollit una base $\mathcal{B}' = \{u_1, u_2, \ldots, u_k\}$, i que D és la matriu (de tamany $n \times k$) que relaciona $\mathcal{B}'$ amb $\mathcal{B}$. Aleshores, la *restricció* de T al subespai H, que es defineix de manera òbvia, queda descrita en la base $\mathcal{B}'$ per la matriu

$$A' = D^t A D\,,$$

perquè aquest producte matricial calcula els coeficients $a'_{ij} = T(u_i, u_j)$ de la matriu A'. Com a il·lustració, restringirem el tensor T de $\mathbb{R}^3$ definit per

$$A = M(T, \mathcal{B}_c) = \begin{pmatrix} 6 & 2 & 4 \\ 2 & -2 & -2 \\ 4 & -2 & 0 \end{pmatrix}$$

al pla d'equació $2x + 3y - z = 0$. Una base d'aquest pla és $\mathcal{B}' = \{(1, 0, 2), (0, 1, 3)\}$, i la matriu de la restricció de T en aquesta base és

$$A' = D^t A D = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 6 & 2 & 4 \\ 2 & -2 & -2 \\ 4 & -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 22 & 10 \\ 10 & -14 \end{pmatrix}.$$

La tècnica de restricció és útil per a l'estudi dels extrems condicionats de les funcions de diverses variables.

Exemples 7.5 (a) Si la posició d'una partícula de massa m referida a una base ortonormal positiva de l'espai és $r = (x, y, z)$, la matriu del tensor d'inèrcia de la partícula en aquesta posició és

$$A = m \begin{pmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{pmatrix}.$$

En el cas d'un sistema format per un nombre finit de partícules de masses m_k i posicions respectives $r_k = (x_k, y_k, z_k)$, la matriu és

$$A = \begin{pmatrix} \sum_k m_k (y_k^2 + z_k^2) & -\sum_k m_k x_k y_k & -\sum_k m_k x_k z_k \\ -\sum_k m_k x_k y_k & \sum_k m_k (x_k^2 + z_k^2) & -\sum_k m_k y_k z_k \\ -\sum_k m_k x_k z_k & -\sum_k m_k y_k z_k & \sum_k m_k (x_k^2 + y_k^2) \end{pmatrix}.$$

Els coeficients de la diagonal principal són els *moments d'inèrcia*, i els restants s'anomenen *productes d'inèrcia*.

(b) La presentació geomètrica de la teoria de la relativitat restringida fa servir un espai vectorial real, l'espai–temps $E_4 = V_3 \oplus T_1$, on V_3 és l'espai ordinari (de les posicions) i T_1 és l'eix temporal. La formulació de la teoria es fonamenta en el *tensor de Minkowski*, que en una base $\{i, j, k, h\}$ adaptada a la suma directa queda definit per la matriu

$$A = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -c^2 \end{pmatrix},$$

on c és la velocitat de la llum. L'equació de la forma quadràtica associada al tensor de Minkowski en la mateixa base és

$$q(u) = q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2 ,$$

i la cinemàtica de la relativitat restringida equival essencialment a l'estudi de les "isometries"de E_4, endomorfismes que deixen invariant el tensor de Minkowski i reben el nom de transformacions de Lorentz.

La mecànica relativista dels medis continus demostra que el tensor de l'energia mecànica d'un fluid perfecte és un múltiple escalar del tensor de Minkowski.

7.2 *Diagonalització de tensors i formes*

Exposarem ara un dels processos de diagonalització de tensors i formes quadràtiques.

Teorema 7.6 *Si T és un tensor arbitrari sobre $\mathbb{R}^n$, existeixen bases $\mathcal{B}'$ de $\mathbb{R}^n$ on la matriu de T és diagonal i, per tant, la forma quadràtica associada q s'expressa com a combinació lineal de quadrats:*

$$A' = M(T, \mathcal{B}') = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \quad i \quad q(x_1', x_2', \ldots, x_n') = \alpha_1(x_1')^2 + \alpha_2(x_2')^2 + \cdots + \alpha_n(x_n')^2 .$$

DEMOSTRACIÓ. La matriu A de T en una base $\mathcal{B}$ de $\mathbb{R}^n$ és simètrica. Aplicant el teorema espectral generalitzat, existeix una matriu ortogonal de canvi de base C tal que

$$A' = C^{-1}AC = C^tAC$$

és diagonal. □

Observacions 7.7 (a) Si només considerem bases ortonormals (és a dir, fem canvis ortogonals), el resultat final del procés —la matriu A'— és únic llevat de l'ordre en què disposem els valors propis. L'equació corresponent per a la forma quadràtica associada,

$$q(x_1', x_2', \ldots, x_n') = \alpha_1 (x_1')^2 + \alpha_2 (x_2')^2 + \cdots + \alpha_n (x_n')^2 \,,$$

s'anomena l'*expressió euclidiana canònica* de q.

(b) A efectes pràctics, recordem que els coeficients de la diagonal de A' són els valors propis de la matriu A, i que les columnes de la matriu de canvi C són els vectors propis ortonormals de A escrits en el mateix ordre que els valors propis.

Exemple 7.8 Considerem un cub rígid i homogeni, de costat 1 i densitat 12, que gira al voltant d'un eix que passa per un dels seus vèrtexs, i suposem que els vectors de la base ortonormal positiva de referència $\mathcal{B}$ segueixen les direccions de les arestes del cub. Aleshores, la matriu del tensor d'inèrcia del sòlid en la base $\mathcal{B}$ és

$$A = \begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix} .$$

El polinomi característic és $\det(tI - A) = t^3 - 24t^2 + 165t - 242$, i els valors propis són 2 (simple) i 11 (doble). Escollim com a vectors propis respectius

$$u_1 = \frac{1}{\sqrt{3}}(1, 1, 1), \quad u_2 = \frac{1}{\sqrt{2}}(1, -1, 0) \quad \text{i} \quad u_3 = \frac{1}{\sqrt{6}}(1, 1, -2) .$$

En la base ortonormal positiva $\mathcal{B}' = \{u_1, u_2, u_3\}$, la matriu del tensor d'inèrcia del cub és

$$A' = \begin{pmatrix} 2 & & \\ & 11 & \\ & & 11 \end{pmatrix} .$$

Els valors propis són els *moments principals d'inèrcia* del cub, i les direccions definides pels vectors u_1, u_2, u_3 són els *eixos principals d'inèrcia*. Si l'eix de rotació és un d'aquests eixos aleshores no hi ha productes d'inèrcia, i això significa que la figura es troba en *equilibri dinàmic*. L'expressió euclidiana canònica de la forma quadràtica associada és

$$q(x', y', z') = 2x'^2 + 11y'^2 + 11z'^2 .$$

Si la figura gira, per exemple, al voltant del vector $w = u_1 + 2u_2 + 2u_3$, amb velocitat angular 3, la seva energia cinètica en qualsevol moment és

$$E = \frac{1}{2} q(1, 2, 2) = 45 .$$

Aplicant factors correctors adequats als vectors de la base on un tensor T diagonalitza, sempre podem fer que els coeficients de la matriu diagonal de T siguin exclusivament 1, -1 ó 0. A l'exemple anterior, si prenem com a base

$$v_1 = \frac{1}{\sqrt{6}}(1,1,1), \quad v_2 = \frac{1}{\sqrt{22}}(1,-1,0) \quad \text{i} \quad v_3 = \frac{1}{\sqrt{66}}(1,1,-2)$$

queda com a matriu del tensor

$$A' = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix},$$

i com a expressió de la forma quadràtica associada

$$q(x', y', z') = x'^2 + y'^2 + z'^2 .$$

Observem que hem sacrificat la normalització dels vectors.

7.3 *Índexs d'inèrcia i classificació de formes*

Hem vist que, donat un tensor T sobre $\mathbb{R}^n$, existeixen bases on la matriu de T és diagonal, i apareixen r coeficients positius, s coeficients negatius i $t = n - r - s$ zeros a la diagonal principal. De fet, hi ha altres procediments de diagonalització, i no condueixen a un mateix resultat.

Exemple 7.9 Considerem la forma quadràtica definida per

$$q(x, y, z) = 6x^2 + 4xy - 2y^2 + 2xz + 2yz .$$

En la base formada pels vectors

$$u_1 = \frac{1}{\sqrt{6}}(1,0,0), \qquad u_2 = \frac{1}{\sqrt{24}}(1,-3,0) \qquad \text{i} \qquad u_3 = (-1,1,4)$$

la matriu del tensor associat és

$$\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix},$$

mentre que els valors propis són $2+\sqrt{22}$, $2-\sqrt{22}$ i 0. No importa la manera com hem obtingut els vectors u_1, u_2, u_3: la qüestió és que existeixen diferents matrius diagonals d'un mateix tensor. Aleshores, el resultat següent esdevé essencial.

Teorema 7.10 *(Llei d'inèrcia de Sylvester.) Els nombres r i s no depenen de la base escollida per diagonalitzar el tensor T.*

DEMOSTRACIÓ. Sigui $\mathcal{B} = \{u_1, \ldots, u_r, v_1, \ldots, v_s, w_1, \ldots, w_t\}$ una base on la matriu de T és diagonal, i suposem que, a la diagonal, els primers r coeficients són positius, els s següents negatius i els t finals zeros. Això ens dóna una descomposició de l'espai en suma directa:

$$\mathbb{R}^n = P_r \oplus N_s \oplus R_t\,,$$

on $T(u, v) = 0$ si u i v pertanyen a subespais diferents i, a més,

$$q(u) > 0 \quad \forall u \in P_r, \qquad q(v) < 0 \quad \forall v \in N_s \qquad \text{i} \qquad q(w) = 0 \quad \forall w \in R_t\,.$$

Sigui $\mathcal{B}'$ una altra base on la matriu de T és diagonal, ordenada de forma anàloga. D'entrada,

$$R_t = \{u \in \mathbb{R}^n : T(u, v) = 0 \;\; \forall v \in \mathbb{R}^n\} = R'_{t'}\,,$$

per tant, $t = t'$ i $r + r' = s + s'$.

Comprovem ara que la projecció de $\mathbb{R}^n$ sobre P_r paral·lelament a $N_s \oplus R_t$ és injectiva quan es restringeix a $P'_{r'}$. En efecte: si $0 \neq x \in P'_{r'}$, tenim $x = u + v + w$ amb $u \in P_r$, $v \in N_s$ i $w \in R_t$; si fós $u = 0$, resultaria que

$$0 < q(x) = q(v + w) = q(v) + q(w) + 2T(v, w) = q(v) < 0\,,$$

absurd. Així doncs, $r' \leq r$. Per la simetria del raonament, $r = r'$, i per tant també $s = s'$. □

Els nombres r i s són els *índexs d'inèrcia* de T (i de q). La llei d'inèrcia ens diu que l'expressió d'una forma quadràtica en qualsevol base on la matriu sigui diagonal amb coeficients 1, -1 i 0 és única, i queda justificat denominar–la l'*expressió canònica afí* de la forma. Una manera de calcular els índexs d'inèrcia, i d'escriure per tant l'expressió canònica afí de la forma, és comptar els signes dels valors propis. El teorema de Descartes (vegeu l'apèndix B) sembla fet a mida per a aquesta finalitat, ja que per ser simètrica la matriu A que representa a T sabem *a priori* que el seu polinomi característic té totes les arrels reals, i podem determinar fàcilment r i s.

Exemple 7.11 Considerem la forma quadràtica de l'exemple anterior:

$$q(x, y, z) = 6x^2 + 4xy - 2y^2 + 2xz + 2yz\,.$$

El polinomi característic és $\det(tI - A) = t^3 - 4t^2 - 18t = 0$. Aquí hi ha una variació de signe, i per tant un únic valor propi positiu; com que 0 és un valor propi simple, hi ha un tercer valor propi negatiu. L'expressió canònica afí de q és

$$q(x', y', z') = x'^2 - y'^2\,.$$

Es diu que una forma quadràtica q és *definida positiva* (respectivament, *definida negativa*) si $q(u) > 0$ (resp., $q(u) < 0$) per a tot $u \neq 0$. Es diu que q és *semidefinida positiva* si $q(u) \geq 0$ per a tot $u \in \mathbb{R}^n$ (amb algun u tal que $q(u) > 0$) però existeix algun $v \neq 0$ tal que $q(v) = 0$ —el lector proveïrà la definició de *semidefinida negativa*. Finalment, q és *no definida* si existeixen $u, v \in \mathbb{R}^n$ tals que $q(u) > 0$ i $q(v) < 0$. L'única forma quadràtica que escapa, doncs, a aquesta classificació és la *forma nul·la*. A continuació donem diversos criteris per classificar una forma quadràtica arbitrària.

Teorema 7.12 *Els índexs d'inèrcia r i s d'una forma quadràtica q definida sobre $\mathbb{R}^n$, calculables a partir dels valors propis $\lambda_1, \lambda_2, \ldots, \lambda_n$, permeten classificar–la segons l'esquema de la taula 7.1.*

Tot $\lambda_i > 0$	$r = n$	$s = 0$	definida positiva
Tot $\lambda_i < 0$	$r = 0$	$s = n$	definida negativa
Tot $\lambda_i \geq 0$, algun $\lambda_k > 0$ i algun $\lambda_j = 0$	$0 < r < n$	$s = 0$	semidefinida positiva
Tot $\lambda_i \leq 0$, algun $\lambda_k < 0$ i algun $\lambda_j = 0$	$r = 0$	$0 < s < n$	semidefinida negativa
Algun $\lambda_i > 0$ i algun $\lambda_j < 0$	$r > 0$	$s > 0$	no definida
Tot $\lambda_i = 0$	$r = 0$	$s = 0$	forma nul·la

Taula 7.1: Classificació de formes quadràtiques

DEMOSTRACIÓ. Només cal pensar en l'expressió canònica afí de q, amb la qual es fan paleses les sis possibilitats, i en la relació entre els índexs d'inèrcia i els signes dels valors propis. $\square$

Exemples 7.13 (a) El tensor d'inèrcia d'una partícula dóna sempre una forma quadràtica semidefinida positiva, perquè el polinomi característic és

$$t^3 - 2mr^2t^2 + m^2r^4t = 0\,,$$

on $r^2 = x^2 + y^2 + z^2$, i els valors propis són mr^2 (doble) i 0 (simple).

(b) En canvi, si tenim tres partícules de masses respectives 2, 1 i 1, situades a les posicions (1,1,0), (1,0,1) i (0,1,1), la matriu del tensor d'inèrcia del sistema és

$$A = \begin{pmatrix} 5 & -2 & -1 \\ -2 & 5 & -1 \\ -1 & -1 & 6 \end{pmatrix}.$$

El polinomi característic és $\det(tI - A) = t^3 - 16t^2 + 79t - 112$, de manera que la forma quadràtica associada és definitiva positiva perquè $r = 3$ i $s = 0$.

Teorema 7.14 *(Criteri de Sylvester.) Si A és la matriu del tensor T en alguna base, i els nombres $D_1, D_2, \ldots, D_n$ són els menors diagonals creixents de la matriu $A = (a_{ij})$, és a dir,*

$$D_k = \det(A_k)\,, \qquad \text{on} \qquad A_k = (a_{ij}) \quad \text{per a} \quad 1 \le i,\, j \le k\,,$$

aleshores:

(a) q és definida positiva si, i només si, $D_1, D_2, \ldots, D_n > 0$.

(b) q és definida negativa si, i només si, $D_k > 0$ quan k és parell i $D_k < 0$ quan k és senar.

DEMOSTRACIÓ. (a) Denotem per $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ la base on la matriu de T és A.

($\Longrightarrow$) Per a cada $k = 1, 2, \ldots, n$, la restricció de q al subespai $\langle\, e_1, e_2, \ldots, e_k \,\rangle$ és definida positiva i la matriu és A_k; en alguna base d'aquest subespai la matriu serà diagonal,

$$A'_k = \begin{pmatrix} t_1 & & \\ & \ddots & \\ & & t_k \end{pmatrix} \qquad \text{amb} \qquad t_1, \ldots, t_k > 0\,,$$

de manera que $\det(A'_k) > 0$ i, per la fórmula del canvi de base, $D_k = \det(A_k) > 0$.

($\Longleftarrow$) Suposem que q no és definida positiva i considerem el més petit dels subespais $H_k = \langle\, e_1, e_2, \ldots, e_k \,\rangle$ on la restricció de q no és definida positiva. Aquesta restricció diagonalitza en una base $\langle\, e'_1, e'_2, \ldots, e'_k \,\rangle$ de H_k amb matriu A'_k, i no tots els $t_1, \ldots, t_k$ poden ser positius. Ara bé, si existissin $t_i, t_j \le 0$, la restricció de q al pla $\langle\, e'_i, e'_j \,\rangle$ seria definida (o semidefinida) negativa o nul·la, en contradicció amb el fet que $\langle\, e'_i, e'_j \,\rangle \cap H_{k-1} \neq \{0\}$. Per tant, existeix un únic $t_i \le 0$, i $D_k = \det(A_k) \le 0$ en contradicció amb les hipòtesis.

(b) El raonament és molt semblant en els dos casos.

($\Longrightarrow$) La diagonal de A'_k constarà de $t_1, \ldots, t_k < 0$, de manera que $D_k = \det(A_k) > 0$ si k és parell i $D_k = \det(A_k) < 0$ si k és senar.

($\Longleftarrow$) Aquí q no seria definida negativa en algun H_k i, anàlogament al cas (a), resultaria que $D_k \leq 0$ si k és parell o $D_k \geq 0$ si k és senar, en contradicció amb les hipòtesis. □

Corol·lari 7.15 *Si $n = 2$, considerem la matriu del tensor associat T, així com el determinant i la traça d'aquesta matriu:*

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \qquad D = \det(A) = ac - b^2 \qquad i \qquad T = tr(A) = a + c\,.$$

Aleshores tenim el criteri de classificació següent:

$D > 0$	$a > 0$	definida positiva
$D > 0$	$a < 0$	definida negativa
$D = 0$	$T > 0$	semidefinida positiva
$D = 0$	$T < 0$	semidefinida negativa
$D < 0$		no definida
$D = 0$	$T = 0$	forma nul·la

Taula 7.2: Classificació de formes quadràtiques ($n = 2$)

DEMOSTRACIÓ. Si $D > 0$, ha de ser $a \neq 0$, i podem aplicar el criteri de Sylvester. Si $D = 0$, un valor propi és 0 i l'altre T, i apliquem el criteri dels valors propis. Finalment, si $D < 0$ els valors propis són de signe diferent i fem servir el mateix criteri. □

7.4 Extrems locals de les funcions de diverses variables

Aquesta secció es dedica a una aplicació interessant de la classificació de formes. Suposem que el lector ha seguit —o està seguint— un curs de càlcul infinitesimal que proporciona els coneixements sobre diferenciabilitat dels que farem ús.

Sigui $f : D \longrightarrow \mathbb{R}$ una funció real de diverses variables definida en un domini $D \subset \mathbb{R}^n$ i de classe $\mathcal{C}^2$ en algun entorn U d'un punt $p \in D$, entorn que podem suposar una bola oberta de centre p.

Es diu que f té un *màxim* (respectivament, *mínim*) *local* en el punt p si $f(p) \geq f(x)$ (resp., $f(p) \leq f(x)$) per a tot x d'un cert entorn de p. En tots dos casos direm que f té un *extrem local* a p.

Si f té un extrem local a p, és *condició necessària* que s'anul·li la diferencial en aquell punt,

$$(df)_p = 0\,,$$

o, equivalentment, que totes les derivades parcials de f en el punt p siguin nul·les:

$$\left(\frac{\partial f}{\partial x_i}\right)_p = 0 \quad \text{per a} \quad i = 1, 2, \ldots, n$$

(es diu que p és un *punt estacionari* de f).

D'altra banda, la *fórmula de Taylor* aplicada a f en el punt p, amb terme complementari de segon grau, és

$$f(p+t) = f(p) + (df)_p(t) + \tfrac{1}{2}\left(d^2 f\right)_q(t,t)$$

per a $t \in \mathbb{R}^n$ suficientment petit en norma per tal que $p + t \in U$, i essent q un punt del segment que uneix p amb $p + t$. Així com df és un *camp* que en cada punt, diguem p, dóna una *forma lineal* $(df)_p$ —és a dir, una transformació lineal de $\mathbb{R}^n$ a $\mathbb{R}$—, $d^2 f$ és un *camp tensorial* de segon ordre, que en cada punt p dóna un tensor $(d^2 f)_p$ (simètric pel teorema de Schwartz), ja que els coeficients de la matriu d'aquest tensor en la base canònica són les derivades parcials segones de f en el punt:

$$a_{ij} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_p\,.$$

Escrivint la fórmula de Taylor en termes de les derivades parcials tenim

$$f(p+t) = f(p) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_p t_i + \frac{1}{2} \sum_{i,j=1}^{n} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_q t_i t_j\,,$$

on $t = (t_1, \ldots, t_n)$.

Si suposem que f satisfà la condició necessària d'extrem local en el punt p, la fórmula es redueix a

$$f(p+t) - f(p) = \tfrac{1}{2}(d^2 f)_q(t,t) = \frac{1}{2} \sum_{i,j=1}^{n} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_q t_i t_j\,.$$

Per a obtenir *condicions suficients* d'extrem local és necessari estudiar el signe d'aquest increment prop del punt p. El resultat següent dóna un criteri força útil, i per justificar-lo només necessitem observar que la condició $\mathcal{C}^2$ exigida a f assegura que les derivades parcials segones són contínues.

Teorema 7.16 *(Condicions suficients d'extrem local.) Sigui $f : D \longrightarrow \mathbb{R}$ una funció definida en un domini $D \subset \mathbb{R}^n$ i de classe $\mathcal{C}^2$ en un entorn U d'un punt $p \in D$. Suposem que $(df)_p = 0$. Aleshores:*

(a) Si $(d^2 f)_p$ és definida positiva, f té un mínim local en el punt p.

(b) Si $(d^2 f)_p$ és definida negativa, f té un màxim local en el punt p.

(c) Si $(d^2 f)_p$ és no definida, f no té extrem local en el punt p.

DEMOSTRACIÓ. (a) Si $(d^2 f)_p$ és definida positiva, per les condicions que reflecteixen aquest fet —desigualtats, básicament— i la continuïtat de les derivades parcials segones de f, $(d^2 f)_q$ seguirà essent definida positiva en tot un entorn de p, i retornant a la fórmula de l'increment escrita abans d'aquest teorema arribem a la conclusió que $f(p+t) - f(p) \geq 0$ per a tot punt $p+t$ de l'entorn; per tant, f presenta un mínim local en el punt p.

(b) L'argument és, *mutatis mutandis*, idèntic.

(c) Si $(d^2 f)_p$ és no definida, existeixen vectors $u, v \in \mathbb{R}^n$ tals que

$$(d^2 f)_p(u,u) > 0 \qquad \text{i} \qquad (d^2 f)_p(v,v) < 0\,.$$

Considerant la funció $g(\lambda) = f(p+\lambda u)$ resulta que

$$g'(0) = (df)_p(u) = 0 \qquad \text{i} \qquad g''(0) = (d^2 f)_p(u,u) > 0$$

i, per tant, f té, segons la direcció de u, un mínim local en el punt p; però un raonament anàleg per a la funció $h(\lambda) = f(p+\lambda v)$ ens diu que f té un màxim local en el punt p segons la direcció de v. En definitiva, f no té extrem local en el punt p. □

Observem que si $(d^2 f)_p$ és semidefinida o nul·la, el teorema anterior no és aplicable, i cal fer un estudi basat en altres mètodes. I també que per classificar $(d^2 f)_p$ podem fer ús, segons convingui, de qualsevol dels criteris establerts a la secció anterior.

Exemples 7.17 (a) Considerem $f(x,y) = x^2 - y^2$. Busquem els punts que satisfan la condició necessària $df = 0$, és a dir,

$$\frac{\partial f}{\partial x} = 2x = 0 \qquad \text{i} \qquad \frac{\partial f}{\partial y} = -2y = 0\,,$$

i trobem $p = (0, 0)$. La matriu hessiana —matriu de les derivades parcials segones en aquest punt— és

$$H(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} .$$

És evident, doncs, que $(d^2 f)_p$ és no definida i, per tant, f no té extrem local en el punt $(0, 0)$: es diu que hi ha un *punt de sella* (màxim en una direcció, quan $x = 0$, i mínim en una altra, quan $y = 0$).

(b) Considerem ara la funció $z = f(x, y) = (x^2 + y^2 - 2x)(x^2 + y^2 - 6x)$. Seguint el mateix procediment, tenim com a condiciones necessàries

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= (4x - 8)(x^2 + y^2) - 16x^2 + 24x = 0 \\ \frac{\partial f}{\partial y} &= 4y(x^2 + y^2 - 4x) = 0 \end{aligned}\right\} .$$

De la segona equació resulta que $y = 0$ o bé $x^2 + y^2 - 4x = 0$. (1) Si $y = 0$, la primera equació es converteix en $4x(x^2 - 6x + 6) = 0$, d'on

$$x = 0 , \qquad x = 3 + \sqrt{3} \qquad \text{ó} \qquad x = 3 - \sqrt{3} .$$

(2) Si $x^2 + y^2 - 4x = 0$, tenim que $x^2 + y^2 = 4x$ i la primera equació es transforma en $-8x = 0$, per tant $x = 0$ i $y = 0$.

Resumint, caldrà estudiar la matriu hessiana en els punts

$$a = (0, 0) , \qquad b = (3 + \sqrt{3}, 0) \qquad \text{i} \qquad c = (3 - \sqrt{3}, 0) .$$

Les derivades parcials segones de f són

$$\frac{\partial^2 f}{\partial x^2} = 12(x^2 - 4x + 2) + 4y^2 , \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 8y(x - 2) \qquad \text{i}$$
$$\frac{\partial^2 f}{\partial y^2} = 4(3y^2 + x^2 - 4x) .$$

En el punt b, la matriu hessiana és

$$H(b) = \begin{pmatrix} 24(\sqrt{3} + 1) & 0 \\ 0 & 8\sqrt{3} \end{pmatrix} ,$$

i per tant f presenta un *mínim local*, d'altura $-36 - 24\sqrt{3}$.

En el punt c, la matriu hessiana és

$$H(c) = \begin{pmatrix} 24(1-\sqrt{3}) & 0 \\ 0 & -8\sqrt{3} \end{pmatrix},$$

i per tant f presenta un *màxim local*, d'altura $-36 + 24\sqrt{3}$.

Hem deixat el punt a per al final, ja que

$$H(a) = \begin{pmatrix} 24 & 0 \\ 0 & 0 \end{pmatrix}$$

i el teorema no decideix perquè $H(a)$ és semidefinida positiva. Cal dur a terme un estudi particular, que en aquest cas queda resolt de la forma següent. El primer factor s'anul·la sobre la circumferència $x^2 + y^2 - 2x = 0$, de centre $(1, 0)$ i radi 1, mentre que el segon ho fa sobre la circumferència $x^2+y^2-6x = 0$, de centre $(3, 0)$ i radi 3. A qualsevol entorn del punt $(0, 0)$ hi ha punts dels interiors de les dues circumferències, en els quals f prendrà valors positius, i també punts interiors a una però exteriors a l'altra, on f prendrà valors negatius. Tenint en compte que $f(0, 0) = 0$, s'arriba a la conclusió que f no presenta extrem local en el punt $(0, 0)$.

8

Geometria lineal

La percepció primària del món real es fonamenta en els punts i les figures que amb ells es formen, i no pas en els vectors. Però l'ús de coordenades per determinar la posició dels punts, i per tant d'equacions per descriure les figures elementals, converteix la geometria analítica en un camp de fecunda aplicació de l'àlgebra lineal. En aquest capítol estudiarem les figures lineals, és a dir, les definides per equacions de primer grau. Com que el lector ja coneix segurament alguns dels conceptes i tècniques que apareixen, aprofitarem la seva intuició i la seva experiència prèvia per estalviar-li un excés de formalisme analític. Presentarem la teoria només per a l'espai tridimensional; les referències al pla es reduiran a aquells aspectes en els quals l'analogia requereix algun comentari. (D'altra banda, no sembla oportú en aquest context fer distincions entre geometria estrictament afí i geometria mètrica.)

8.1 L'espai euclidià puntual tridimensional

La primera definició estableix les relacions bàsiques entre punts i vectors que cal prendre com a punt de partida.

Definició 8.1 *L'espai euclidià puntual tridimensional* és una terna $E_3 = (P_3, V_3, +)$, on P_3 és un conjunt d'elements anomenats *punts*, V_3 és l'espai vectorial euclidià tridimensional —que suposarem orientat sempre que convingui— i hi ha definida una operació de la forma

$$\begin{aligned} P_3 \times V_3 &\xrightarrow{+} P_3 \\ (p, \vec{v}) &\longmapsto p + \vec{v} \end{aligned}$$

que assigna un punt $p + \vec{v}$ a cada parell format per un punt p i un vector $\vec{v}$ (en aquest capítol i el següent els vectors duran fletxa). Es demana que es compleixin els axiomes següents:

(a) Associativitat: $(p + \vec{v}) + \vec{w} = p + (\vec{v} + \vec{w})$.

(b) Determinació: donats $p, q \in P_3$ existeix un únic $\vec{v} \in V_3$ tal que $p + \vec{v} = q$.

D'ara endavant representarem aquest vector per $\vec{v} = \overrightarrow{pq}$, i direm que q és l'*extrem* de $\vec{v}$ quan l'*origen* és p; amb aquesta nova notació,

$$p + \overrightarrow{pq} = q\,.$$

El lector probablement entén —i això és positiu— que aquesta definició representa en certa manera la construcció recíproca del procés de formació dels vectors lliures que ha vist en els seus estudis anteriors.

Un exemple —sense connotacions gràfiques si no es vol— és el següent: P_3 és el *conjunt* $\mathbb{R}^3$, V_3 és l'*espai vectorial euclidià* $\mathbb{R}^3$ i l'operació $+$ és la suma ordinària component a component: és immediat comprovar que es compleixen els axiomes.

Proposició 8.2 *Per a qualssevol $p, q, r \in P_3$ es compleix que:*

(a) $\overrightarrow{pp} = \vec{0}\,.$

(b) $\overrightarrow{qp} = -\overrightarrow{pq}\,.$

(c) $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}\,.$

DEMOSTRACIÓ. (a) Existeix un únic vector $\vec{v}$ tal que $p + \vec{v} = p$; per tant

$$p + (\vec{v} + \vec{v}) = (p + \vec{v}) + \vec{v} = p + \vec{v}\,,$$

i d'aquí resulta que $\vec{v} + \vec{v} = \vec{v}$, és a dir, $\vec{v} = \vec{0}$.

(b) Aplicant l'associativitat, tenim que

$$p + (\overrightarrow{pq} + \overrightarrow{qp}) = (p + \overrightarrow{pq}) + \overrightarrow{qp} = q + \overrightarrow{qp} = p\,,$$

i això implica que $\overrightarrow{pq} + \overrightarrow{qp} = \vec{0}$ i, per tant, $\overrightarrow{qp} = -\overrightarrow{pq}$.

(c) Novament per l'associativitat,

$$p + (\overrightarrow{pq} + \overrightarrow{qr}) = (p + \overrightarrow{pq}) + \overrightarrow{qr} = q + \overrightarrow{qr} = r\,,$$

és a dir, $\overrightarrow{pq} + \overrightarrow{qr} = \overrightarrow{pr}$. □

Una *referència cartesiana* de E_3 és una quaterna $\mathcal{R} = \{o\,;\, \vec{e}_1,\, \vec{e}_2,\, \vec{e}_3\}$, on $o \in P_3$ és un punt escollit arbitràriament, anomenat *origen* de la referència, i $\mathcal{B} = \{\vec{e}_1,\, \vec{e}_2,\, \vec{e}_3\}$ és una base de V_3. Es diu que la referència és *rectangular* si $\mathcal{B}$ és una base ortonormal, que suposarem sempre positiva. Farem servir sempre referències rectangulars, tot i que molts resultats —els que no depenen del producte escalar ni de conceptes derivats d'ell— són extrapolables a referències de qualsevol mena.

Donada una referència $\mathcal{R}$, cada punt p de P_3 es pot descriure pel seu *vector de posició* respecte a l'origen de coordenades:

$$\overrightarrow{op} = p_1\vec{e}_1 + p_2\vec{e}_2 + p_3\vec{e}_3\,.$$

Els coeficients $(p_1,\, p_2,\, p_3)$ són les *coordenades* del punt p en la referència $\mathcal{R}$, i el determinen unívocament i sense restriccions: qualsevol terna $(p_1,\, p_2,\, p_3)$ defineix un punt, concretament

$$p = o + p_1\vec{e}_1 + p_2\vec{e}_2 + p_3\vec{e}_3\,.$$

Es important traduir a coordenades dels punts i components dels vectors les dues operacions bàsiques definides entre ells: donats

$$p = (p_1,\, p_2,\, p_3)\,, \quad q = (q_1, q_2, q_3) \quad \text{i} \quad \vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + v_3\vec{e}_3 = (v_1,\, v_2,\, v_3)\,,$$

tenim les relacions

$$\vec{v} = \overrightarrow{pq} \iff p + \vec{v} = q \iff p_i + v_i = q_i \quad \text{per a} \quad i = 1, 2, 3\,.$$

Exemples 8.3 (a) Considerant les relacions entre coordenades i components, és immediat comprovar la *propietat del paral·lelogram*:

$$\overrightarrow{pq} = \overrightarrow{rs} \qquad \text{si, i només si,} \qquad \overrightarrow{pr} = \overrightarrow{qs}\,.$$

(b) El *punt mitjà* de p i q es defineix com

$$m = p + \frac{1}{2}\overrightarrow{pq}\,.$$

L'aparent asimetria de la definició desapareix en observar que les coordenades de m són les semisumes de les corresponents coordenades de p i q, d'on resulta que

$$m = q + \frac{1}{2}\overrightarrow{qp}\,.$$

(c) Anàlogament, el *baricentre* de tres punts p, q i r es defineix com

$$g = r + \frac{2}{3}\overrightarrow{rm}\,,$$

on m és el punt mitjà de p i q; aquí trobem que les coordenades de g són les mitjanes aritmètiques de les coordenades corresponents de p, q i r, i per tant els papers de p, q i r en aquesta construcció son permutables.

Fórmula del canvi de referència. Sigui $\mathcal{R}' = \{o_1;\ \vec{u}_1,\ \vec{u}_2,\ \vec{u}_3\}$ una altra referència, i suposem conegudes tant la posició del nou origen com la dependència de la nova base:

$$\overrightarrow{oo_1} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3 \qquad \text{i} \qquad \vec{u}_j = c_{1j}\vec{e}_1 + c_{2j}\vec{e}_2 + c_{3j}\vec{e}_3 \quad \text{per a} \quad j = 1, 2, 3\,.$$

La relació entre les coordenades velles (x, y, z) i les noves (x', y', z') d'un punt genèric $p \in P_3$ resulta simplement de la igualtat

$$\overrightarrow{op} = \overrightarrow{oo_1} + \overrightarrow{o_1p}$$

i és

$$\left.\begin{aligned} x &= b_1 + c_{11}x' + c_{12}y' + c_{13}z' \\ y &= b_2 + c_{21}x' + c_{22}y' + c_{23}z' \\ z &= b_3 + c_{31}x' + c_{32}y' + c_{33}z' \end{aligned}\right\}\,, \qquad \text{o bé} \qquad X = B + CX'\,,$$

on

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \qquad \text{i} \qquad C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}.$$

La matriu C és regular ($\det C \neq 0$) perquè descriu un canvi de base, ortogonal ($C^{-1} = C^t$) si les dues referències són rectangulars i amb $\det(C) = 1$ si les bases de les referències tenen la mateixa orientació. Dins d'aquest model general, són freqüents dues situacions més simples: (a) si $C = I$, la fórmula queda $X = B + X'$ i es diu que hem fet una *transl·lació d'eixos*; (b) si $B = 0$, la fórmula és $X = CX'$ i parlem de *rotació d'eixos*.

8.2 Varietats lineals

En la resta del capítol suposarem donada una referència rectangular $\mathcal{R} = \{o;\ \vec{e}_1,\ \vec{e}_2,\ \vec{e}_3\}$. Cada punt quedarà identificat per les seves coordenades i cada vector per les seves components:

$$p = (x, y, z)\,, \qquad \vec{u} = (a, b, c)\,.$$

Una *varietat lineal* de l'espai E_3 és un conjunt de punts de la forma $L = p + D$, on $p \in P_3$ i $D \subset V_3$ és un subespai vectorial, anomenat *direcció* de L. Segons sigui $\dim(D) = 1$ o $\dim(D) = 2$, es diu que L és una *recta* o un *pla*, mentre que els punts són les varietats de dimensió 0.

Equacions d'una recta. Donada una recta

$$R\ :\quad p_0 + D_1 = p_0 + \langle\, \vec{u}\, \rangle$$

es diu que $\vec{u}$ és un *vector director* de R. Qualsevol punt de R és de la forma

$$p_t = p_0 + t\vec{u} \qquad \text{amb} \quad t \in \mathbb{R}\,,$$

i això és l'*equació paramètrica vectorial* de la recta R. Si referim a $\mathcal{R}$ el punt genèric $p_t = (x, y, z)$ i també $p_0 = (x_0, y_0, z_0)$ i $\vec{u} = (a, b, c)$, l'equació origina les *equacions paramètriques* (escalars)

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}$$

de les quals, eliminant el paràmetre t, passem a

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}\,,$$

que és l'*equació contínua* de la recta R.

Equacions d'un pla. Donat un pla

$$P\ :\quad p_0 + D_2 = p_0 + \langle\, \vec{u},\ \vec{v}\, \rangle$$

qualsevol dels seus punts és de la forma

$$p = p_0 + t\vec{u} + s\vec{v} \qquad \text{amb} \quad t, s \in \mathbb{R}\,,$$

i això és l'*equació paramètrica vectorial* del pla. Els paràmetres s'eliminen escrivint la relació equivalent o *equació implícita*

$$\det(\overrightarrow{p_0 p},\ \vec{u},\ \vec{v}) = 0\,.$$

Referint a $\mathcal{R}$ el punt genèric $p = (x, y, z)$ i també $p_0 = (x_0, y_0, z_0)$, $\vec{u} = (u_1, u_2, u_3)$ i $\vec{v} = (v_1, v_2, v_3)$ resulta

$$\begin{vmatrix} x - x_0 & u_1 & v_1 \\ y - y_0 & u_2 & v_2 \\ z - z_0 & u_3 & v_3 \end{vmatrix} = 0\,,$$

que una vegada desenvolupada adopta la forma

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0\,, \quad \text{o bé} \quad Ax + By + Cz + D = 0\,.$$

Observacions 8.4 (a) Notem que l'equació contínua d'una recta condueix a un sistema de dues equacions lineals de rang 2 (equacions *implícites* de la recta) i que, recíprocament, un tal sistema defineix una recta, ja que la seva solució general ens dóna les equacions paramètriques. Geomètricament, donar el sistema significa expressar la recta com a intersecció de dos plans. També es posa de manifest que els plans i les rectes són precisament les figures definides per equacions lineals, les més senzilles, doncs, segons el criteri del grau però també les més usuals a la pràctica.

Així, donada la recta definida per

$$\left.\begin{aligned} 3x + y + 2z - 6 = 0 \\ x + y + z - 4 = 0 \end{aligned}\right\}\,,$$

resolem aquest sistema i obtenim, prenent $x = t$ com a variable lliure,

$$x = t\,, \qquad y = 2 + t\,, \qquad z = 2 - 2t\,,$$

que són les equacions paramètriques; l'equació contínua serà

$$\frac{x - 0}{1} = \frac{y - 2}{1} = \frac{z - 2}{-2}\,.$$

(b) És convenient recordar que els coeficients A, B i C de l'equació d'un pla satisfan la proporcionalitat derivada del seu origen,

$$\frac{A}{\begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix}} = \frac{B}{\begin{vmatrix} u_3 & v_3 \\ u_1 & v_1 \end{vmatrix}} = \frac{C}{\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}}\,,$$

de manera que la direcció del pla és $D_2 = \langle \vec{w} \rangle^{\perp}$, on $\vec{w} = \vec{u} \wedge \vec{v} = (A, B, C)$, i la seva equació implícita en la base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de V_3 s'obté senzillament suprimint el terme independent de l'equació de P:

$$Ax + By + Cz = 0\,.$$

(c) En el pla euclidià puntual $E_2 = (P_2, V_2, +)$ les úniques varietats lineals no trivials són les rectes, que comparteixen característiques de les rectes de l'espai (direcció de dimensió 1) i dels plans (direcció de dimensió $n-1$), i per tant admeten tres tipus d'equacions:

(1) *paramètrica*: $p_t = p_0 + t\vec{u}$,

(2) *contínua*: $\dfrac{x-x_0}{a} = \dfrac{y-y_0}{b}$,

(3) *implícita*: $Ax + By + C = 0$,

a més d'altres també usuals que suposem conegudes pel lector, com l'*explícita* $y = mx + n$ o l'equació *punt-pendent* $y - y_0 = m(x - x_0)$.

8.3 Posicions relatives de les varietats lineals

Estudiarem en aquest apartat el paral·lelisme, la intersecció i la perpendicularitat de varietats, donant condicions operatives. Inclourem també l'equació del feix de plans que passen per una recta.

Dues varietats lineals L i L', de direccions respectives D i D', són *paral·leles*, i s'escriu $L \parallel L'$, si $D \subset D'$ o $D' \subset D$.

En particular, dos plans (o dues rectes) són paral·lels si, i només si, tenen la mateixa direcció —per dimensions—, mentre que una recta i un pla són paral·lels si la direcció de la recta està continguda en la del pla. Esquemàticament tenim:

$\parallel$	R	P'
R'	$\dfrac{a}{a'} = \dfrac{b}{b'} = \dfrac{c}{c'}$	——
P	$Aa + Bb + Cc = 0$	$\dfrac{A}{A'} = \dfrac{B}{B'} = \dfrac{C}{C'}$

Taula 8.1: Paral·lelisme de varietats lineals

on $\vec{u} = (a, b, c)$ és un vector director de R, $\vec{v} = (a', b', c')$ ho és de R', $Ax + By + Cz + D = 0$ és l'equació implícita de P i $A'x + B'y + C'z + D' = 0$ és la de P'.

Analitzem ara les possibles situacions d'*incidència* (intersecció) entre varietats.

(A) Donades dues rectes

$$R: \quad p_0 + t\vec{u} \qquad \text{i} \qquad R': \quad q_0 + s\vec{v}\,,$$

la primera possibilitat és que siguin *paral·leles*, quan $\vec{u}$ i $\vec{v}$ són proporcionals; en particular, seran *coincidents* si $q_0 \in R$.

En cas contrari, els vectors $\vec{u}$ i $\vec{v}$ són linealment independents, de manera que si existeix un punt comú $p \in R \cap R'$ és únic (rectes *secants*): vegem que això és equivalent a que

$$\det(\overrightarrow{p_0q_0},\ \vec{u},\ \vec{v}) = 0\,.$$

En efecte: si $p \in R \cap R'$ tindrem que $\overrightarrow{p_0q_0} = \overrightarrow{p_0p} + \overrightarrow{pq_0} = t\vec{u} - s\vec{v}$ per a certs $t, s \in \mathbb{R}$, dependència lineal que anul·la el determinant; recíprocament, si $\overrightarrow{p_0q_0}$, $\vec{u}$ i $\vec{v}$ són linealment dependents, existirà una relació de la forma

$$\overrightarrow{p_0q_0} = t\vec{u} - s\vec{v}\,,$$

i prenent $p = p_0 + t\vec{u}$ resulta que

$$p = p_0 + t\vec{u} = p_0 + \overrightarrow{p_0q_0} + s\vec{v} = q_0 + s\vec{v} \in R'\,,$$

i per tant $p \in R \cap R'$.

Finalment, quan

$$\det(\overrightarrow{p_0q_0},\ \vec{u},\ \vec{v}) \neq 0\,,$$

les rectes no són paral·leles ni secants, i es diu que *s'encreuen* (possibilitat que no existeix entre rectes del pla).

(B) Donats una recta i un pla,

$$R: \quad p_0 + t\vec{u} \qquad \text{i} \qquad P: \quad Ax + By + Cz + D = 0\,,$$

els possibles punts de $R \cap P$ s'obtindran en substituir les equacions paramètriques de la recta

$$x = x_0 + at\,, \qquad y = y_0 + bt\,, \qquad z = z_0 + ct$$

a l'equació del pla: això dóna

$$(Aa + Bb + Cc)t + (Ax_0 + By_0 + Cz_0 + D) = 0\,.$$

Si $Aa + Bb + Cc = 0$, el vector director de la recta, $\vec{u}$, pertany a la direcció del pla, D_2, i les varietats són *paral·leles*; si, a més, $Ax_0 + By_0 + Cz_0 + D = 0$, el punt p_0 i tota la recta R es troben dins el pla, i es diu que P *passa per* R o que R està inclosa en el pla.

L'alternativa és que sigui $Aa + Bb + Cc \neq 0$: aleshores hi ha un únic punt comú $p \in R \cap P$, i es diu que les varietats són *secants*.

Si la recta ve donada per les seves equacions implícites, formem amb elles i l'equació del pla un sistema 3×3:

$$\left.\begin{array}{r} A_1x + B_1y + C_1z + D_1 = 0 \\ A_2x + B_2y + C_2z + D_2 = 0 \\ Ax + By + Cz + D = 0 \end{array}\right\} .$$

Si els rangs de la matriu dels coeficients i la matriu ampliada són r i r', tenim tres possibilitats: (a) si $r = r' = 3$, la recta i el pla són secants (solució única); (b) si $r = 2$, però $r' = 3$, són paral·lels sense inclusió (incompatibilitat); (c) si $r = r' = 2$, el pla passa per la recta (infinites solucions amb un grau de llibertat).

(C) Donats dos plans

$$P_1 : \quad A_1x + B_1y + C_1z + D_1 = 0 \qquad \text{i} \qquad P_2 : \quad A_2x + B_2y + C_2z + D_2 = 0\,,$$

considerem els rangs r i r' del sistema que formen. Hi ha també tres casos: si $r = 1$, els plans són *paral·lels*, *coincidents* si $r' = 1$ i diferents si $r' = 2$; si $r = r' = 2$, són *secants*, i els punts comuns formen una recta com ja hem dit en una observació de la secció anterior. Com que la referència és rectangular, el vector director de la recta comuna és $\vec{w}_1 \wedge \vec{w}_2$, on $\vec{w}_1 = (A_1, B_1, C_1)$ i $\vec{w}_2 = (A_2, B_2, C_2)$.

Feix de plans. Aquesta és la denominació del conjunt de plans que passen per una recta determinada R. Suposem que R està definida com a intersecció de dos plans

$$P_1 : \quad A_1x + B_1y + C_1z + D_1 = 0 \qquad \text{i} \qquad P_2 : \quad A_2x + B_2y + C_2z + D_2 = 0\,.$$

Com hem vist a l'apartat (B) d'incidència, la condició necessària i suficient per tal que un pla arbitrari P passi per R és que afegint la seva equació $Ax + By + Cz + D = 0$ al sistema format per les altres dues no canviï el conjunt de solucions (els punts de R): això significa que l'equació de P serà combinació lineal de les altres dues, i per tant

$$\alpha(A_1x + B_1y + C_1z + D_1) + \beta(A_2x + B_2y + C_2z + D_2) = 0\,,$$

variant $\alpha,\ \beta \in \mathbb{R}$, és l'equació (de tots els plans) del feix que passa per R. Una expressió equivalent, en la que només perdem P_2, resulta de dividir per $\alpha \neq 0$ i prendre $\lambda = \beta/\alpha$:

$$(A_1x + B_1y + C_1z + D_1) + \lambda(A_2x + B_2y + C_2z + D_2) = 0\,.$$

Exemple 8.5 Busquem l'equació del pla que passa pel punt $p = (1, 1, 3)$ i per la recta

$$\left.\begin{array}{r} 2x + y + z = 0 \\ x + y - 3z = 1 \end{array}\right\} .$$

Preparem l'equació del feix que passa per aquesta recta:

$$(2x + y + z) + \lambda(x + y - 3z - 1) = 0\,.$$

Imposant que un pla d'aquesta forma passi per p, determinem

$$6 + \lambda(-8) = 0\,, \qquad \lambda = \frac{3}{4}$$

i, per tant, el pla serà $(2x + y + z) + {}^{3}/_{4}(x + y - 3z - 1) = 0$ o, equivalentment,

$$4(2x + y + z) + 3(x + y - 3z - 1) = 0\,, \qquad \text{és a dir,} \qquad 11x + 7y - 5z - 3 = 0\,.$$

Dues varietats lineals L i L', de direccions respectives D i D', són *perpendiculars*, i s'escriu $L \perp L'$, si $D^{\perp} \subset D'$ o $D' \subset D^{\perp}$.

Esquemàticament tenim:

$\perp$	R	P'
R'	$aa' + bb' + cc' = 0$	——
P	$\dfrac{a}{A} = \dfrac{b}{B} = \dfrac{c}{C}$	$AA' + BB' + CC' = 0$

Taula 8.2: Perpendicularitat de varietats lineals

on $\vec{u} = (a, b, c)$ és un vector director de R, $\vec{v} = (a', b', c')$ ho és de R', $Ax + By + Cz + D = 0$ és l'equació implícita de P i $A'x + B'y + C'z + D' = 0$ és la de P'.

Teorema 8.6 *Considerem un punt p i una varietat lineal $L = p_0 + D$. Aleshores:*

(a) Existeix un únic punt $q \in L$ tal $\overrightarrow{pq} \perp D$.

(b) Existeix un únic punt s tal que q és el punt mitjà de p i s.

DEMOSTRACIÓ. (a) El punt q ha de ser de la forma $p_0 + \vec{v}$, amb $\vec{v} \in D$. La condició de perpendicularitat és

$$\overrightarrow{p_0 p} - \vec{v} \perp D\,,$$

és a dir, $\vec{v}$ ha de ser la projecció ortogonal del vector $\overrightarrow{p_0p}$ sobre el subespai D. El teorema de la projecció ortogonal assegura l'existència i unicitat d'aquest vector, per tant el punt q queda unívocament determinat.

(b) És fàcil veure que $s = p + 2\overrightarrow{pq}$ és el punt que compleix la condició de l'enunciat. $\square$

Es diu que q és la *projecció ortogonal* de p sobre la varietat L i s el *simètric* de p respecte a L.

Exemples 8.7 (a) Si R és la recta $p_0 + \langle \vec{u} \rangle$, la projecció ortogonal de $\overrightarrow{p_0p}$ sobre el subespai $\langle \vec{u} \rangle$ és

$$\vec{v} = \frac{\overrightarrow{p_0p} \cdot \vec{u}}{\| \vec{u} \|^2} \vec{u} ,$$

per tant la projecció de p sobre R és el punt

$$q = p_0 + \frac{\overrightarrow{p_0p} \cdot \vec{u}}{\| \vec{u} \|^2} \vec{u} .$$

Aquest punt també es pot obtenir fent la intersecció de la recta R i el pla perpendicular que passa per p.

(b) Si P és el pla $p_0 + \langle \vec{w} \rangle^{\perp}$, la projecció ortogonal de $\overrightarrow{p_0p}$ sobre el subespai $\langle \vec{w} \rangle^{\perp}$ és

$$\vec{v} = \overrightarrow{p_0p} - \frac{\overrightarrow{p_0p} \cdot \vec{w}}{\| \vec{w} \|^2} \vec{w} ,$$

per tant la projecció de p sobre P és el punt

$$q = p_0 + \overrightarrow{p_0p} - \frac{\overrightarrow{p_0p} \cdot \vec{w}}{\| \vec{w} \|^2} \vec{w} = p - \frac{\overrightarrow{p_0p} \cdot \vec{w}}{\| \vec{w} \|^2} \vec{w} .$$

Anàlogament al cas anterior, podem obtenir aquest punt com a intersecció del pla P i la recta perpendicular que passa per p. Com a il·lustració, donats el punt $p = (5, 3, -2)$ i el pla P d'equació $2x + y - z + 3 = 0$ busquem q i s.

La recta que passa per p amb vector director $\vec{w} = (2, 1, -1)$, ortogonal a P, és

$$x = 5 + 2t , \qquad y = 3 + t , \qquad z = -2 - t ,$$

i fent la intersecció amb el pla resulta

$$2(5 + 2t) + (3 + t) - (-2 - t) + 3 = 0 , \qquad 6t + 18 = 0 , \qquad t = -3 ,$$

de manera que obtenim la projecció ortogonal de p sobre P, que és $q = (-1, 0, 1)$. Finalment, considerant $\overrightarrow{pq} = (-6, -3, 3)$ i fent

$$s = p + 2\overrightarrow{pq} = (5, 3, -2) + 2(-6, -3, 3) = (-7, -3, 4) ,$$

trobem el simètric de p respecte a P.

8.4 Angles, distàncies, àrees i volums

Tots els *angles* (no orientats) entre varietats es redueixen a angles entre vectors. En efecte: (a) l'angle entre dues rectes és el que formen els respectius vectors directors, i això inclou la possibilitat d'obtenir–ne un o el seu suplementari; (b) l'angle entre una recta i un pla és el que forma la recta amb la seva projecció ortogonal sobre el pla (o 90° si són perpendiculars), però aquest angle coincideix amb el complementari de l'angle entre la recta i les perpendiculars al pla; (c) l'angle entre dos plans és el que formen sengles perpendiculars traçades dins de cada un a la recta intersecció (o 0° si són plans paral·lels), però coincideix amb el que forma qualsevol parell de rectes perpendiculars respectivament a cada pla. Com es pot observar, l'angle és invariant per paral·lelisme.

Exemple 8.8 Per calcular l'angle entre el pla $3x - 4y + 12z - 1 = 0$ i la recta

$$\frac{x-1}{20} = \frac{y+2}{-21} = \frac{z-3}{0},$$

considerem el vector perpendicular al pla i el vector director de la recta:

$$\vec{w} = (3, -4, 12) \qquad \text{i} \qquad \vec{u} = (20, -21, 0)\,.$$

L'angle entre aquests vectors és

$$\beta = \arccos \frac{\vec{w} \cdot \vec{u}}{\| \vec{w} \| \| \vec{u} \|} = \arccos \frac{144}{13 \cdot 29} = 67^\circ 32'\,,$$

i l'angle entre la recta i el pla és

$$\alpha = 90^\circ - 67^\circ 32' = 22^\circ 28'\,.$$

La *distància entre dos punts* p i q de P_3 és la norma del vector que defineixen:

$$d(p,q) = \| \overrightarrow{pq} \|\,.$$

Les propietats de la distància resulten de traduir les de la norma (capítol 4), i per tant són de comprovació immediata:

(a) $d(p,p) = 0$; $d(p,q) > 0$ si $p \neq q$.

(b) $d(p,q) = d(q,p)$.

(c) *Desigualtat triangular*: $d(p,r) \leq d(p,q) + d(q,r)$.

(d) *Teorema de Pitàgores*: si p, q, r són els vèrtexs d'un triangle rectangle a q, és a dir, $\overrightarrow{qp} \perp \overrightarrow{qr}$, aleshores

$$d(p,r)^2 = d(p,q)^2 + d(q,r)^2 .$$

La *distància entre dues varietats lineals* L i L' és el mínim de les distàncies entre punts de l'una i de l'altra:

$$d(L, L') = \min\{ d(p,q) : p \in L,\ q \in L' \} .$$

Aquesta definició de distància exigeix que comprovem en cada cas l'existència del mínim. Analitzarem seguidament tots els casos possibles.

Teorema 8.9 *La distància d'un punt p a una varietat lineal L és la del punt a la seva projecció ortogonal sobre L.*

DEMOSTRACIÓ. Si q és la projecció ortogonal de p sobre L i $p_0 \neq q$ és un punt qualsevol de L, aplicant el teorema de Pitàgores al triangle de vèrtexs p_0, q i p, rectangle a q, tenim

$$d(p_0, p)^2 = d(p_0, q)^2 + d(q, p)^2 ,$$

de manera que $d(p_0, p) > d(q, p)$ i per tant $d(p, L) = d(p, q)$. □

Exemple 8.10 Si $p = (x_0, y_0, z_0)$ i P és el pla d'equació $Ax + By + Cz + D = 0$,

$$d(p, P) = \frac{| Ax_0 + By_0 + Cz_0 + D |}{\sqrt{A^2 + B^2 + C^2}} .$$

En efecte: si $\vec{w} = (A, B, C)$, la projecció ortogonal q del punt p sobre P és la intersecció de la recta $p + \langle \vec{w} \rangle$, d'equacions

$$x = x_0 + At , \qquad y = y_0 + Bt , \qquad z = z_0 + Ct ,$$

amb el pla: substituint trobem

$$A(x_0 + At) + B(y_0 + Bt) + C(z_0 + Ct) + D = 0 , \qquad t = -\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2} .$$

Finalment, recordant que $\| w \|^2 = A^2 + B^2 + C^2$,

$$d(p, P) = d(p, q) = \| t\vec{w} \| = \frac{| Ax_0 + By_0 + Cz_0 + D |}{\sqrt{A^2 + B^2 + C^2}} .$$

Aquesta fórmula és aprofitable per al pla: la distància del punt $p = (x_0, y_0)$ a la recta R d'equació $Ax + By + C = 0$ és

$$d(p, R) = \frac{| Ax_0 + By_0 + C |}{\sqrt{A^2 + B^2}} .$$

La definició de distància implica que *la distància entre varietats secants és nul·la*. Queden, doncs, quatre casos possibles. En els tres referits a *varietats paral·leles* L i L' (dues rectes, dos plans o recta i pla), amb direccions respectives $D \subset D'$, la distància és la de qualsevol punt $p \in L$ a L'. En efecte: si q és la projecció ortogonal de p sobre L' i $p_1 \in L$ i $q_1 \in L'$ són punts arbitraris,

$$\overrightarrow{p_1q_1} = \overrightarrow{p_1p} + \overrightarrow{pq} + \overrightarrow{qq_1} = \overrightarrow{pq} + (\overrightarrow{p_1p} + \overrightarrow{qq_1})\,,$$

i, pel teorema de Pitàgores,

$$d(p_1, q_1) \geq d(p, q)\,,$$

és a dir, $d(L, L') = d(p, L')$ per a qualsevol $p \in L$.

La darrera situació requereix un estudi particular.

Proposició 8.11 *Donades dues rectes que s'encreuen*

$$R\ :\quad p_t = p_0 + t\vec{u}\,, \qquad i \qquad R'\ :\quad q_s = q_0 + s\vec{v}\,,$$

existeix un únic parell de punts $p_t \in R$ i $q_s \in R'$ tals que la recta que els uneix és perpendicular a R i a R'. A més,

$$d(R, R') = d(p_t, q_s)\,.$$

DEMOSTRACIÓ. Escrivim

$$\overrightarrow{p_tq_s} = \overrightarrow{p_tp_0} + \overrightarrow{p_0q_0} + \overrightarrow{q_0q_s} = -t\vec{u} + \overrightarrow{p_0q_0} + s\vec{v}$$

i imposem les condicions de perpendicular comuna,

$$\vec{u} \cdot \overrightarrow{p_tq_s} = 0\,, \qquad \text{i} \qquad \vec{v} \cdot \overrightarrow{p_tq_s} = 0\,,$$

que condueixen a un sistema lineal 2×2 per a t i s:

$$\left.\begin{aligned} -t(\vec{u}\cdot\vec{u}) + s(\vec{u}\cdot\vec{v}) + \vec{u}\cdot\overrightarrow{p_0q_0} = 0 \\ -t(\vec{v}\cdot\vec{u}) + s(\vec{v}\cdot\vec{v}) + \vec{v}\cdot\overrightarrow{p_0q_0} = 0 \end{aligned}\right\}\,.$$

El determinant del sistema és

$$\begin{vmatrix} -\vec{u}\cdot\vec{u} & \vec{u}\cdot\vec{v} \\ -\vec{v}\cdot\vec{u} & \vec{v}\cdot\vec{v} \end{vmatrix} < 0\,,$$

per la desigualtat de Schwarz. El sistema té, per tant, solució única, i els punts p_t i q_s defineixen la perpendicular comuna. Per veure que $d(p_t, q_s)$ és la mínima distància possible entre les dues rectes, considerem altres punts $p_0 \in R$ i $q_0 \in R'$: aplicant el teorema de Pitàgores a la relació

$$\overrightarrow{p_0q_0} = \overrightarrow{p_0p_t} + \overrightarrow{p_tq_s} + \overrightarrow{q_sq_0} = \overrightarrow{p_tq_s} + (\overrightarrow{p_0p_t} + \overrightarrow{q_sq_0})$$

obtenim que

$$d(p_0, q_0) > d(p_t, q_s)\,. \quad \square$$

El lector pot comprovar que, fent passar un pla P per R' que sigui paral·lel a R, es compleix que $d(R, R') = d(R, P)$, que és un dels casos previs. Més endavant veurem també una fórmula per calcular la distància entre rectes que s'encreuen, basada en el producte vectorial. De tota manera, inclourem ara una aplicació del teorema anterior.

Exemple 8.12 Calculem la distància entre les rectes donades per les equación contínues

$$R : \quad \frac{x-3}{0} = y = \frac{z+2}{-1}, \qquad \text{i} \qquad R' : \quad \frac{x}{4} = y+1 = \frac{z-5}{-3}.$$

Amb les expressions paramètriques preparem $\overrightarrow{p_t q_s}$:

$$p_t = (3, 0, -2) + t(0, 1, -1), \qquad q_s = (0, -1, 5) + s(4, 1, -3),$$

$$\overrightarrow{p_t q_s} = (-3, -1, 7) + s(4, 1, -3) - t(0, 1, -1).$$

Imposant la perpendicularitat amb $\vec{u} = (0, 1, -1)$ i $\vec{v} = (4, 1, -3)$ tenim

$$\left.\begin{aligned} -8 + 4s - 2t &= 0 \\ -34 + 26s - 4t &= 0 \end{aligned}\right\},$$

d'on resulten $t = -2$, $s = 1$, $\overrightarrow{p_{-2} q_1} = (1, 2, 2)$ i, finalment,

$$d(R, R') = \| \overrightarrow{p_{-2} q_1} \| = 3.$$

El producte vectorial resulta útil per a obtenir certes fórmules sobre distàncies, àrees i volums. Recordem que, orientant V_3 d'una de les dues maneres possibles, queda definit unívocament el producte vectorial $\vec{u} \wedge \vec{v}$ de qualsevol parell de vectors $\vec{u}, \vec{v} \in V_3$ (de fet, canviar l'orientació no influeix en cap de les fórmules).

Proposició 8.13 *(a) La distància d'un punt p a una recta $R : p_0 + \langle \vec{u} \rangle$ és*

$$d(p, R) = \frac{\| \overrightarrow{p_0 p} \wedge \vec{u} \|}{\| \vec{u} \|}.$$

(b) La distància d'un punt p a un pla $P : p_0 + \langle \vec{u}, \vec{v} \rangle$ és, posant $\vec{w} = \vec{u} \wedge \vec{v}$,

$$d(p, P) = \frac{| \overrightarrow{p_0 p} \cdot \vec{w} |}{\| \vec{w} \|} = \frac{| \overrightarrow{p_0 p} \cdot (\vec{u} \wedge \vec{v}) |}{\| \vec{u} \wedge \vec{v} \|} = \frac{| \det(\overrightarrow{p_0 p}, \vec{u}, \vec{v}) |}{\| \vec{u} \wedge \vec{v} \|}.$$

(c) La distància entre dues rectes que s'encreuen, $R : p_0 + \langle \vec{u} \rangle$ i $R' : q_0 + \langle \vec{v} \rangle$, és

$$d(R, R') = \frac{| \overrightarrow{p_0 q_0} \cdot (\vec{u} \wedge \vec{v}) |}{\| \vec{u} \wedge \vec{v} \|} = \frac{| \det(\overrightarrow{p_0 q_0}, \vec{u}, \vec{v}) |}{\| \vec{u} \wedge \vec{v} \|}.$$

DEMOSTRACIÓ. (a) Sigui q la projecció ortogonal de p sobre R: el triangle de vèrtexs p_0, q i p és rectangle a q, per tant, si α és l'angle entre $\overrightarrow{p_0p}$ i $\vec{u}$, tenim

$$d(p,R) = d(p,q) = \|\overrightarrow{p_0p}\| \sin\alpha = \frac{\|\overrightarrow{p_0p} \wedge \vec{u}\|}{\|\vec{u}\|}.$$

(b) Sigui q la projecció ortogonal de p sobre P: $\overrightarrow{qp}$ és la projecció ortogonal de $\overrightarrow{p_0p}$ sobre el subespai $\langle \vec{w} \rangle$, de manera que

$$\overrightarrow{qp} = \frac{\overrightarrow{p_0p} \cdot \vec{w}}{\|\vec{w}\|^2} w, \quad \text{d'on} \quad d(p,P) = \|\overrightarrow{qp}\| = \frac{|\overrightarrow{p_0p} \cdot \vec{w}|}{\|\vec{w}\|}.$$

(c) $\overrightarrow{p_tq_s}$ és la projecció ortogonal de $\overrightarrow{p_0q_0}$ sobre el subespai $\langle \vec{u} \wedge \vec{v} \rangle$; així doncs,

$$d(R,R') = \|\overrightarrow{p_tq_s}\| = \frac{|\overrightarrow{p_0q_0} \cdot (\vec{u} \wedge \vec{v})|}{\|\vec{u} \wedge \vec{v}\|}. \quad \square$$

Acabem aquest breu formulari del producte vectorial amb aplicacions al càlcul de casos elementals però freqüents d'àrees i volums, estalviant al lector disquisicions teòriques sobre aquests dos conceptes, que li són prou familiars.

Proposició 8.14 *(a) L'àrea del triangle de vèrtexs $p, q, r \in P_3$ és*

$$A = \frac{1}{2}\|\overrightarrow{pq} \wedge \overrightarrow{pr}\|.$$

(b) L'àrea d'un paral·lelogram, on p és un vèrtex i q i r són els vèrtexs adjacents a p, és

$$A = \|\overrightarrow{pq} \wedge \overrightarrow{pr}\|.$$

DEMOSTRACIÓ. (a) Prenent com a base del triangle el costat pq, l'altura corresponent és $h = d(r, pq)$ i, aplicant la fórmula de la distància d'un punt a una recta,

$$A = \frac{1}{2}\|\overrightarrow{pq}\| h = \frac{1}{2}\|\overrightarrow{pq}\| \frac{\|\overrightarrow{pq} \wedge \overrightarrow{pr}\|}{\|\overrightarrow{pq}\|} = \frac{1}{2}\|\overrightarrow{pq} \wedge \overrightarrow{pr}\|.$$

(b) El raonament és anàleg, però aquí $A = \|\overrightarrow{pq}\| h$. $\square$

Proposició 8.15 *(a) El volum d'un tetràedre de vèrtexs $p, q, r, s \in P_3$ és*

$$V = \frac{1}{6} | \overrightarrow{ps} \cdot (\overrightarrow{pq} \wedge \overrightarrow{pr}) | .$$

(b) El volum d'un paral·lelepípede, on p és un vèrtex i q, r i s són els vèrtexs adjacents a p, és

$$V = | \overrightarrow{ps} \cdot (\overrightarrow{pq} \wedge \overrightarrow{pr}) | .$$

DEMOSTRACIÓ. (a) Prenem com a base el triangle de vèrtexs p, q, r. L'altura corresponent serà $h = d(s, pqr)$ i, aplicant fórmules anteriors,

$$V = \frac{1}{3} A(pqr)\, h = \frac{1}{3}\frac{1}{2} \| \overrightarrow{pq} \wedge \overrightarrow{pr} \| \frac{| \overrightarrow{ps} \cdot (\overrightarrow{pq} \wedge \overrightarrow{pr}) |}{\| \overrightarrow{pq} \wedge \overrightarrow{pr} \|} = \frac{1}{6} | \overrightarrow{ps} \cdot (\overrightarrow{pq} \wedge \overrightarrow{pr}) | .$$

(b) El raonament és anàleg, però aquí el volum és simplement l'àrea de la base per l'altura. $\square$

Observacions 8.16 (a) Cal fer notar que, en les quatre fórmules precedents, el vèrtex destacat p pot ser qualsevol de la figura mentre es respecti l'adjacència en les fórmules, i també que podem calcular el producte mixt fent servir determinants:

$$\overrightarrow{ps} \cdot (\overrightarrow{pq} \wedge \overrightarrow{pr}) = \det(\overrightarrow{pq},\ \overrightarrow{pq},\ \overrightarrow{ps}) .$$

(b) Les fórmules de les àrees també són aplicables a figures de pla P_2, si el considerem submergit dins de P_3 afegint simplement una tercera coordenada nul·la a cada punt.

Per exemple, per calcular l'àrea del triangle de vèrtexs $p = (p_1, p_2)$, $q = (q_1, q_2)$ i $r = (r_1, r_2)$, submergint P_2 dins de P_3 escrivim

$$\overrightarrow{pq} = (q_1 - p_1, q_2 - p_2, 0) , \qquad \text{i} \qquad \overrightarrow{pr} = (r_1 - p_1, r_2 - p_2, 0) .$$

L'àrea del triangle pqr és

$$\frac{1}{2} \| \overrightarrow{pq} \wedge \overrightarrow{pr} \| .$$

D'altra banda,

$$\overrightarrow{pq} \wedge \overrightarrow{pr} = \left(0,\ 0,\ \begin{vmatrix} q_1 - p_1 & q_2 - p_2 \\ r_1 - p_1 & r_2 - p_2 \end{vmatrix} \right)$$

i, tenint en compte que

$$\begin{vmatrix} q_1 - p_1 & q_2 - p_2 \\ r_1 - p_1 & r_2 - p_2 \end{vmatrix} = \begin{vmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{vmatrix} ,$$

resulta que l’àrea del triangle pqr és el valor absolut de

$$\frac{1}{2}\begin{vmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{vmatrix}.$$

9

Corbes i superfícies de segon grau

En el capítol anterior hem estudiat a fons les rectes i els plans, és a dir, les figures geomètriques definides per equacions de primer grau. Aquest capítol tracta de les còniques i les quàdriques, que són, respectivament, les corbes del pla i les superfícies de l'espai definides per equacions de segon grau.

Discutirem amb detall les còniques, i més breument les quàdriques per raons de concisió: el lector trobarà analogies formals suficients per justificar aquesta diferència de tractament.

Continuarem denotant per $E_2 = (P_2, V_2, +)$ el pla euclidià puntual i per $E_3 = (P_3, V_3, +)$ l'espai, i fent ús, només, de referències rectangulars positives.

9.1 El·lipse, hipèrbola i paràbola

Donats dos punts F i F' del pla P_2, anomenarem *el·lipse* el conjunt de punts $p \in P_2$ tals que

$$d(p, F) + d(p, F') = 2a ,$$

on $2a > d(F, F')$. F i F' són els *focus* de l'el·lipse i a el *semieix major*.

El punt mitjà del segment FF' és el *centre* de l'el·lipse, i la constant $2c = d(F, F')$ és la *distància focal*. La recta que uneix els focus és l'*eix principal* i la recta que passa pel centre i és perpendicular a l'eix principal s'anomena *eix secundari*.

Quan $F = F'$, l'el·lipse es redueix a una *circumferència*, i en aquest cas es poden prendre com a eixos dues rectes perpendiculars qualssevol que passin pel centre.

Escollint una referència $\mathcal{R} = \{ o\,;\ \vec{u},\ \vec{v}\,\}$, on o és el centre de l'el·lipse, $\vec{u}$ un vector unitari sobre l'eix principal i $\vec{v}$ un vector unitari sobre l'eix secundari, els focus són $F = (c, 0)$ i $F' = (-c, 0)$, i els punts de l'el·lipse són els que compleixen l'equació

$$\sqrt{(x-c)^2 + y^2} + \sqrt{(x+c)^2 + y^2} = 2a\,.$$

Aquesta equació és equivalent a

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2)\,,$$

i si posem $b^2 = a^2 - c^2$ i dividim tota l'equació per a^2b^2 es converteix en

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,,$$

que és l'*equació canònica* de l'el·lipse. En el cas d'una circumferència s'escriu $x^2 + y^2 = r^2$, on $r = a = b$ és el *radi*.

De la mateixa manera, anomenarem *hipèrbola* el conjunt de punts $p \in P_2$ tals que

$$|\, d(p, F) - d(p, F')\,| = 2a\,,$$

on ara $0 < 2a < d(F, F')$. F i F' són els *focus* de la hipèrbola, i a el *semieix real*. El centre, la distància focal i els eixos principal i secundari d'una hipèrbola es defineixen igual que per a l'el·lipse.

Escollint una referència com en el cas anterior, l'equació de la hipèrbola és

$$|\,\sqrt{(x-c)^2 + y^2} - \sqrt{(x+c)^2 + y^2}\,| = 2a\,,$$

que és equivalent a

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2)\,.$$

Prenent $b^2 = c^2 - a^2$ i dividint per a^2b^2 obtenim

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1\,,$$

que és l'*equació canònica* de la hipèrbola.

Observacions 9.1 (a) De les equacions canòniques de l'el·lipse i la hipèrbola es dedueix que els eixos principal i secundari són eixos de simetria, ja que si (x, y) compleix l'equació els seus simètrics respecte als eixos, $(-x, y)$ i $(x, -y)$, també la compleixen.

(b) Les interseccions d'aquestes còniques amb els eixos s'anomenen *vèrtexs*. L'el·lipse té quatre vèrtexs: dos sobre l'eix principal, els punts de coordenades $(-a, 0)$ i $(a, 0)$, i dos sobre l'eix secundari, $(0, -b)$ i $(0, b)$. La hipèrbola només en té dos sobre l'eix principal, els punts de coordenades $(-a, 0)$ i $(a, 0)$.

(c) Aïllant la y de les equacions canòniques de l'el·lipse i la hipèrbola obtenim

$$y = \pm\frac{b}{a}\sqrt{a^2 - x^2} \quad \text{i} \quad y = \pm\frac{b}{a}\sqrt{x^2 - a^2}$$

respectivament. Això ens diu que l'el·lipse és una figura fitada i la hipèrbola no. A més, les rectes d'equacions

$$y = \pm\frac{b}{a}x$$

són *asímptotes* de la hipèrbola. Si les dues asímptotes són perpendiculars, és a dir, si $a = b$, direm que la hipèrbola és *equilàtera*.

Donats un punt F i una recta R que no el conté, anomenarem *paràbola* el conjunt de punts p del pla que compleixen la condició

$$d(p, F) = d(p, R)\,.$$

F és el *focus* de la paràbola, R és la *directriu*, i el punt mitjà de la perpendicular de F a R és el *vèrtex*. La distància $m = d(F, R)$ és el *paràmetre* de la paràbola.

Escollim una referència $\mathcal{R} = \{\,o\,;\ \vec{u},\ \vec{v}\,\}$, on o és el vèrtex de la paràbola, $\vec{u}$ un vector unitari amb la direcció de R i $\vec{v}$ un vector unitari i perpendicular a $\vec{u}$ en el sentit de R a F. L'equació de R és $y = -\frac{m}{2}$, el focus és $F = (0, \frac{m}{2})$ i els punts de la paràbola són els que compleixen l'equació

$$\sqrt{x^2 + \left(y - \frac{m}{2}\right)^2} = \left|\, y + \frac{m}{2}\,\right|\,,$$

equivalent a

$$x^2 - 2my = 0 \quad \text{o bé} \quad y = \frac{x^2}{2m}\,,$$

que és l'*equació canònica* de la paràbola.

Observacions 9.2 (a) Aquesta equació ens diu que la paràbola té un únic eix de simetria, l'eix *oy*, i un únic vèrtex, que coincideix amb l'origen de coordenades de $\mathcal{R}$.

(b) A diferència de l'el·lipse i la hipèrbola, en les que el sentit dels vectors $\vec{u}$ i $\vec{v}$ (i, per tant, el sentit positiu dels eixos de coordenades) no influeix en l'equació canònica, si en l'elecció de la referència $\mathcal{R}$ per a la paràbola invertim el sentit del vector $\vec{v}$ (i, per tant, el sentit de l'eix oy), la recta R tindrà per equació $y = \frac{m}{2}$, el focus serà $F = (0, -\frac{m}{2})$ i l'equació de la paràbola serà $x^2 + 2my = 0$.

9.2 Classificació de les còniques

Una *cònica* del pla és el conjunt de punts que en una referència $\mathcal{R}$ satisfan una equació de segon grau

$$F(x, y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0\,.$$

El nostre objectiu és trobar una altra referència on l'equació de la cònica adopti la forma més senzilla possible. Aquesta serà l'*equació reduïda* de la cònica, que en el cas d'una el·lipse, una hipèrbola o una paràbola haurà de coincidir, essencialment, amb la seva equació canònica.

Aquest procés de reducció ens permetrà donar una classificació de les còniques i introduir el concepte d'*invariant* d'una cònica per canvis de coordenades.

Si posem

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \qquad X = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{i} \quad L = (d \quad e)\,,$$

l'equació $ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0$ es pot escriure

$$X^t AX + 2LX + f = 0\,,$$

que és l'*equació matricial* de la cònica; si posem

$$A_3 = \begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix} \quad \text{i} \quad X_3 = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix},$$

la mateixa equació té la forma

$$X_3^t\, A_3\, X_3 = 0\,,$$

equació matricial compacta de la cònica. L'equivalència de les tres formes es comprova directament efectuant els productes de les matrius que intervenen a les equacions.

Donada una nova referència $\mathcal{R}' = \{o\,;\ \vec{u},\ \vec{v}\,\}$, on $o = (x_0, y_0)$, $\vec{u} = (c_{11}, c_{21})$ i $\vec{v} = (c_{12}, c_{22})$, les equacions del canvi de coordenades són

$$X = B + CX'\,,$$

on

$$B = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \qquad C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \quad \text{i} \quad X' = \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

La forma matricial compacta d'aquest canvi de coordenades és

$$X_3 = C_3\, X_3', \quad \text{on} \quad C_3 = \begin{pmatrix} c_{11} & c_{12} & x_0 \\ c_{21} & c_{22} & y_0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{i} \quad X_3' = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}.$$

Com que les dues referències són rectangulars, C és una matriu ortogonal, és a dir, $C^tC = I$ i per tant $C^t = C^{-1}$. A més, $\det(C) = 1$, ja que també suposem sempre que les referències utilitzades són positives.

Proposició 9.3 *Aplicant a l'equació $X^tAX + 2LX + f = 0$ el canvi de coordenades $X = B + CX'$ obtenim*

$$(X')^t A'X' + 2L'X' + f' = 0\,,$$

on

$$\begin{cases} A' = C^tAC \\ L' = (B^tA + L)C \\ f' = B^tAB + 2LB + f = F(x_0, y_0)\,. \end{cases}$$

L'equació matricial compacta de la cònica en les noves coordenades és

$$(X_3')^t\, A_3'\, X_3' = 0\,, \quad \textit{on} \quad A_3' = C_3^t\, A_3\, C_3\,.$$

DEMOSTRACIÓ. En substituir $X = B + CX'$ a l'equació matricial de la cònica s'obté

$$(B + CX')^t A(B + CX') + 2L(B + CX') + f = 0$$

i, en efectuar els productes d'aquestes matrius,

$$B^tAB + B^tACX' + (X')^tC^tAB + (X')^tC^tACX' + 2LB + 2LCX' + f = 0\,.$$

D'altra banda, $(X')^tC^tAB = B^tACX'$, ja que aquests productes donen escalars; per tant, agrupant els termes obtenim

$$(X')^tC^tACX' + 2(B^tA + L)CX + B^tAB + 2LB + f = 0\,,$$

és a dir,

$$(X')^tA'X' + 2L'X' + f' = 0\,.$$

De la mateixa manera, la substitució de $X_3 = C_3 X_3'$ a l'equació matricial compacta de la cònica dóna

$$(X_3')^t C_3^t A_3 C_3 X_3' = 0\,,$$

és a dir,

$$(X_3')^t A_3' X_3' = 0\,. \quad \square$$

La noció de cònica no depèn del sistema de referència escollit ja que, segons la proposició anterior, el fet de quedar definida per una equació de segon grau no varia en fer un canvi de coordenades.

Les matrius A i A_3 d'una banda, i les matrius A' i A_3' de l'altra, corresponen a les formes matricial i matricial compacta de la mateixa cònica en dos sistemes de referència diferents. Estan relacionades per les igualtats

$$A' = C^t A C \qquad \text{i} \qquad A_3' = C_3^t A_3 C_3\,,$$

la qual cosa ens permetrà introduir el concepte d'*invariant* a partir del resultat següent.

Proposició 9.4 *En tot canvi de coordenades es compleix:*

(a) A *i* A' *tenen els mateixos valors propis.*

(b) $\det(A) = \det(A')$, $\quad tr(A) = tr(A')$.

(c) $\det(A_3) = \det(A_3')$, $\quad \mathrm{rang}(A_3) = \mathrm{rang}(A_3')$.

DEMOSTRACIÓ. (a) Com que C és una matriu ortogonal, $A' = C^t A C = C^{-1} A C$; per tant, A i A' tenen el mateix polinomi característic i, en conseqüència, els mateixos valors propis.

(b) Només cal observar que el determinant i la traça són els coeficients del polinomi característic d'una matriu d'ordre 2.

(c) És evident que $\det(C_3) = \det(C) = 1$, per tant $\det(A_3) = \det(A_3')$. Finalment, $\mathrm{rang}(A_3) = \mathrm{rang}(A_3')$ perquè la matriu C_3 és regular. $\square$

Definim ara com *invariants* de la cònica:

(a) α i β: valors propis de A.

(b) $D_2 = \det(A)$ i $T = \mathrm{tr}(A)$.

(c) $D_3 = \det(A_3)$ i $r_3 = \mathrm{rang}(A_3)$.

Observació 9.5 T, D_2 i D_3 són invariants de la funció $F(x, y)$ per canvis de coordenades, però no podem dir que siguin invariants de la cònica descrita per l'equació $F(x, y) = 0$. Només cal observar que la multiplicació d'aquesta equació per un escalar $t \neq 0$ no modifica la cònica, mentre que les quantitats anteriors es transformen en tT, $t^2 D_2$ i $t^3 D_3$ respectivament. En canvi, el signe de D_2, l'anul·lació de D_3, el rang r_3 i el signe de $D_3 T$ són *invariants estrictes* de la cònica, ja que són independents del sistema de referència i dels factors $t \neq 0$, i són els que farem servir per donar un criteri pràctic de classificació de les còniques en aquesta mateixa secció.

La classificació de les còniques es basa en fer l'elecció del canvi de coordenades $X = B + CX'$ de manera que l'equació inicial de la cònica es transformi en la més senzilla possible. La matriu C ens dóna les direccions dels nous eixos de coordenades, que han de coincidir amb els de la cònica, mentre que l'objectiu de la translació B és que el nou origen de coordenades coincideixi amb el centre o el vèrtex de la cònica.

Per tal que l'equació resultant sigui la més senzilla possible, la matriu A' ha de ser diagonal i, si és possible, hem de fer que $L' = 0$ o $f' = 0$. Si $\det(A) \neq 0$, el sistema d'equacions $AX + L^t = 0$ és compatible determinat i, per tant, si B n'és la solució tindrem que $L' = 0$. El lema següent descriu els casos en que $D_2 = \det(A) = 0$.

Lema 9.6 *(a) Si $D_2 = D_3 = 0$, el sistema $AX + L^t = 0$ és compatible indeterminat.*
(b) Si $D_2 = 0$ i $D_3 \neq 0$, el sistema $AX + L^t = 0$ és incompatible.

DEMOSTRACIÓ. Suposant que $D_2 = 0$, és suficient demostrar que el sistema és compatible indeterminat si, i només si, $D_3 = 0$. Com que $D_2 = 0$, la matriu A ha de ser de la forma

$$A = \begin{pmatrix} a & ta \\ ta & t^2 a \end{pmatrix} \quad \text{amb} \quad a \neq 0 \qquad \text{ó} \qquad A = \begin{pmatrix} 0 & 0 \\ 0 & c \end{pmatrix} \quad \text{amb} \quad c \neq 0\,.$$

En el primer cas tindrem que

$$D_3 = \begin{vmatrix} a & ta & d \\ ta & t^2 a & e \\ d & e & f \end{vmatrix} = d \begin{vmatrix} ta & d \\ t^2 a & e \end{vmatrix} - e \begin{vmatrix} a & d \\ ta & e \end{vmatrix} = -a(dt - e)^2\,;$$

aleshores, si $D_3 = 0$ ha de ser $e = td$, i el sistema $AX + L^t = 0$ queda

$$\begin{pmatrix} a & ta \\ ta & t^2 a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d \\ dt \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que, evidentment, és compatible indeterminat.

Recíprocament, si el sistema és compatible indeterminat s'ha de complir que

$$\begin{vmatrix} a & d \\ ta & e \end{vmatrix} = 0\,, \qquad \text{d'on} \qquad D_3 = 0\,.$$

La demostració en el segon cas és anàloga i més senzilla. □

Teorema 9.7 *Donada una cònica $F(x, y) = 0$, existeix una referència rectangular positiva $\mathcal{R}' = \{o\,; \vec{u},\ \vec{v}\}$ on la seva equació és d'un dels tres tipus següents:*

(a) $\alpha x'^2 + \beta y'^2 + f' = 0$,

(b) $\alpha x'^2 + f' = 0$,

(c) $\alpha x'^2 + 2e'y' = 0$,

amb α, β i e' no nuls.

DEMOSTRACIÓ. Hem de trobar un canvi de coordenades rectangular $X = B + CX'$ tal que l'equació de la cònica en les noves coordenades sigui d'un dels tres tipus de l'enunciat.

Sigui C una matriu ortogonal amb $\det(C) = 1$ tal que $A' = C^tAC$ sigui diagonal:

$$A' = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix}$$

amb $\alpha \neq 0$.

Per escollir la translació B distingirem tres casos:

(a) Si $D_2 \neq 0$, sigui $B = (x_0, y_0)^t$ l'única solució del sistema $AX + L^t = 0$. Aleshores, l'equació de la cònica en les noves coordenades és $\alpha x'^2 + \beta y'^2 + f' = 0$, on $f' = F(x_0, y_0)$.

(b) Si $D_2 = D_3 = 0$, sigui $B = (x_0, y_0)^t$ una solució qualsevol del sistema $AX + L^t = 0$. Aleshores, l'equació de la cònica en les noves coordenades és $\alpha x'^2 + f' = 0$, on $f' = F(x_0, y_0)$, ja que un dels valors propis de A ha de ser nul.

(c) Si $D_2 = 0$ i $D_3 \neq 0$, sigui $B = (x_0, y_0)^t$ l'única solució del sistema d'equacions

$$\left.\begin{aligned} F(x, y) &= 0 \\ L'_1 &= 0 \end{aligned}\right\}\,,$$

on L'_1 és el primer coeficient de L' (vegeu el lema següent). Aleshores, l'equació de la cònica en les noves coordenades és $\alpha x'^2 + 2e'y' = 0$, on

$$e' = (d, e) \cdot \vec{v}$$

i $\vec{v}$ és el vector propi unitari de A de valor propi 0 que ocupa la segona columna de la matriu de canvi C. □

Lema 9.8 *Si $D_2 = 0$ i $D_3 \neq 0$, el sistema*

$$\left.\begin{array}{r} F(x,y) = 0 \\ L_1' = 0 \end{array}\right\}$$

té una única solució.

DEMOSTRACIÓ. Farem la demostració només en el cas on

$$A = \begin{pmatrix} a & ta \\ ta & t^2a \end{pmatrix} \qquad \text{amb} \qquad a \neq 0\,,$$

deixant la de l'altre cas per al lector. Els valors propis de A són $\alpha = a(1+t^2)$ i $\beta = 0$, i la matriu de canvi és

$$C = \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} 1 & -t \\ t & 1 \end{pmatrix}.$$

Aleshores és immediat comprovar que $F(x,y) = a(x+ty)^2 + 2dx + 2ey + f$, i que l'equació $L_1' = 0$ es pot escriure $a(1+t^2)(x+ty) + d + et = 0$; per tant, el sistema

$$\left.\begin{array}{r} F(x,y) = 0 \\ L_1' = 0 \end{array}\right\}$$

es transforma en un sistema lineal de dues incògnites i determinant

$$\begin{vmatrix} 1 & t \\ d & e \end{vmatrix} = e - dt \neq 0\,,$$

ja que $D_3 = -a(e-dt)^2 \neq 0$. En conseqüència, el sistema és compatible determinat. □

Les matrius de la cònica en les noves coordenades són

$$A' = \begin{pmatrix} \alpha & \\ & \beta \end{pmatrix} \quad \text{ó} \quad A' = \begin{pmatrix} \alpha & \\ & 0 \end{pmatrix}$$

i

$$A_3' = \begin{pmatrix} \alpha & & \\ & \beta & \\ & & f' \end{pmatrix}, \qquad A_3' = \begin{pmatrix} \alpha & & \\ & 0 & \\ & & f' \end{pmatrix} \quad \text{ó} \quad A_3' = \begin{pmatrix} \alpha & & \\ & 0 & e' \\ & e' & 0 \end{pmatrix}$$

A més, els invariants de la cònica es poden calcular a partir de les matrius en les coordenades originals i a partir d'aquestes matrius. A continuació analitzem els diferents tipus de figura que corresponen a les equacions del teorema anterior, fent una primera distinció segons $D_3 \neq 0$ o $D_3 = 0$.

Còniques no degenerades ($D_3 \neq 0$*)*

Només ho poden ser les del tipus (a) amb $f' \neq 0$ i les del tipus (c). Les del primer tipus tenen equació reduïda $\alpha x'^2 + \beta y'^2 + f' = 0$, i distingim tres possibilitats:

1. Valors propis del mateix signe ($D_2 = \alpha\beta > 0$) i terme independent amb signe contrari ($\dfrac{D_3}{T} = \dfrac{\alpha\beta f'}{\alpha+\beta} < 0$). L'equació es transforma en
$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = 1$$
i la cònica és una *el·lipse real*. Si els valors propis coincideixen, és una *circumferència*, ja que queda $p = q$ (radi de la circumferència).

2. Valors propis i terme independent del mateix signe ($D_2 = \alpha\beta > 0$ i $\dfrac{D_3}{T} = \dfrac{\alpha\beta f'}{\alpha+\beta} > 0$). L'equació es transforma en
$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = -1$$
i la cònica és una *el·lipse imaginària*.

3. Valors propis amb signe diferent ($D_2 = \alpha\beta < 0$). Permutant els eixos si és necessari, l'equació es transforma en
$$\frac{x'^2}{p^2} - \frac{y'^2}{q^2} = 1$$
i la cònica és una *hipèrbola*. Si $T = 0$, els valors propis compleixen $\beta = -\alpha$ i la hipèrbola és *equilàtera*.

Les còniques del segon tipus tenen equació reduïda $\alpha x'^2 + 2e'y' = 0$ ($D_2 = 0$), que en aillar la y' es transforma en
$$y' = \frac{x'^2}{2m}\,.$$
Per tant, es tracta d'una *paràbola*.

Còniques degenerades ($D_3 = 0$*)*

Ho són les còniques del tipus (a) amb $f' = 0$ i les del tipus (b). Les del primer tipus tenen equació reduïda $\alpha x'^2 + \beta y'^2 = 0$, i distingim dos casos:

1. Valors propis amb signe diferent ($D_2 = \alpha\beta < 0$). L'equació es pot escriure
$$p^2x'^2 - q^2y'^2 = 0 \quad \text{o bé} \quad \begin{cases} px' + qy' = 0 \\ px' - qy' = 0 \end{cases}$$
i la cònica és un *parell de rectes reals secants*. Si $T = 0$, els valors propis compleixen $\beta = -\alpha$ i les dues rectes són perpendiculars.

2. Valors propis del mateix signe ($D_2 = \alpha\beta > 0$). L'equació es pot escriure

$$p^2 x'^2 + q^2 y'^2 = 0 \quad \text{o bé} \quad \begin{cases} px' + iqy' = 0 \\ px' - iqy' = 0 \end{cases}$$

i la cònica és un *parell de rectes imaginàries secants* amb vèrtex real (el punt $x' = y' = 0$).

Les còniques del segon tipus tenen equació reduïda $\alpha x''^2 + f' = 0$ ($D_2 = 0$), i ara distingim tres casos:

1. $f' = 0$ ($r_3 = 1$). L'equació reduïda és $x'^2 = 0$ i la cònica és una *recta doble*.
2. α i f' tenen signe diferent ($r_3 = 2$). L'equació reduïda es transforma en

$$x'^2 = p^2 \quad \text{o bé} \quad \begin{cases} x' = p \\ x' = -p \end{cases}$$

i la cònica és un *parell de rectes reals paral·leles* .

3. α i f' tenen el mateix signe ($r_3 = 2$). L'equació reduïda es transforma en

$$x'^2 = -p^2 \quad \text{o bé} \quad \begin{cases} x' = pi \\ x' = -pi \end{cases}$$

i la cònica és un *parell de rectes imaginàries paral·leles*.

Les observacions que hem fet sobre els invariants en aquesta classificació de les còniques a partir de l'equació reduïda ens permeten establir el criteri de classificació de la taula 9.1.

$$D_3 \neq 0 \begin{cases} D_2 > 0 & \begin{cases} D_3 T > 0 & \text{el·lipse imaginària} \\ D_3 T < 0 & \text{el·lipse real} \end{cases} \\ D_2 < 0 & \text{hipèrbola} \\ D_2 = 0 & \text{paràbola} \end{cases}$$

$$D_3 = 0 \begin{cases} D_2 > 0 & \text{parell de rectes imaginàries secants amb vèrtex real} \\ D_2 < 0 & \text{parell de rectes reals secants} \\ D_2 = 0 & \begin{cases} r_3 = 2 & \text{parell de rectes paral·leles (reals o imaginàries)} \\ r_3 = 1 & \text{recta doble} \end{cases} \end{cases}$$

Taula 9.1: Classificació de còniques

Observacions 9.9 (a) El cas $\alpha = \beta = 0$ fa que l'equació es redueixi a una de primer grau ($A' = A = 0$) i, com que es tracta d'una recta, no l'hem inclòs a l'estudi de les còniques, on hem suposat sempre $\alpha \neq 0$.

(b) La referència $\mathcal{R}' = \{\, 0\,;\ \vec{u},\ \vec{v}\,\}$ on la cònica adopta l'equació reduïda compleix les propietats següents: (1) $\vec{u}$ és un vector propi unitari de valor propi α i determina la direcció positiva de l'eix ox'; (2) $\vec{v}$ és un vector propi unitari de valor propi β, determina la direcció positiva de l'eix oy' i el podem escollir de manera que $\det(C) = 1$ per tal que la referència $\mathcal{R}'$ sigui positiva; (3) el nou origen de coordenades és el punt $x' = y' = 0$, per tant a partir d'aquí podem trobar les seves coordenades en la referència original i coincideix amb el centre de l'el·lipse o el de la hipèrbola o amb el vèrtex de la paràbola.

(c) Si permutem l'ordre en que escollim els valors propis i vectors propis corresponents, el sistema de referència i l'equació reduïda també varien, intercanviant-se els papers de les coordenades x' i y'.

Exemples 9.10 (a) La cònica d'equació

$$10x^2 - 12xy - 6y^2 - 12x - 12y - 129 = 0$$

té matrius

$$A = \begin{pmatrix} 10 & -6 \\ -6 & -6 \end{pmatrix} \quad \text{i} \quad A_3 = \begin{pmatrix} 10 & -6 & -6 \\ -6 & -6 & -6 \\ -6 & -6 & -129 \end{pmatrix}$$

i invariants $D_2 = -96$ i $D_3 = 11808$, per tant és una *hipèrbola*.

El polinomi característic de la matriu A és $\det(xI - A) = x^2 - 4x - 96$ i els valors propis 12 i -8. Com a vectors propis escollim $(3, -1)$ de valor propi 12 i $(1, 3)$ de valor propi -8. El centre de la hipèrbola és la solució del sistema d'equacions $AX + L^t = 0$, és a dir,

$$\left.\begin{aligned} 10x - 6y - 6 &= 0 \\ -6x - 6y - 6 &= 0 \end{aligned}\right\} ,$$

és a dir, el punt $(0, -1)$, i els eixos són les rectes que passen pel centre i tenen direccions $(3, -1)$ i $(1, 3)$ respectivament. L'equació reduïda i el canvi de coordenades corresponent són

$$12x'^2 - 8y'^2 - 123 = 0 \quad \text{i} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} .$$

Finalment, les asímptotes de la hipèrbola són les rectes que passen pel centre i tenen pendent

$$m = \lim_{x \to \infty} \frac{y}{x} .$$

Dividint l'equació de la hipèrbola entre x^2 i prenent el límit obtenim que $10 - 12m - 6m^2 = 0$, és a dir, els pendents de les asímptotes són

$$\frac{-3 + 2\sqrt{6}}{3} \quad \text{i} \quad \frac{-3 - 2\sqrt{6}}{3}\,.$$

(b) Les matrius de la cònica d'equació

$$x^2 - 4xy + 4y^2 - 2x - y = 0$$

són

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \quad \text{i} \quad A_3 = \begin{pmatrix} 1 & -2 & -1 \\ -2 & 4 & -1/2 \\ -1 & -1/2 & 0 \end{pmatrix}$$

i els seus invariants $D_2 = 0$ i $D_3 = -25/4$, per tant la cònica és una *paràbola*.

Els valors propis de la matriu A són 5 i 0, i els vectors propis corresponents $(1, -2)$ i $(2, 1)$. Per calcular el vèrtex de la paràbola hem de resoldre el sistema d'equacions

$$\left.\begin{aligned} x^2 - 4xy + 4y^2 - 2x - y = 0 \\ x - 2y = 0 \end{aligned}\right\}\,,$$

que té per solució $(0, 0)$. D'altra banda, tenim que

$$e' = (-1, -1/2) \cdot \frac{1}{\sqrt{5}}(2, 1) = -\frac{\sqrt{5}}{2}$$

i l'equació reduïda de la paràbola és

$$5x'^2 - \sqrt{5}\,y' = 0\,, \quad \text{o bé} \quad y' = \sqrt{5}\,x'^2\,.$$

El canvi de coordenades que transforma l'equació inicial en l'equació reduïda és

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}\,.$$

(c) La cònica d'equació

$$2x^2 - 2xy - 4y^2 - 2x + 10y - 4 = 0$$

té matrius

$$A = \begin{pmatrix} 2 & -1 \\ -1 & -4 \end{pmatrix} \quad \text{i} \quad A_3 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & -4 & 5 \\ -1 & 5 & -4 \end{pmatrix}$$

i invariants $D_2 = -9$ i $D_3 = 0$, per tant és un *parell de rectes reals secants*. Per calcular les seves equacions, tractem l'equació de la cònica com una equació de segon grau per a y,

$$-4y^2 + (10 - 2x)y + (2x^2 - 2x - 4) = 0\,,$$

que té per solució

$$\begin{aligned} y &= \frac{2x - 10 \pm \sqrt{(10-2x)^2 + 16(2x^2 - 2x - 4)}}{-8} \\ &= \frac{2x - 10 \pm \sqrt{36x^2 - 72x + 36}}{-8} \\ &= \frac{2x - 10 \pm (6x - 6)}{-8}\,. \end{aligned}$$

Aleshores, la cònica està formada per les rectes

$$y = -x + 2 \quad \text{i} \quad y = \frac{x+1}{2}\,,$$

és a dir, $x + y - 2 = 0$ i $x - 2y + 1 = 0$.

9.3 Llocs geomètrics

Un *lloc geomètric* en el pla és un conjunt de punts que satisfan una propietat determinada, expressada habitualment en termes geomètrics (distàncies entre punts, distàncies entre punts i rectes, intersecció de rectes, etc.). El problema de determinar un lloc geomètric consisteix a trobar l'equació d'aquest conjunt de punts en una referència donada o en una adaptada al problema.

Les definicions d'el·lipse, hipèrbola i paràbola de la primera secció d'aquest capítol i el càlcul de les seves equacions canòniques en unes referències adequades són exemples de llocs geomètrics on la propietat que els determina ve donada per distàncies. En ocasions semblants, l'aplicació directa de la condició que han de complir els punts del lloc geomètric dóna la seva equació.

Exemple 9.11 A partir d'una recta R i un punt F exterior a ella, calculem l'equació del lloc geomètric dels punts p que satisfan, per a un $t \geq 0$ fix,

$$d(p, F) = t\,d(p, R)\,.$$

Escollim una referència de manera que la recta R tingui equació $y = 0$ i el punt F coordenades $(0, 1)$. Aleshores l'equació és

$$\sqrt{x^2 + (y-1)^2} = t|\,y\,|\,.$$

Elevant al quadrat els dos termes obtenim

$$x^2 + (y-1)^2 = t^2 y^2 \,,$$

i, agrupant els termes,

$$x^2 + (1-t^2)y^2 - 2y + 1 = 0 \,,$$

que és una cònica per a qualsevol valor de t. Tenint en compte que $D_3 = -t^2$ i $D_2 = 1 - t^2$, resulta: si $t = 0$, la cònica és un parell de rectes imaginàries secants amb vèrtex real (el punt F); si $0 < t < 1$, és una el·lipse; si $t = 1$, és una paràbola; i si $t > 1$, és una hipèrbola.

En altres ocasions, les propietats que defineixen els punts d'un lloc geomètric depenen d'elements variables que es poden expressar en termes d'un o més paràmetres. Aleshores, les equacions que descriuen aquests punts també depenen dels paràmetres, i l'equació del lloc geomètric s'obté en eliminar-los d'aquest sistema. (Observem que per poder eliminar–los el sistema ha de tenir una equació més que paràmetres.)

Exemple 9.12 Donats el punt $F = (0,1)$ i la paràbola $x^2 - 4y = 0$, calculem l'equació del lloc geomètric dels punts mitjans dels segments que uneixen dos punts de la paràbola i passen per F.

Aquests segments s'obtenen en fer passar una recta variable pel punt F i calcular la seva intersecció amb la paràbola. Una recta que passa pel punt F té equació

$$y - 1 = mx \quad \text{o bé} \quad y = mx + 1$$

(les rectes variables depenen del paràmetre m). La intersecció d'aquesta recta i la paràbola ve donada per l'equació

$$x^2 - 4(mx+1) = 0 \,,$$

i les seves solucions són

$$x = \frac{4m \pm \sqrt{16m^2+16}}{2} = \frac{4m \pm 4\sqrt{m^2+1}}{2} = 2m \pm 2\sqrt{1+m^2} \,.$$

Per tant, les interseccions de la recta i la paràbola són els punts

$$P = (2m + 2\sqrt{1+m^2}, 1 + 2m^2 + 2m\sqrt{1+m^2}) \qquad \text{i}$$
$$Q = (2m - 2\sqrt{1+m^2}, 1 + 2m^2 - 2m\sqrt{1+m^2}) \,,$$

i el punt mitjà del segment PQ té per coordenades

$$\left.\begin{aligned} x &= 2m \\ y &= 1 + 2m^2 \end{aligned}\right\} \,.$$

Aquest és el sistema que descriu els punts del lloc geomètric en funció del paràmetre m. Per eliminar-lo, aïllem m de la primera equació,

$$m = \frac{x}{2},$$

i substituim a la segona: resulta

$$\begin{aligned} y &= 1 + 2\left(\frac{x}{2}\right)^2 \\ 2y &= 2 + x^2 \\ x^2 - 2y + 2 &= 0, \end{aligned}$$

que és l'equació d'una paràbola.

9.4 Quàdriques

Una *quàdrica* és una figura $\mathcal{C}$ de l'espai que en algun sistema de coordenades queda definida per una equació de la forma

$$F(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0.$$

Introduint les notacions

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \qquad \text{i} \qquad L = (b_1 \quad b_2 \quad b_3),$$

l'equació $F(x, y, z) = 0$ és equivalent a l'*equació matricial* de $\mathcal{C}$,

$$X^t AX + 2LX + c = 0,$$

i escrivint

$$X_4 = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \qquad \text{i} \qquad A_4 = \begin{pmatrix} A & L^t \\ L & c \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{12} & a_{22} & a_{23} & b_2 \\ a_{13} & a_{23} & a_{33} & b_3 \\ b_1 & b_2 & b_3 & c \end{pmatrix}$$

obtenim l'*equació matricial compacta*

$$X_4^t\, A_4\, X_4 = 0.$$

D'altra banda, un canvi de referència de les coordenades (x, y, z) a les coordenades (x', y', z') es descriu per $X = B + CX'$, on B és una transl·lació de l'origen al punt (x_0, y_0, z_0) i C una rotació dels eixos. En forma compacta resulta

$$X_4 = C_4\, X_4', \quad \text{on} \quad C_4 = \begin{pmatrix} C & B \\ 0 & 1 \end{pmatrix}.$$

La demostració dels dos resultats següents és la mateixa que en el cas de les còniques.

Proposició 9.13 *Aplicant a la quàdrica $\mathcal{C}$ el canvi de referència $X = B + CX'$, l'equació de la figura en la nova referència és $(X')^t A' X' + 2L'X' + c' = 0$, on*

$$\begin{cases} A' = C^t AC \\ L' = (B^t A + L)C \\ c' = B^t AB + 2LB + c = F(x_0, y_0, z_0) \end{cases}$$

En forma compacta, l'expressió passa a ser $(X_4')^t\, A_4'\, X_4' = 0$, on $A_4' = C_4^t\, A_4\, C_4$. □

Corol·lari 9.14 *Són invariants de la quàdrica $\mathcal{C}$ per canvis de coordenades:*

(a) Els valors propis α, β i γ de la matriu A.

(b) El polinomi característic de la matriu A, $\det(\lambda I - A) = \lambda^3 - t_1\lambda^2 + t_2\lambda - t_3$, on $t_1 = tr(A)$, t_2 és la suma dels tres menors diagonals d'ordre 2 i $t_3 = \det(A) = D_3$.

(c) El determinant $D_4 = \det(A_4)$.

(d) El rang $r_4 = \operatorname{rang}(A_4)$. □

Recordem que, com en el cas de les còniques, per parlar d'*invariants estrictes* cal tenir present l'efecte de multiplicar per $t \neq 0$ l'equació de la quàdrica. Els valors propis passen a ser $t\alpha$, $t\beta$ i $t\gamma$, els coeficients del polinomi característic de A queden multiplicats respectivament per t, t^2 i t^3 i el determinant total es converteix en $t^4 D_4$, mentre que el rang de A_4 no varia.

El procés de reducció de l'equació d'una quàdrica és semblant al de les còniques. Com que la matriu A és simètrica, existeix una matriu ortogonal C amb $\det(C) = 1$ i tal que $C^t AC = C^{-1}AC$ és diagonal; en llenguatge geomètric, una rotació d'eixos elimina els termes no rectangulars, és a dir, fa que

$$a_{12}' = a_{13}' = a_{23}' = 0\,,$$

mentre que $a_{11}' = \alpha$, $a_{22}' = \beta$ i $a_{33}' = \gamma$, valors propis de A. D'altra banda, escollint la translació B de manera adequada (o aplicant una segona rotació d'eixos en un dels casos) podem anul·lar els termes de la part lineal quan el sistema $AX + L^t = 0$ és compatible, o bé dos coeficients de la part lineal i el terme independent quan aquest sistema és incompatible. Arribem així al resultat següent.

Teorema 9.15 *Amb un canvi de coordenades adequat, l'equació de qualsevol quàdrica pot reduir-se a un dels cinc tipus següents:*

(a) $\alpha x'^2 + \beta y'^2 + \gamma z'^2 + c' = 0.$

(b) $\alpha x'^2 + \beta y'^2 + c' = 0.$

(c) $\alpha x'^2 + c' = 0.$

(d) $\alpha x'^2 + \beta y'^2 + 2b_3' z' = 0.$

(e) $\alpha x'^2 + 2b_3' y' = 0.$

Els coeficients α, β, γ *i* b_3' *són no nuls.* □

A continuació analitzem els diversos tipus de superfícies que representen aquestes equacions, i el criteri bàsic per distingir–les és el signe dels coeficients. Prèviament, però, seguim el criteri del rang per a una primera classificació.

Quàdriques no degenerades ($r_4 = 4$*,* és a dir, $D_4 \neq 0$*)*

Ho són les quàdriques del tipus (a) amb $c' \neq 0$ i les del tipus (d). Les del primer tipus tenen equació reduïda $\alpha x'^2 + \beta y'^2 + \gamma z'^2 + c' = 0$, i hem de distingir quatre possibilitats:

1. Els tres valors propis tenen el mateix signe i el terme independent signe contrari ($D_4 = \alpha\beta\gamma c' < 0$). Amb petites modificacions aritmètiques, l'equació adopta la forma
$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} + \frac{z'^2}{r^2} = 1$$
i la quàdrica és un *el·lipsoide real* (de *revolució* si coincideixen dos dels valors propis i una *esfera* si coincideixen tots tres).

2. Valors propis i terme independent del mateix signe ($D_4 > 0$). L'equació es podrà escriure
$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} + \frac{z'^2}{r^2} = -1$$
i la quàdrica és un *el·lipsoide imaginari* (sense punts reals).

3. Valors propis de signe diferent i terme independent del mateix signe que el valor propi aïllat ($D_4 > 0$). L'equació es pot escriure
$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} - \frac{z'^2}{r^2} = 1\,,$$

i la quàdrica és un *hiperboloide d'una fulla*.

4. Valors propis de signe diferent i terme independent amb el mateix signe que dos d'ells ($D_4 < 0$). L'equació quedarà de la forma

$$\frac{x'^2}{p^2} - \frac{y'^2}{q^2} - \frac{z'^2}{r^2} = 1\,,$$

que és l'*hiperboloide de dues fulles*.

Les quàdriques del tipus (d) tenen equació reduïda $\alpha x'^2 + \beta y'^2 + 2b'_3 z' = 0$, i distingim dues possibilitats:

1. Els valors propis no nuls són del mateix signe ($t_2 = \alpha\beta > 0$ i $D_4 = -\alpha\beta {b'_3}^2 < 0$). Resulta

$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = z'\,,$$

anomenada *paraboloide el·líptic* (*de revolució* o *circular* si $p = q$).

2. Els valors propis no nuls tenen signe diferent ($t_2 < 0$ i $D_4 > 0$). Prenent el primer positiu, el segon negatiu i el coeficient b'_3 negatiu, apareix

$$\frac{x'^2}{p^2} - \frac{y'^2}{q^2} = z'\,,$$

que és el *paraboloide hiperbòlic*.

Quàdriques poc degenerades ($r_4 = 3$ i, en particular, $D_4 = 0$)

Ho són les quàdriques del tipus (a) amb $c' = 0$, les del tipus (b) amb $c' \neq 0$ i les del tipus (e). Les del tipus (a) tenen equació reduïda $\alpha x'^2 + \beta y'^2 + \gamma z'^2 = 0$ ($D_3 \neq 0$), i distingim dues possibilitats:

1. Valors propis de signe diferent. L'equació s'escriu

$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = z'^2$$

i la quàdrica és un *con real* (*circular* o *de revolució* si $p = q$).

2. Valors propis del mateix signe. Això dóna

$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = -z'^2\,,$$

que és un *con imaginari* (amb *vèrtex real* $x' = y' = z' = 0$).

Les del tipus (b) tenen equació reduïda $\alpha x'^2 + \beta y'^2 + c' = 0$ ($D_3 = 0$), i hem de distingir tres possibilitats:

1. Valors propis no nuls del mateix signe ($t_2 > 0$) i terme independent de signe contrari ($\dfrac{\alpha\beta c'}{\alpha+\beta} < 0$). L'equació reduïda es transforma en

$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = 1$$

i la quàdrica és un *cilindre el·líptic real* (*circular* si $p = q$).

2. Valors propis no nuls i terme independent del mateix signe ($t_2 > 0$ i $\dfrac{\alpha\beta c'}{\alpha+\beta} > 0$). L'equació queda

$$\frac{x'^2}{p^2} + \frac{y'^2}{q^2} = -1$$

i això és un *cilindre el·líptic imaginari* (sense punts reals).

3. Valors propis no nuls de signe diferent ($t_2 < 0$). Permutant, si és necessari, les coordenades x' i y' tenim que

$$\frac{x'^2}{p^2} - \frac{y'^2}{q^2} = 1\,,$$

anomenat *cilindre hiperbòlic*.

Les quàdriques del tipus (e) tenen equació reduïda $\alpha x'^2 + 2b_3' y' = 0$ ($D_3 = t_2 = 0$), i l'equació es pot escriure com

$$x'^2 = 2py'\,,$$

que correspon a un *cilindre parabòlic*.

Quàdriques molt degenerades ($r_4 \le 2$ i, en particular, $D_4 = R_4 = D_3 = 0$*)*

Ho són les quàdriques del tipus (b) amb $c' = 0$ i les del tipus (c). Les del tipus (b) tenen equació reduïda $\alpha x'^2 + \beta y'^2 = 0$, i hem de distingir dues possibilitats:

1. Valors propis de signe diferent ($t_2 < 0$). Podem escriure

$$p^2 x'^2 - q^2 y'^2 = 0\,, \quad \text{o bé} \quad \begin{cases} px' + qy' = 0 \\ px' - qy' = 0 \end{cases}$$

que defineix un *parell de plans reals secants* (perpendiculars si $p = q$).

2. Valors propis del mateix signe ($t_2 > 0$). Tindrem

$$p^2 x'^2 + q^2 y'^2 = 0\,, \quad \text{o bé} \quad \begin{cases} px' + iqy' = 0 \\ px' - iqy' = 0 \end{cases}$$

que és un *parell de plans imaginaris secants* (amb intersecció real: la recta d'equacions $x' = y' = 0$).

Finalment, les quàdriques del tipus (c) tenen equació reduïda $\alpha x'^2 + c' = 0$ ($t_2 = 0$), i distingim tres possibilitats:

1. Si α i c' tenen signe diferent, tenim

$$x'^2 = p^2 , \quad \text{o bé} \quad \begin{cases} x' = p \\ x' = -p \end{cases}$$

   que és un *parell de plans reals paral·lels* (diferents).

2. Si α i c' tenen el mateix signe,

$$x'^2 = -p^2 , \quad \text{o bé} \quad \begin{cases} x' = ip \\ x' = -ip \end{cases}$$

   i això és un *parell de plans imaginaris paral·lels* (sense punts reals).

3. Si $c' = 0$ ($r_4 = 1$), dividint l'equació per α obtenim

$$x'^2 = 0$$

   que és un *pla doble*.

Aquesta classificació de les quàdriques es pot expressar en termes dels invariants d'acord amb l'esquema de la taula 9.2.

Observacions 9.16 (a) Per classificar una quàdrica no és indispensable calcular efectivament els valors propis de la matriu A: només cal tenir en compte els signes, i per a això és força útil el teorema de Descartes (vegeu l'apèndix B).

(b) Recomanem al lector que estudiï les seccions de les quàdriques per plans perpendiculars als eixos: és un exercici instructiu per familiaritzar–se amb els noms de les figures i per notar determinades característiques: la connexió d'el·lipsoides, paraboloides i hiperboloides d'una fulla, l'afitament de l'el·lipsoide, el paper del vèrtex del con i del paraboloide o els eixos dels cilindres.

(c) En comparació amb les còniques, aquí també trobem el cas no degenerat (el·lipsoides, hiperboloides i paraboloides) i un cas degenerat (parell de plans) anàleg al dels parells de rectes del pla; hi ha, però, un grup intermig, format pels cons i els cilindres, que és una mena d'híbrid sense correspondència en el cas de les còniques.

$$
D_4 > 0 \begin{cases} D_3 \neq 0 & \begin{cases} \textit{Valors propis amb el mateix signe} & \text{el·lipsoide imaginari} \\ \textit{Valors propis amb signe diferent} & \text{hiperboloide d'una fulla} \end{cases} \\ D_3 = 0 & \text{paraboloide hiperbòlic} \end{cases}
$$

$$
D_4 < 0 \begin{cases} D_3 \neq 0 & \begin{cases} \textit{Valors propis amb el mateix signe} & \text{el·lipsoide real} \\ \textit{Valors propis amb signe diferent} & \text{hiperboloide de dues fulles} \end{cases} \\ D_3 = 0 & \text{paraboloide el·líptic} \end{cases}
$$

$$
D_4 = 0 \begin{cases} D_3 \neq 0 & \begin{cases} \textit{Valors propis amb el mateix signe} & \text{con imaginari} \\ \textit{Valors propis amb signe diferent} & \text{con real} \end{cases} \\ D_3 = 0 & \begin{cases} r_4 = 3 & \begin{cases} t_2 > 0 & \text{cilindre el·líptic (real o imaginari)} \\ t_2 < 0 & \text{cilindre hiperbòlic} \\ t_2 = 0 & \text{cilindre parabòlic} \end{cases} \\ r_4 = 2 & \begin{cases} t_2 > 0 & \text{parell de plans imaginaris secants} \\ t_2 < 0 & \text{parell de plans reals secants} \\ t_2 = 0 & \text{parell de plans paral·lels (reals o imaginaris)} \end{cases} \\ r_4 = 1 & \text{pla doble} \end{cases} \end{cases}
$$

Taula 9.2: Classificació de quàdriques

Exemples 9.17 (a) Considerem la quàdrica d'equació

$$2x^2 + 2xy - 2xz - 2yz + 6x + 10y + 4z - 20 = 0\,.$$

Calculant els invariants a partir de la matriu

$$A_4 = \begin{pmatrix} 2 & 1 & -1 & 3 \\ 1 & 0 & -1 & 5 \\ -1 & -1 & 0 & 2 \\ 3 & 5 & 2 & -20 \end{pmatrix}$$

resulta $D_4 = 0$, $r_4 = 2$ i $t_2 = -3$, per tant la quàdrica és un *parell de plans reals secants*.

(b) Estudiem ara la quàdrica

$$x^2 + y^2 - 2z^2 + 2xy + 4xz + 2z - 4 = 0\,.$$

La matriu és

$$A_4 = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & -2 & 1 \\ 0 & 0 & 1 & -4 \end{pmatrix}.$$

Com que $D_4 = 16$, és una quàdrica no degenerada. El polinomi característic de la matriu A,

$$\det(\lambda I - A) = \lambda^3 - 8\lambda + 4\,,$$

ens indica que hi ha dos valors propis positius i un de negatiu, per tant aquesta quàdrica és un *hiperboloide d'una fulla*.

(c) Finalment, considerem la quàdrica

$$x^2 + y^2 - z^2 + 4x - 4z = 0\,.$$

Novament, partim de la matriu

$$A_4 = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -2 \\ 2 & 0 & -2 & 0 \end{pmatrix}$$

i obtenim $D_4 = 0$ i $D_3 = -1$. Els valors propis de A són 1 (doble) i -1, per tant es tracta d'un *con real circular*.

A

Permutacions

Una *permutació* del conjunt $N = \{1, 2, \ldots, n\}$ és una aplicació bijectiva $\sigma : N \longrightarrow N$, que escriurem amb la imatge de cada element a sota d'ell:

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma_1 & \sigma_2 & \ldots & \sigma_n \end{pmatrix}; \qquad \text{per exemple,} \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 1 & 6 & 3 & 2 \end{pmatrix}.$$

Representarem per $\mathcal{P}_n$ el conjunt de les $n!$ permutacions possibles de N, que són simplement reordenacions dels n elements. La *identitat*, definida per

$$\text{id} = \begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix},$$

és l'única permutació que conserva l'ordre natural. Donada una permutació σ, es diu que un parell $i < j$ és una *inversió* de σ si $\sigma_i > \sigma_j$. El *signe* de σ és $\epsilon(\sigma) = (-1)^{\text{inv}(\sigma)}$, on $\text{inv}(\sigma)$ és el nombre d'inversions de σ. Per exemple, $\epsilon(\text{id}) = 1$, i si σ és l'exemple concret del principi aleshores $\epsilon(\sigma) = -1$, ja que σ presenta 9 inversions.

Una *transposició* és una permutació que deixa fixos tots els elements de N excepte dos, i direm que és *bàsica* si aquests dos elements són consecutius; per exemple, si

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 3 & 4 & 2 & 6 \end{pmatrix} \qquad \text{i} \qquad \rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 4 & 3 & 5 & 6 \end{pmatrix},$$

τ és una transposició i ρ és una transposició bàsica. En qualsevol cas, escriurem simplement transposicions com aquestes en la forma $\tau = (2, 5)$ i $\rho = (3, 4)$. El signe de tota transposició és -1.

El *producte* de dues permutacions σ i τ és l'aplicació composta, escrita $\sigma\tau$ (recordem que primer opera τ i després σ). El producte és una operació associativa, i la identitat n'és l'element neutre. La *inversa* d'una permutació σ és l'aplicació inversa, escrita σ^{-1}, i és l'única permutació tal que $\sigma\sigma^{-1} = \mathrm{id} = \sigma^{-1}\sigma$. Amb els exemples anteriors,

$$\sigma\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 3 & 2 \end{pmatrix} \quad \text{i} \quad \sigma^{-1} = \begin{pmatrix} 4 & 5 & 1 & 6 & 3 & 2 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 5 & 1 & 2 & 4 \end{pmatrix}.$$

Proposició A.1 $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$ *per a qualsevol* $\sigma \in \mathcal{P}_n$.

DEMOSTRACIÓ. Cada inversió de σ equival a una inversió de σ^{-1}, ja que

$$i < j \quad \text{amb} \quad k = \sigma_i > \ell = \sigma_j \qquad \Longleftrightarrow \qquad \ell < k \quad \text{amb} \quad j = {\sigma^{-1}}_\ell > i = {\sigma^{-1}}_k . \quad \square$$

Proposició A.2 *Sigui* τ *una transposició. Aleshores:*

(a) $\epsilon(\sigma\tau) = -\epsilon(\sigma) \quad \forall\sigma \in \mathcal{P}_n$.

(b) L'aplicació $\mathcal{P}_n \longrightarrow \mathcal{P}_n$ *definida per* $\sigma \mapsto \sigma\tau$ *transforma les permutacions de signe positiu en permutacions de signe negatiu i viceversa; per tant, hi ha la meitat de permutacions de cada signe.*

DEMOSTRACIÓ. (a) Si τ és bàsica, $\sigma\tau$ presenta una inversió menys o una inversió més que σ, de manera que $\epsilon(\sigma\tau) = -\epsilon(\sigma)$. Si τ no és bàsica, τ és el producte d'un nombre senar de transposicions bàsiques, ja que per a qualsevol parell $i < j$ de N podem escriure

$$(i, j) = (i, i+1)(i+1, i+2)\cdots(j-2, j-1)(j-1, j)(j-2, j-1)\cdots(i+1, i+2)(i, i+1).$$

Aplicant reiteradament la propietat establerta abans en el cas de les transposicions bàsiques, arribem també aquí a la relació $\epsilon(\sigma\tau) = -\epsilon(\sigma)$.

(b) Segons l'apartat anterior, el signe de $\sigma\tau$ és el contrari del de σ; com que l'aplicació donada per $\sigma \mapsto \sigma\tau$ és injectiva, hi ha tantes permutacions de signe positiu com de signe negatiu. $\square$

Proposició A.3 *Tota permutació és producte de transposicions. Si* $\sigma = \tau_1\tau_2\cdots\tau_p$ *és una descomposició d'aquest tipus,*

$$\epsilon(\sigma) = (-1)^p .$$

DEMOSTRACIÓ. Aplicarem recurrència sobre el nombre d'elements que no queden fixos. Si σ mou només i i j, aleshores $\sigma = (i, j)$. Suposem ara que σ mou $k > 2$ elements de N, i escollim i tal que $\sigma_i = j \neq i$. La permutació $(i, j)\sigma$ deixa invariants els mateixos elements que σ i, a més, l'element i: per recurrència, sabem que $(i, j)\sigma = \tau_2 \cdots \tau_p$ (producte de transposicions), per tant $\sigma = \tau_1\tau_2 \cdots \tau_p$ amb $\tau_1 = (i, j)$. L'afirmació final sobre el signe de la permutació és una conseqüència immediata de l'apartat (a) de la proposició precedent. $\square$

B

Polinomis

Un *monomi* és una expressió de la forma ax^m, on el coeficient a és un nombre real o complex, x és una indeterminada i m és un nombre natural o zero. Un *polinomi* és una suma finita de monomis, que escrivim com

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \, ,$$

amb $a_n \neq 0$ si $n > 0$. Es diu que n és el *grau* de $p(x)$. Els polinomis de grau 0, també anomenats *constants*, s'identifiquen amb els nombres. Els polinomis es poden sumar, restar i multiplicar seguint les lleis ordinàries de l'àlgebra, i només tenen invers els polinomis constants no nuls.

Divisió entera. Donats dos polinomis $p(x)$ i $d(x) \neq 0$, existeixen dos únics polinomis $q(x)$ (*quocient*) i $r(x)$ (*residu*) tals que:

(1) $p(x) = d(x)q(x) + r(x)$.

(2) $r(x)$ és 0 o de grau més petit que el grau de $d(x)$.

La demostració és senzilla, per inducció sobre el grau de $p(x)$. Sens dubte, el lector també coneix la *regla de Ruffini*, que esquematitza la divisió d'un polinomi $p(x)$ per $x - \alpha$.

Quan $r(x) = 0$ la divisió és *exacta*, i es diu que $d(x)$ és un *divisor* de $p(x)$ o bé que $p(x)$ és un *múltiple* de $d(x)$.

Si α és un nombre real o complex, el *valor numèric* d'un polinomi $p(x)$ quan $x = \alpha$ és

$$p(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n\,,$$

i α és una *arrel* de $p(x)$ si $p(\alpha) = 0$.

Proposició B.1 *(a) (Teorema del residu.) El residu de la divisió d'un polinomi $p(x)$ per $x - \alpha$ és el valor numèric $p(\alpha)$.*

(b) (Teorema del factor.) Un nombre α és arrel d'un polinomi $p(x)$ si, i només si, $x - \alpha$ és un divisor de $p(x)$.

DEMOSTRACIÓ. (a) La divisió de $p(x)$ per $x - \alpha$ dóna $p(x) = (x - \alpha)q(x) + r$, i substituint arreu $x = \alpha$ queda $p(\alpha) = r$.

(b) Es una conseqüència immediata de (a). $\square$

Exemples B.2 (a) Els valors numèrics d'un polinomi es poden calcular amb la regla de Ruffini. Per exemple, si volem trobar $p(\sqrt{2})$ quan $p(x) = x^3 + 2x^2 - 3x + 1$, fem

$$\begin{array}{c|cccc}
 & 1 & 2 & -3 & 1 \\
\sqrt{2} & & \sqrt{2} & 2+2\sqrt{2} & 4-\sqrt{2} \\
\hline
 & 1 & 2+\sqrt{2} & -1+2\sqrt{2} & \boxed{5-\sqrt{2}}
\end{array}$$

i tenim que $p(\sqrt{2}) = 5 - \sqrt{2}$.

(b) Segons el teorema del factor, si α és una arrel de $p(x)$ tenim

$$p(x) = (x - \alpha)q(x)\,.$$

Si β és una arrel de $q(x)$ resulta $q(x) = (x - \beta)r(x)$, i per tant

$$p(x) = (x - \alpha)(x - \beta)r(x)\,.$$

Si continuem aquest procés de descomposició, i tenim en compte el comportament del grau davant de la multiplicació de polinomis, observem que el darrer quocient serà de grau zero, i per tant *un polinomi de grau $n > 0$ pot tenir, com a màxim, n arrels diferents.*

D'altra banda, una arrel pot aparèixer més d'una vegada durant aquesta descomposició d'un polinomi $p(x)$ com a producte de factors de primer grau, per tant és convenient introduir la noció següent: es diu que α és una arrel de $p(x)$ amb *multiplicitat m* si

$$p(x) = (x-\alpha)^m q(x) \quad \text{amb} \quad q(\alpha) \neq 0\,.$$

Així, direm que α és *simple* si $m = 1$, i *múltiple* (*doble, triple, ...*) si $m > 1$.

Per completar l'estudi de les arrels i la descomposició dels polinomis en factors, serà útil distingir entre els polinomis de coeficients reals, el conjunt dels quals representarem per $\mathbb{R}[x]$, i els polinomis de coeficients complexos, que formaran el conjunt denotat per $\mathbb{C}[x]$. Com que es pot considerar $\mathbb{R} \subset \mathbb{C}$, també tenim que $\mathbb{R}[x] \subset \mathbb{C}[x]$. El motiu de la distinció és el següent: mentre que molts polinomis de coeficients reals no tenen arrels reals ($x^2 + 1$ n'és l'exemple més senzill), o no tantes com indica el seu grau encara que comptem les multiplicitats, la situació és totalment diferent en el cas de coeficients complexos. El resultat següent permet afirmar que els complexos representen la culminació de les ampliacions de la idea de nombre.

Teorema B.3 *(Teorema fonamental de l'àlgebra.) Tot polinomi no constant de coeficients complexos té alguna arrel complexa.*

DEMOSTRACIÓ. Gauss va donar cinc demostracions diferents d'aquesta propietat. Lamentablement, cap d'elles és prou elemental per incloure–la aquí. □

Corol·lari B.4 *Tot polinomi no constant $p(x)$ de coeficients complexos es descompon totalment dins de $\mathbb{C}[x]$ en producte de factors de primer grau:*

$$p(x) = a_n(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)$$

o, agrupant les arrels segons les seves multiplicitats,

$$p(x) = a_n(x-\alpha_1)^{m_1}(x-\alpha_2)^{m_2}\cdots(x-\alpha_r)^{m_r}\,.$$

En particular, tot polinomi no constant té, comptant les multiplicitats, tantes arrels com indica el seu grau:

$$m_1 + m_2 + \cdots + m_r = n\,.$$

DEMOSTRACIÓ. Només cal aplicar reiteradament el teorema fonamental de l'àlgebra i el teorema del factor. □

El teorema de Gauss també dóna informació sobre els polinomis de coeficients reals, i ens ocuparem immediatament d'aquesta qüestió.

Lema B.5 *Si tenim un polinomi no constant $p(x) \in \mathbb{R}[x]$, i $\alpha \in \mathbb{C}$ és una arrel de $p(x)$ amb multiplicitat m, el complex conjugat $\overline{\alpha}$ també és una arrel de $p(x)$, i amb la mateixa multiplicitat m.*

DEMOSTRACIÓ. La conjugació, és a dir, l'aplicació

$$\alpha = u + iv \quad \mapsto \quad \overline{\alpha} = u - iv\,,$$

conserva les quatre operacions aritmètiques. Denotarem per $\overline{q}(x)$ el polinomi que resulta de conjugar els coeficientes de $q(x)$. Conjugant les relacions

$$p(x) = (x-\alpha)^m q(x) \qquad \text{i} \qquad q(\alpha) \neq 0$$

obtenim les corresponents a $\overline{\alpha}$,

$$p(x) = (x-\overline{\alpha})^m \overline{q}(x) \qquad \text{i} \qquad \overline{q}(\overline{\alpha}) \neq 0\,,$$

i per tant $\overline{\alpha}$ és una arrel de $p(x)$ amb multiplicitat m. □

Corol·lari B.6 *Tot polinomi no constant $p(x)$ de coeficients reals es descompon totalment dins de $\mathbb{R}[x]$ en producte de factors de primer grau o de segon grau amb discriminant negatiu:*

$$p(x) = a_n(x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}(x^2+b_1x+c_1)^{n_1}\cdots(x^2+b_sx+c_s)^{n_s}\,,$$

amb $\alpha_1,\ldots,\alpha_r, b_1, c_1, \ldots, b_s, c_s \in \mathbb{R}$ i ${b_j}^2 - 4c_j < 0$ per a $j = 1,\ldots,s$.

DEMOSTRACIÓ. Si $\beta = u + iv$ i $\overline{\beta} = u - iv$ són arrels complexes conjugades de $p(x)$, podem agrupar els factors corresponents així:

$$\begin{aligned}(x-\beta)(x-\overline{\beta}) &= (x-u-iv)(x-u+iv) = (x-u)^2 + v^2\\ &= x^2 - 2ux + (u^2+v^2) = x^2 + bx + c\,,\end{aligned}$$

on $b = -2u$, $c = u^2 + v^2$ i $b^2 - 4c < 0$. Partim ara de la descomposició de $p(x)$ a $\mathbb{C}[x]$, que segons el lema anterior serà de la forma

$$p(x) = a_n(x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}(x-\beta_1)^{n_1}(x-\overline{\beta}_1)^{n_1}\cdots(x-\beta_s)^{n_s}(x-\overline{\beta}_s)^{n_s}\,,$$

amb $\alpha_1,\ldots,\alpha_r \in \mathbb{R}$ i $\beta_1,\ldots,\beta_s \notin \mathbb{R}$. Agrupant de la manera indicada abans cada arrel complexa amb la seva conjugada resulta la descomposició de $p(x)$ dins de $\mathbb{R}[x]$. □

Un aspecte bàsic dels polinomis és la seva interpretació com a funcions. Cada polinomi $p(x)$ defineix una funció, denotada per p, que assigna a cada α el valor numèric $p(\alpha)$. El domini de la funció p pot ser $\mathbb{R}$ o $\mathbb{C}$ segons convingui; el recorregut, sempre dins de $\mathbb{C}$, es trobarà dins de $\mathbb{R}$ o no depenent del domini escollit i del tipus de coeficients del polinomi (tots reals o algun complex). Es important notar que si p és la funció nul·la aleshores $p(x)$ és el polinomi 0; per tant, polinomis diferents defineixen funcions diferents. Els polinomis de grau 0 s'anomenen *constants* perquè ho són les funcions associades.

Si $p(x) \in \mathbb{R}[x]$, habitualment es considera $p : \mathbb{R} \longrightarrow \mathbb{R}$. Aquesta *funció polinòmica* és contínua i, de fet, infinitament derivable.

Exemples B.7 (a) Tot polinomi $p(x) \in \mathbb{R}[x]$ de grau senar té alguna arrel real. Això és conseqüència de l'aparellament d'arrels conjugades, però també es pot demostrar considerant la funció polinòmica p. Suposant $a_n > 0$ observem que

$$\lim_{t \to -\infty} p(t) = -\infty \qquad \text{i} \qquad \lim_{t \to +\infty} p(t) = +\infty\,,$$

perquè el grau de $p(x)$ és senar. Per tant, existeixen nombres t_0 i t_1 tals que $p(t_0) < 0$ i $p(t_1) > 0$. Aplicant el teorema de Bolzano, es dedueix que $p(\alpha) = 0$ per a algun $\alpha \in (t_0, t_1)$.

(b) α és una arrel del polinomi $p(x)$ amb multiplicitat m si, i només si,

$$p(\alpha) = Dp(\alpha) = D^2 p(\alpha) = \cdots = D^{m-1} p(\alpha) = 0 \quad \text{però} \quad D^m p(\alpha) \neq 0\,,$$

on $D^i p(\alpha)$ representa la i–èssima derivada de la funció p en el punt α. La demostració d'aquesta propietat és senzilla, observant que si α és una arrel de $p(x)$ amb multiplicitat m aleshores també ho és de $Dp(x)$ amb multiplicitat $m - 1$.

Càlcul elemental d'arrels. Recordarem alguns criteris senzills per a l'obtenció d'arrels de polinomis. En primer lloc, si $p(x)$ és un polinomi de coeficients enters, les arrels enteres de $p(x)$ són divisors del terme independent a_0. Això suposa un nombre finit de proves amb la regla de Ruffini. Les arrels racionals de $p(x)$ són tals que, escrites en forma irreductible r/s, el numerador r ha de ser un divisor del terme independent a_0 i el denominador s ha de ser–ho del coeficient principal a_n; per tant, també en aquest cas el nombre de proves és finit.

Si $p(x)$ és un polinomi de coeficients racionals, és fàcil trobar un polinomi de coeficients enters amb les mateixes arrels que $p(x)$, al qual podrem aplicar les idees anteriors: només cal multiplicar $p(x)$ pel denominador comú de tots els seus coeficients.

Per al càlcul d'arrels irracionals, esmentem només el mètode de les aproximacions successives basat en el teorema de Bolzano (si és possible, convé fer primer un esboç de la gràfica de p).

Finalment, recordem algunes idees elementals que sovint són útils per a la descomposició d'un polinomi en factors: les anomenades igualtats notables, la resolució d'equacions de segon grau (aplicable també a les equacions biquadrades) i treure factor comú la màxima potència possible de x quan el terme independent és zero.

(Per desgràcia, cap de les demostracions del teorema fonamental de l'àlgebra és constructiva, en el sentit de permetre la localització efectiva de l'arrel de la qual se'n garanteix l'existència.)

L'última propietat que indicarem és un resultat interessant, sobre tot per a l'estudi de les formes quadràtiques i les quàdriques, ja que en aquests dos casos treballem amb el polinomi característic d'una matriu simètrica que, segons el teorema espectral, sabem que té totes les arrels reals.

Teorema B.8 *(Teorema de Descartes.) Si $p(x)$ és un polinomi de coeficients reals i totes les arrels de $p(x)$ són reals, el nombre de variacions de signe entre coeficients consecutius (no nuls, lògicament) de $p(x)$ coincideix amb el nombre d'arrels positives de $p(x)$ —comptant les multiplicitats—.*

DEMOSTRACIÓ. La justificació no és difícil, i es pot trobar a qualsevol tractat d'àlgebra que inclogui un estudi detallat dels polinomis. □

Exemples B.9 (a) El polinomi $p(x) = x^3 - 16x^2 - 9x + 28$ és el polinomi característic d'una matriu simètrica, i per tant té tres arrels reals. Com que no té l'arrel 0 i presenta dues variacions de signe, sabem que tindrà dues arrels positives i una arrel negativa (el lector pot confirmar que les arrels de $p(x)$ són, aproximadament, $-1'54$, $1'10$ i $16'44$).

(b) Per la mateixa raó, sabem que el polinomi $q(x) = x^5 - 8x^3 + 4x^2$ té cinc arrels reals. Hi ha dues variacions de signe, per tant té dues arrels positives, l'arrel 0 doble i una arrel negativa.

(c) En canvi, el polinomi $r(x) = 2x^3 - 6x + 5$ té una sola arrel real i dues complexes conjugades. L'aplicació (incorrecta) del teorema de Descartes diria que té dues arrels positives i una negativa.

Bibliografia

[1] **Amer, R. i Domínguez, J. M.**
Àlgebra Lineal. Problemes resolts.
Ed. UPC. Barcelona. 1992.

[2] **Amer, R. Carreras, F. i Tudurí, J.**
Àlgebra Lineal. Problemes, exercicis i qüestions.
Ed. Cardellach Còpies. Terrassa. 1998.

[3] **Anton, H.**
Introducción al álgebra lineal.
Ed. Limusa. México. 1989.

[4] **Burgos, J. de**
Algebra Lineal.
Ed. McGraw Hill. Madrid. 1993.

[5] **Castellet, M. i Llerena, I.**
Àlgebra lineal i geometria.
Publicacions de la Universitat Autònoma de Barcelona. 1988.

[6] **Granero, F.**
Algebra y geometría analítica.
Ed. McGraw. Madrid. 1985.

[7] **Lang, S.**
Algebra lineal.
Ed. Addison–Wesley. México. 1986.

[8] **Moreno, J. M.**
Una introducción al álgebra lineal elemental.
Dep. de Matemáticas. U.A.B. Bellaterra. 1988.

[9] **Xambó, S.**
Algebra lineal y geometrias lineales (Tomos I y II).
Ed. Eunibar. Barcelona.

Índex alfabètic