

Apunts d'Àlgebra Lineal

MatCAD - Universitat Autònoma de Barcelona

28 de novembre de 2018

Índex

1	Introducció	2
2	Matrius i equacions lineals	2
2.1	Operacions amb matrius. Matriu invertible	4
2.2	Transformacions elementals en matrius	7
2.3	Criteri d'invertibilitat. Rang d'una matriu	10
2.4	Resolució de sistemes d'equacions lineals	12
3	Espais vectorials i aplicacions lineals	16
3.1	Matrius com a aplicacions lineals	16
3.2	Aplicacions lineals i geometria	18
3.3	Subespais, generadors i bases	20
3.4	Suma i intersecció de subespais vectorials	30
3.5	Aplicacions injectives, exhaustives i bijectives	34
3.6	Coordenades de vectors	35
3.7	Espais vectorials	39
4	Diagonalització	45
4.1	Motivació	46
4.2	Determinants	46
4.3	Polinomi característic. Valors i vectors propis	53
4.4	Vectors propis associats a un valor propi	57
4.5	Matrius sobre \mathbb{R}	62
5	Ortogonalitat	63
5.1	Ortogonalitat a \mathbb{R}^n	63
5.2	El teorema de Pitàgores	65
5.3	Mètode de Gram-Schmidt	66
5.4	Aplicacions i matrius ortogonals	69
5.5	Matriu d'una projecció ortogonal en una base ortonormal	71
5.6	Mínims quadrats	71
5.7	Formes bilineals i productes escalars	76
5.8	Tota matriu simètrica sobre \mathbb{R} diagonalitza	78
5.9	Descomposició en valors singulars	80
5.10	Classificació de formes bilineals simètriques sobre \mathbb{R}^n	82
A	Nombres complexos	86

1 Introducció

Encara que el curs serà força autocontingut es requerirà que l'alumne conegui la resolució de sistemes d'equacions lineals, l'aritmètica bàsica de números i polinomis, i que tingui destresa de càlcul amb expressions algebraiques simbòliques.

A tot aquest curs suposarem que treballem sobre un cos commutatiu \mathbb{K} fixat, que podeu pensar és \mathbb{Q} , \mathbb{R} o \mathbb{C} . Els elements de K els anomenarem nombres o escalars. Les propietats que utilitzarem són:

- És commutatiu amb la suma: $a + b = b + a \ \forall a, b \in \mathbb{K}$.
- És commutatiu amb el producte: $ab = ba \ \forall a, b \in \mathbb{K}$.
- La suma té un element neutre que anomenem zero: $0 + a = a \ \forall a \in \mathbb{K}$.
- El producte té un elements neutre que anomenem u: $1a = a \ \forall a \in \mathbb{K}$.
- Tot element $a \in \mathbb{K}$ té un invers per la suma que anomenem $-a$: $a + (-a) = 0$.
- Tot element a diferent de zero té un invers per la multiplicació que anomenem $1/a$ o bé a^{-1} : $aa^{-1} = 1$.
- Hi ha les propietats associatives a la suma i al producte: $(a + b) + c = a + (b + c)$ i $(ab)c = a(bc) \ \forall a, b, c \in \mathbb{K}$.
- Hi ha la propietat distributiva: $a(b + c) = ab + ac \ \forall a, b, c \in \mathbb{K}$.

També suposarem certa familiaritat amb el llenguatge dels conjunts. Si A és un conjunt i B un subconjunt d' A , escriurem $B \subset A$. $a \in A$ voldrà dir que a és un element d' A . També escriurem $A \setminus B = \{a \in A \mid a \notin B\}$ i llegirem els a que pertanyen a A i que no pertanyen a B (o bé el complementari de B en A).

2 Matrius i equacions lineals

El contingut d'aquesta secció el podem trobar a [1, Tema 1] i a [2, Tema 2].

Definició 2.1. Si m i n són dos nombres naturals, una *matriu* $m \times n$ amb entrades a \mathbb{K} és una taula rectangular d'elements de \mathbb{K} amb m files i n columnes. Denotem $M_{m \times n}(\mathbb{K})$ al conjunt de matrius que tenen m files i n columnes i els seus elements són de \mathbb{K} .

Notació 2.2. Denotarem amb lletres majúscules en nom de les matrius i amb la mateixa lletra i subíndexs cadascun dels coeficients: si A és una matriu, anomenarem a_{ij} al nombre de la fila i , columna j . En el producte de matrius, a vegades utilitzarem la notació $(AB)_{ij}$ per a fer referència al coeficient de la posició (i, j) després de fer el producte.

Exemple 2.3. Si $A = \begin{pmatrix} 1 & 3 & 0 \\ 0 & -1 & 1 \end{pmatrix} \in M_{2 \times 3}(\mathbb{Q})$, llavors $a_{11} = 1$, $a_{12} = 3$, $a_{13} = 0$, $a_{21} = 0$, $a_{22} = -1$ i $a_{23} = 1$.

A continuació fixem algunes notacions i definicions de casos particulars:

- Una *matriu quadrada* és una matriu amb el nombre de columnes igual al nombre de files. Denotarem per $M_n(\mathbb{K}) = M_{n \times n}(\mathbb{K})$.

- Un element està a la *diagonal d'una matriu quadrada* si la posició que ocupa té el mateix nombre de fila que de columna: si la matriu és A , els elements de la diagonal són els a_{ii} .
- Una *matriu diagonal* és una matriu quadrada on els únics elements no nuls estan a la diagonal: A , matriu quadrada, és diagonal si $a_{ij} = 0 \forall i \neq j$.
- La *matriu identitat* $n \times n$ és una matriu diagonal on tots els elements de la diagonal valen 1 (i per tant els altres valen 0). La denotem per $\mathbf{1}_n$ la matriu identitat $n \times n$.
- En general, escriurem els vectors per columnes: un *vector de* \mathbb{K}^n és una matriu amb n files i 1 columna.
- Direm que una matriu quadrada A és *triangular superior* si tots els coeficients per sota de la diagonal valen 0, o sigui, $a_{ij} = 0$ si $i > j$.
- Direm que una matriu quadrada A és *triangular inferior* si tots els coeficients per sobre de la diagonal valen 0, o sigui, $a_{ij} = 0$ si $i < j$.
- Donada una matriu $A \in M_{m \times n}(\mathbb{K})$, definim la *transposada d'A* i la denotem per A^T com la matriu de $M_{n \times m}(\mathbb{K})$ que té per columnes les files d' A , o sigui, que a la posició (i, j) té el coeficient a_{ji} . Tenim la propietat:

$$(A^T)^T = A.$$

- Diem que una matriu A és *simètrica* si $A = A^T$ (en particular, ha de ser quadrada). Per exemple, la matriu identitat $\mathbf{1}_n$ és simètrica.

Exemple 2.4. Considerem un sistema d'equacions amb m equacions i n incògnites:

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

D'aquí podem treure la matriu associada als coeficients del sistema:

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

el vector de termes independents:

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

o bé escriure-ho tot en una sola matriu (matriu ampliada), on habitualment separem els termes independents:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right)$$

2.1 Operacions amb matrius. Matriu invertible

Considerem $\lambda \in \mathbb{K}$ i $A, B \in M_{m \times n}(\mathbb{K})$. Les primeres operacions que podem fer són les que corresponen a veure $M_{m \times n}(\mathbb{K})$ com \mathbb{K} -espai vectorial (més endavant veurem què vol dir):

- Definim la matriu λA com la matriu que té per coeficients λa_{ij} .
- Definim la matriu $A + B$ com la matriu que té per coeficients $a_{ij} + b_{ij}$.

Exemple 2.5.

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

Exemple 2.6.

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

Aquestes definicions, més les propietats dels elements de \mathbb{K} impliquen:

1. Existeix una matriu $\mathbf{0}_{mn}$ complint $A + \mathbf{0}_{mn} = A$, $\forall A \in M_{mn}(\mathbb{K})$ ($\mathbf{0}_{mn}$ té tots els coeficients zero).
2. $0A = \mathbf{0}_{mn}$, $\forall A \in M_{m \times n}(\mathbb{K})$.
3. $A + B = B + A$, $\forall A, B \in M_{m \times n}(\mathbb{K})$.
4. $1A = A$, $\forall A \in M_{m \times n}(\mathbb{K})$.
5. $\lambda(\mu A) = (\lambda\mu)A$, $\forall \lambda, \mu \in \mathbb{K}$ i $\forall A \in M_{m \times n}(\mathbb{K})$

Abans de la definició del producte de matrius és convenient utilitzar el llenguatge següent:

Definició 2.7. Diem que un vector \vec{w} és combinació lineal de $\{\vec{v}_1, \dots, \vec{v}_n\}$ si existeixen escalars $\lambda_1, \dots, \lambda_n$ tals que $w = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$.

Si un dels vectors de $\{\vec{v}_1, \dots, \vec{v}_n\}$ es pot escriure com a combinació lineal dels altres, diem que la família de vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ és linealment dependent.

En cas contrari, diem que la família de vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ és linealment independent.

Parlarem de files (o columnes) linealment dependents o independents d'una matriu A pensades com a vectors amb tantes components com files (o columnes) tingui A .

Exercici 2.1. Demostreu que $\{\vec{v}_1, \dots, \vec{v}_n\}$ són linealment independents si l'única manera d'escriure el vector zero $\vec{0} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$ és amb $\lambda_1 = \dots = \lambda_n = 0$.

A més, a més, podem definir:

Definició 2.8. Si $A \in M_{m \times n}(\mathbb{K})$, $B \in M_{n \times r}(\mathbb{K})$ (o sigui, el nombre de columnes de A és igual al nombre de files de B) podem definir el producte AB com la matriu $C \in M_{m \times r}$ que té per coeficients:

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

La matriu C és pot pensar que té per columnes combinacions lineals de columnes de A (B ens diu quines són aquestes combinacions lineals): la columna j de la matriu C és:

$$\begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{mj} \end{pmatrix} = b_{1j} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + b_{2j} \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + b_{nj} \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

Anàlogament, la matriu C es pot pensar que té per files combinacions lineals de files de B (A ens diu quines són aquestes combinacions lineals): la fila i de la matriu C és (separem amb comes per a que quedi més clar):

$$(c_{i1}, c_{i2}, \dots, c_{ir}) = a_{i1}(b_{11}, b_{12}, \dots, b_{1r}) + a_{i2}(b_{21}, b_{22}, \dots, b_{2r}) + \dots + a_{in}(b_{n1}, b_{n2}, \dots, b_{nr})$$

Exemple 2.9.

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 12 & 16 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 2 \\ 1 & 11 & 4 \\ 1 & 17 & 6 \end{pmatrix}$$

Exemple 2.10. Podem escriure el sistema d'equacions

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

com $AX = B$, on

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{i} \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Proposició 2.11. El producte de matrius té les propietats següents:

- (a) *Element neutre:* si $A \in M_{m \times n}(\mathbb{K})$, $\mathbf{1}_m A = A \mathbf{1}_n = A$.
- (b) *Propietat associativa:* si $A \in M_{m \times n}(\mathbb{K})$, $B \in M_{n \times r}(\mathbb{K})$, $C \in M_{r \times s}(\mathbb{K})$, llavors

$$(AB)C = A(BC).$$

- (c) *Distributiva respecte el producte:* si $A \in M_{m \times n}(\mathbb{K})$, $B, C \in M_{n \times r}(\mathbb{K})$ i $D \in M_{r \times s}(\mathbb{K})$, llavors $A(B + C) = AB + AC$ i $(B + C)D = BD + CD$.

Demostració. Escriure les fórmules amb els coeficients i surt. Fem la propietat associativa com exemple: volem comparar el coeficient a la posició (i, j) d' $(AB)C$ amb el d' $A(BC)$, que denotem $((AB)C)_{ij}$ i $(A(BC))_{ij}$ respectivament:

$$((AB)C)_{ij} = \sum_{k=1}^r (AB)_{ik} c_{kj} = \sum_{k=1}^r \left(\sum_{l=1}^n a_{il} b_{lk} \right) c_{kj} = \sum_{k=1}^r \sum_{l=1}^n a_{il} b_{lk} c_{kj}$$

Mentre que:

$$(A(BC))_{ij} = \sum_{l=1}^n a_{il} (BC)_{lj} = \sum_{l=1}^n a_{il} \left(\sum_{k=1}^r b_{lk} c_{kj} \right) = \sum_{l=1}^n \sum_{k=1}^r a_{il} b_{lk} c_{kj}$$

I els dos resultats són el mateix ja que podem commutar els sumatoris. □

Observació 2.12. El producte de matrius, en general, no és commutatiu (veure l'Exemple 2.9).

Proposició 2.13. Si $A \in M_{m \times n}(\mathbb{K})$ i $B \in M_{n \times r}(\mathbb{K})$, llavors tenim la relació següent entre productes i transposades

$$(AB)^T = B^T A^T$$

Demostració. Escrivim les fórmules dels coeficients:

$$((AB)^T)_{ij} = (AB)_{ji} = \sum_{k=1}^n a_{jk} b_{ki}$$

mentre que

$$(B^T A^T)_{ij} = \sum_{k=1}^n (B^T)_{ik} (A^T)_{kj} = \sum_{k=1}^n b_{ki} a_{jk}$$

i són iguals per la propietat commutativa del producte a \mathbb{K} . \square

Definició 2.14. Si considerem \vec{v} i \vec{w} vectors de \mathbb{K}^n , que escrivim com una columna cadascun, definim el *producte escalar* $\vec{v} \cdot \vec{w}$ com:

$$\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$$

Per tant, si les coordenades són:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \text{ llavors } \vec{v} \cdot \vec{w} = \sum_{i=1}^n v_i w_i.$$

Volem definir la inversa d'una matriu quadrada, però com que el producte no és commutatiu, hauríem de parlar d'inversa per l'esquerra o per la dreta.

Definició 2.15. Diem que una *matriu quadrada* $A \in M_n(\mathbb{K})$ és *invertible* si existeix una matriu $B \in M_n(\mathbb{K})$ tal que $AB = BA = \mathbf{1}_n$.

Teorema 2.16. Si A és una matriu quadrada $n \times n$ i B és una matriu també $n \times n$. Llavors $AB = \mathbf{1}_n$ si i només si $BA = \mathbf{1}_n$. En particular, si això passa, A és invertible i la inversa és única. Denotem la matriu B com A^{-1} .

Demostració. La primera part la demostrarem quan tinguem el concepte de rang d'una matriu (Observació 2.30). Vegem ara que si B és tal que $AB = \mathbf{1}_n$, i C tal que $CA = \mathbf{1}_n$, llavors $B = C$:

$$C = C\mathbf{1}_n = C(AB) = (CA)B = \mathbf{1}_n B = B.$$

Això implica que la inversa és única: Suposem B' tal que $AB' = \mathbf{1}_n$, llavors, pel raonament d'abans, $B' = C = B$. \square

Proposició 2.17. (a) Si $A \in M_{n \times n}(\mathbb{K})$ és invertible, llavors A^T també ho és i $(A^T)^{-1} = (A^{-1})^T$.

(b) Si $A, B \in M_{n \times n}(\mathbb{K})$ són matrius invertibles, llavors el producte AB també ho és i $(AB)^{-1} = B^{-1}A^{-1}$.

Demostració. Demostrem primer (a): com que ens proposen una inversa, tant sols cal comprovar que ho és:

$$(A^T)(A^{-1})^T = (A^{-1}A)^T = \mathbf{1}_n^T = \mathbf{1}_n.$$

per tant, $(A^{-1})^T$ és inversa d' A^T (com que A^T és quadrada, pel Teorema 2.16, tant sols cal comprovar-ho per un dels costats).

Demostrem ara (b): com que A i B són invertibles, existeixen les matrius inverses A^{-1} i $B^{-1} \in M_{n \times n}(\mathbb{K})$ respectivament. Llavors:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathbf{1}_nA^{-1} = AA^{-1} = \mathbf{1}_n$$

Igual que (a), això ja demostra que $B^{-1}A^{-1}$ és inversa d' AB . □

Observació 2.18. D'aquí es dedueix que si $A_1, \dots, A_r \in M_{n \times n}(\mathbb{K})$ són invertibles, el seu producte també ho és i $(A_1 \cdots A_r)^{-1} = A_r^{-1} \cdots A_1^{-1}$.

2.2 Transformacions elementals en matrius

Considerem una matriu organitzada per files (encara que tot el que farem aquí també es pot fer per columnes) i definim les *transformacions elementals*:

T1. Multiplicar una de les files per $\lambda \neq 0$.

T2. Sumar a una de les files μ vegades una altra fila.

T3. Intercanviar dues files.

Observació 2.19. Si considerem la matriu d'un sistema d'equacions i apliquem qualsevol de les transformacions elementals, no modifiquem les solucions.

Observació 2.20. Les tres transformacions elementals es poden desfer mitjançant una transformació elemental:

1. Multiplicar una de les files per $1/\lambda \neq 0$.

2. Sumar a una de les files $-\mu$ vegades una altra fila.

3. Intercanviar dues files.

Aquestes transformacions elementals es poden fer (i desfer) multiplicant per una matriu invertible P per l'esquerra. Suposem que A és una matriu amb m files:

T1. si apliquem aquest canvi a la fila i d' $\mathbf{1}_m$ tindrem una matriu P on hem modificat el 1 de la fila i per $\lambda \neq 0$. Llavors PA té els mateixos valors que A , però amb la fila i multiplicada per λ . En aquest cas, P^{-1} és una matriu identitat amb un $1/\lambda$ a la posició (i, i) .

T2. si sumem a la fila i de la matriu $\mathbf{1}_n$ μ vegades la fila $k \neq i$ tindrem una matriu P tal que PA té a la fila i la fila i d' A més μ vegades la fila k d' A . En aquest cas, P^{-1} és una matriu amb 1 a la diagonal, 0 fora, excepte la posició (k, i) , que val $-\mu$.

T3. si intercanviem les files i i k de la matriu $\mathbf{1}_n$ ($i \neq k$) obtenim una matriu P tal que PA és el resultat d'intercanviar les files i i k d' A . En aquest cas, $P^{-1} = P$.

Amb aquests raonaments hem demostrat:

Proposició 2.21. Si considerem la matriu identitat $\mathbf{1}_m$ i li apliquem transformacions elementals per files, la matriu P que obtenim és invertible. Si apliquem exactament les mateixes transformacions elementals a una altra matriu $A \in M_{m \times n}(\mathbb{K})$, la matriu que resulta és exactament PA .

Definició 2.22. Diem que dues matrius A i B són equivalents per files si es pot passar d' A a B mitjançant transformacions elementals per files. Escriurem $A \sim B$.

Proposició 2.23. Si A , B i C són matrius de dimensions iguals, aleshores es té:

- $A \sim A$ (reflexiva),
- $A \sim B$ si i només si $B \sim A$ (simètrica),
- $A \sim B$ i $B \sim C$ implica $A \sim C$ (transitiva).

Demostració. La primera és **T2** per a qualsevol fila i $\mu = 0$.

La segona és que la inversa d'una transformació elemental és una transformació elemental (i les posem en ordre invers): si P_1, \dots, P_r són les transformacions elementals que apliquem a A per obtenir B , resulta que:

$$P_r \cdots P_1 A = B$$

llavors

$$A = P_1^{-1} \cdots P_r^{-1} B$$

Però si P_i és una transformació elemental, P_i^{-1} també i per tant $B \sim A$.

La tercera és per definició de \sim , ja que passem de A a C fent primer els canvis elementals que transformen A en B i després els que transformen B en C . \square

Observació 2.24. Si tenim un conjunt S i una relació \sim entre els seus elements complint les propietats de la Proposició 2.23 (reflexiva, simètrica i transitiva) diem que és una *relació d'equivalència*.

Definició 2.25. Diem que una matriu A està en forma esglaonada (per files) si compleix que:

1. El primer element no nul de cada fila val 1, i l'anomenem *pivot*.
2. Si una fila conté un pivot a la columna j , les files superiors també tenen un pivot a una columna $j' < j$.
3. Totes les entrades per sota d'un pivot valen 0.

Diem que una matriu A està en forma reduïda (per files) si està esglaonada i a més compleix que:

- 3' Si una columna té un pivot, aquest és l'únic element no nul de la seva columna.

Exemple 2.26. De les matrius següents:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ i } C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \end{pmatrix}$$

la matriu A està reduïda per files, mentre que B i C no ho són. La matriu D està esglaonada, però no reduïda.

Considerem ara l'algorisme següent (**Mètode de Gauss** o **Mètode de Gauss-Jordan**), que aplica canvis elementals per files a una matriu A fins que obtenim una matriu en forma reduïda per files. A aquest mètode també li diem **triangular la matriu per files**.

Suposem que tenim una matriu $A \in M_{m \times n}(\mathbb{K})$, on la submatriu formada per les i_1 primeres files ja té forma reduïda i si el pivot de la fila i_1 és a la posició j_1 , els coeficients $a_{ij} = 0$ si $i > i_1$ i $j < j_1$.

G1 Busquem la primera columna que no sigui tota zero a la submatriu formada per les files $i_1 + 1$ fins la m (suposem columna j_2), i triem una fila a aquesta submatriu on hi hagi un element diferent de zero a aquesta columna. Si cal, apliquem el canvi **T3** per a que la fila $i_2 := i_1 + 1$ tingui un valor diferent de zero a aquesta columna. Podem suposar, doncs, que $a_{i_2 j_2} \neq 0$ i que $a_{ij} = 0$ si $i > i_2$ i $j < j_2$.

G2 Volem que a la posició (i_2, j_2) hi hagi un pivot ($= 1$), per tant multipliquem tota la fila i_2 per $1/a_{i_2 j_2}$ (canvi **T1**).

G3 Per a totes les files $i \neq i_2$ tals que $a_{ij_2} \neq 0$, hi restem la fila i_2 multiplicada per a_{ij_2} (canvi **T2**), de tal manera que farem que a la posició (i, j_2) hi hagi un zero per a tota $i \neq i_2$.

Ara hem aconseguit que la submatriu formada per les files $1, \dots, i_1 + 1$ sigui en forma reduïda. Iterem aquest procediment fins que $i_1 = m$.

Com que la matriu té un nombre finit de files, aquest algorisme sempre acaba. A més, acabem de demostrar que:

Teorema 2.27. *Tota matriu $A \in M_{m \times n}(\mathbb{K})$ és equivalent a una matriu reduïda.*

Exemple 2.28. Considerem la matriu:

$$A = \begin{pmatrix} 0 & 1 & 3 & 0 \\ 1 & -2 & -5 & 4 \\ 2 & -3 & -7 & 7 \end{pmatrix}$$

Apliquem el canvi **T3**, canviant la primera fila per la segona, obtenint:

$$\begin{pmatrix} 1 & -2 & -5 & 4 \\ 0 & 1 & 3 & 0 \\ 2 & -3 & -7 & 7 \end{pmatrix}$$

Apliquem el canvi **T2**, restant a la tercera fila 2 cops la primera:

$$\begin{pmatrix} 1 & -2 & -5 & 4 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 3 & -1 \end{pmatrix}$$

Com que a la segona fila, la primera posició ja és 1, podem utilitzar-la directament de pivot i sumar-la 2 cops a la primera fila, i restar-la a la tercera:

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

Les dues primeres files ja estan en forma reduïda, pel que podem considerar la tercera fila, i multiplicar-la per -1 per a que l'única posició no nul·la sigui un pivot:

$$\begin{pmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Finalment restem la tercera fila a la primera multiplicada per 4:

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2.3 Criteri d'invertibilitat. Rang d'una matriu

Considerem primer el resultat següent:

Teorema 2.29. *Donada una matriu quadrada $A \in M_{n \times n}(\mathbb{K})$, les condicions següents són equivalents:*

- (a) A és equivalent a $\mathbf{1}_n$.
- (b) A és invertible.

Demostració. Vegem primer (a) implica (b): si A és equivalent a $\mathbf{1}_n$, existeix P tal que $PA = \mathbf{1}_n$. Com que P és invertible, existeix P^{-1} i podem fer:

$$PA = \mathbf{1}_n \Rightarrow P^{-1}PA = P^{-1} \Rightarrow A = P^{-1} \Rightarrow AP = P^{-1}P = \mathbf{1}_n$$

per tant P és la inversa d' A .

Vegem (b) implica (a): suposem que A no és equivalent a $\mathbf{1}_n$. Llavors, forçosament, quan l'esglaonem hi haurà una fila sense pivot, i per tant, tot zeros. Llavors obtenim $PA = A'$, amb P invertible i A' una matriu que té l'última fila tot zeros. Si A fos invertible, voldria dir que existeix Q tal que $AQ = \mathbf{1}_n$, llavors, però llavors:

$$AQ = \mathbf{1}_n \Rightarrow PAQ = P \Rightarrow A'Q = P \Rightarrow A'(QP^{-1}) = \mathbf{1}_n$$

Observem ara que si A' té l'última fila tot zeros $A'(QP^{-1})$ també, i per tant no pot ser $\mathbf{1}_n$. \square

Observació 2.30. Aquests resultats donen la demostració de la primera part del Teorema 2.16.

Observació 2.31. Si considerem $A \in M_{n \times n}(\mathbb{K})$ que es pot subdividir en 4 submatrius:

$$A = \left(\begin{array}{c|c} B & C \\ \hline \mathbf{0} & D \end{array} \right)$$

on la matriu $\mathbf{0}$ està formada per zeros i toca la diagonal (conté un coeficient amb coordenades (i, i)), llavors A no és invertible.

Demostració. Primer observem que per a que una matriu sigui invertible, l'esglaonament ha de fer que a totes les seves columnes (i files) hi hagi un pivot (es dedueix del Teorema 2.29).

Si fos invertible podríem fer els canvis elementals per files i obtenir $\mathbf{1}_n$. Com que la submatriu $\mathbf{0}$ toca la diagonal, llavors B té més columnes que files i D més files que columnes. Quan esglaonem la matriu A , el pivot de les primeres columnes ha de ser a B (sota hi ha zeros) i per tant no hi pot haver més pivots que files a B (a les columnes de B), per tant hi ha columnes de A sense pivot, i per tant no pot ser invertible. \square

Podem utilitzar el concepte de matriu reduïda equivalent per a definir el rang d'una matriu A . Cal tenir en compte que necessitarem demostrar que la definició no depèn de quina matriu reduïda equivalent a A considerem.

Proposició 2.32. *Si A i B són dues matrius en forma reduïda que són equivalents, llavors $A = B$.*

Demostració. Suposem que A i B són diferents. Considerem submatrius A' i B' formades per la primera columna on difereixin, així com per totes les columnes a l'esquerra d'aquesta que continguin pivots. Observem que A' i B' segueixen essent equivalents, mitjançant les mateixes transformacions elementals associades a A i B . Podem interpretar A' i B' com matrius augmentades de sistemes lineals equivalents, posem en r incògnites. Si aquests són compatibles, aleshores necessàriament els termes independents han de coincidir i, per tant obtenim una contradicció. Si aquests dos sistemes són incompatibles, això significa que l'última columna conté zeros en les primeres files (tantes com pivots) i necessàriament tindrem que aquesta última columna tindrà un pivot a l'entrada r -èsima. Per tant, obtenim una nova contradicció. \square

Notació 2.33. D'aquí es dedueix que donada una matriu $A \in M_{n \times n}(\mathbb{K})$, existeix una sola matriu reduïda equivalent a A , i l'anomenem $\text{rref}(A)$ (*reduced row-echelon form*).

Ara ja podem definir el rang d'una matriu:

Definició 2.34. Donada una matriu A , definim el *rang* d' A com el nombre de files diferents de zero (igual al nombre de pivots) de $\text{rref}(A)$.

Corol·lari 2.35. Si fem transformacions elementals a una matriu $A \in M_{m \times n}(\mathbb{K})$, el rang de la matriu resultant és el mateix. Això també ho podem enunciar dient que si $P \in M_{m \times n}(\mathbb{K})$ és una matriu invertible i $A \in M_{m \times n}(\mathbb{K})$, llavors $\text{Rang}(A) = \text{Rang}(PA)$.

Corol·lari 2.36. Una matriu quadrada $A \in M_{n \times n}(\mathbb{K})$ és invertible si i només si té rang n .

Demostració. Considerem el Teorema 2.29, i per tant és invertible si i només si és equivalent a $\mathbf{1}_n$, per tant, si i només si té rang n . \square

El Mètode de Gauss ens dona una manera de calcular la inversa d'una matriu: suposem que ens donen una matriu $A \in M_{n \times n}(\mathbb{K})$. Considerem la matriu formada per una matriu identitat $n \times n$ al costat de la matriu A :

$$(A \mid \mathbf{1}_n) \in M_{n \times 2n}(\mathbb{K})$$

Si apliquem canvis per files a aquesta matriu (cada fila té $2n$ coeficients) obtindrem matrius que podem escriure com:

$$(A' \mid P) \in M_{n \times 2n}(\mathbb{K})$$

amb $A', P \in M_{n \times n}(\mathbb{K})$ i que compleixen que $PA = A'$.

Com a cas particular tenim que, si A és invertible, podem arribar a la situació

$$(\mathbf{1}_n \mid P) \in M_{n \times 2n}(\mathbb{K}) \tag{1}$$

i tindrem que $P = A^{-1}$.

Si la matriu A no fos invertible, el Teorema 2.29 ens diu que no seria possible arribar a la situació de l'Equació (1).

Exemple 2.37. Considerem la matriu:

$$A = \begin{pmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 5 & 6 & 0 \end{pmatrix}$$

Escrivim la matriu amb una còpia de la matriu identitat a la dreta:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 6 & 1 & 0 & 0 \\ 0 & -1 & -8 & 0 & 1 & 0 \\ 5 & 6 & 0 & 0 & 0 & 1 \end{array} \right)$$

Esglaonem la part matriu resultant segon el mètode de Gauss:

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 2 & 6 & 1 & 0 & 0 \\ 0 & -1 & -8 & 0 & 1 & 0 \\ 0 & -4 & -30 & -5 & 0 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 6 & 1 & 0 & 0 \\ 0 & 1 & 8 & 0 & -1 & 0 \\ 0 & -4 & -30 & -5 & 0 & 1 \end{array} \right) \\ \left(\begin{array}{ccc|ccc} 1 & 0 & -10 & 1 & 2 & 0 \\ 0 & 1 & 8 & 0 & -1 & 0 \\ 0 & 0 & 2 & -5 & -4 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 6 & 1 & 0 & 0 \\ 0 & 1 & 8 & 0 & -1 & 0 \\ 0 & 0 & 1 & -5/2 & -2 & 1/2 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -24 & -18 & 5 \\ 0 & 1 & 0 & 20 & 15 & -4 \\ 0 & 0 & 1 & -5/2 & -2 & 1/2 \end{array} \right) \end{aligned}$$

Per tant:

$$A^{-1} = \begin{pmatrix} -24 & -18 & 5 \\ 20 & 15 & -4 \\ -5/2 & -2 & 1/2 \end{pmatrix}$$

Exercici 2.2. Tot el que hem fet per files té el seu anàleg per columnes. Una manera senzilla d'adaptar totes les definicions i resultats és dir que la transposada compleix la definició per files. Per exemple:

La matriu A és *en forma reduïda per columnes* si A^T és en forma reduïda per files.

- (a) Enuncieu l'anàleg per columnes de cada resultat que s'ha vist a aquest capítol per files.
- (b) Demostreu que $\text{Rang}(A) = \text{Rang}(A^T)$.

Exercici 2.3. Considerem $A \in M_{m \times n}(\mathbb{K})$ i hi afegim una fila, obtenint $A' \in M_{(m+1) \times n}(\mathbb{K})$. Demostreu $\text{Rang}(A) = \text{Rang}(A')$ si i només si la fila que hem afegit és combinació lineal de les files d' A .

Exercici 2.4. Demostreu que el rang d'una matriu A és el nombre màxim de files (o columnes) linealment independents que conté A .

2.4 Resolució de sistemes d'equacions lineals

Recordem la notació d'un sistema d'equacions:

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

Que també escrivim com $AX = B$, on

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

Volem esbrinar si el sistema té solució o no, i, en cas de tenir-ne, saber quantes en té i calcular-les.

Una primera interpretació és considerar que tenim una solució X i fer el càlcul següent:

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

Per tant, una solució ens dona B com a combinació lineal de les columnes d' A , i al revés: si podem escriure B com a combinació lineal de les columnes d' A , tenim una solució.

Argumentant amb el rang per columnes, ja tenim un criteri per saber si un sistema d'equacions lineals té solució o no:

Proposició 2.38. *El sistema d'equacions:*

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

té solució si i només si:

$$\text{Rang} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \text{Rang} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

Demostració. Denotem per A la matriu associada al sistema i per \bar{A} la matriu ampliada. Esclaonem per files la matriu \bar{A} fins a tenir una matriu reduïda $P\bar{A}$ i fem exactament les mateixes transformacions a la matriu A , obtenint PA . Com que les primeres n columnes d' \bar{A} són precisament les d' A , tenim que PA també és una matriu en forma reduïda, i tants sols difereixen de que $P\bar{A}$ té una columna més.

Si $\text{Rang}(A) = \text{Rang}(\bar{A})$, vol dir que a l'última columna de $P\bar{A}$ no hi ha cap pivot, i per tant es pot escriure l'última columna de $P\bar{A}$ com a combinació lineal de les n primeres, obtenint una solució del sistema.

Si $\text{Rang}(A) \neq \text{Rang}(\bar{A})$, per força $\text{Rang}(\bar{A}) = \text{Rang}(A) + 1$ i l'última columna té un pivot. La fila on hi ha el pivot de l'última columna correspon a l'equació $0 = 1$, que no té solució. \square

El resultat anterior també es pot enunciar dient que si esclaonem per files la matriu ampliada del sistema, no queda cap fila on l'únic element no nul sigui a la columna de termes independents.

Com que les transformacions elementals per files a la matriu ampliada no modifiquen les solucions del sistema, suposem que hem esclaonat la matriu inicial del sistema. Llavors, obtenint una matriu reduïda \bar{A}' :

- El sistema té solució si i només si l'última columna d' \bar{A}' no té cap pivot. Si el sistema té solució, diem que el sistema és compatible, i quan no té solució, diem que és un sistema incompatible.
- Suposem que el sistema té solució, llavors cada pivot ens permet aïllar la variable corresponent a la seva columna.

- (c) Continuem suposant que el sistema té solució: les columnes (de la matriu sense ampliar) que no tenen pivot definiran el que anomenem *paràmetres lliures*. N'hi haurà tants com $k := n - \text{Rang}(A)$, on n és el número de incògnites. En aquest cas direm que *la solució té dimensió k* .

En el cas particular que $k = 0$ diem que té solució única i diem que és un *sistema compatible determinat*.

Si $k > 0$ diem que és un *sistema compatible indeterminat amb k paràmetres lliures*.

Exemple 2.39. Considerem el sistema d'equacions:

$$\begin{aligned}x - y + 2z + 3t &= 21 \\ -x + 2y + z + 5t &= 26 \\ 3x + y - 2z + t &= -9 \\ 3x + 2y + z + 9t &= 38\end{aligned}$$

Considerem la matriu ampliada i esglaonem:

$$\begin{aligned}\left(\begin{array}{cccc|c}1 & -1 & 2 & 3 & 21 \\ -1 & 2 & 1 & 5 & 26 \\ 3 & 1 & -2 & 1 & -9 \\ 3 & 2 & 1 & 9 & 38\end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|c}1 & -1 & 2 & 3 & 21 \\ 0 & 1 & 3 & 8 & 47 \\ 0 & 4 & -8 & -8 & -72 \\ 0 & 5 & -5 & 0 & -25\end{array}\right) \\ \left(\begin{array}{cccc|c}1 & 0 & 5 & 11 & 68 \\ 0 & 1 & 3 & 8 & 47 \\ 0 & 0 & -20 & -40 & -260 \\ 0 & 0 & -20 & -40 & -260\end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|c}1 & 0 & 5 & 11 & 68 \\ 0 & 1 & 3 & 8 & 47 \\ 0 & 0 & -20 & -40 & -260 \\ 0 & 0 & 0 & 0 & 0\end{array}\right) \\ \left(\begin{array}{cccc|c}1 & 0 & 5 & 11 & 68 \\ 0 & 1 & 3 & 8 & 47 \\ 0 & 0 & 1 & 2 & 13 \\ 0 & 0 & 0 & 0 & 0\end{array}\right) &\rightsquigarrow \left(\begin{array}{cccc|c}1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 & 8 \\ 0 & 0 & 1 & 2 & 13 \\ 0 & 0 & 0 & 0 & 0\end{array}\right)\end{aligned}$$

Com que el rang de la matriu associada és 3 i el de l'ampliada també, el sistema és compatible. Com que tenim 4 incògnites, té $4-3=1$ paràmetre lliures (per tant, sistema compatible indeterminat). Amb l'esglaonament que hem fet, la t és el paràmetre lliure i podem escriure la solució:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \\ 13 \end{pmatrix} - t \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \text{ amb } t \in \mathbb{K}.$$

Exemple 2.40. Per tal d'aprofitar els càlculs del sistema anterior, tant sols fem una petita modificació al l'últim terme independent:

$$\begin{aligned}x - y + 2z + 3t &= 21 \\ -x + 2y + z + 5t &= 26 \\ 3x + y - 2z + t &= -9 \\ 3x + 2y + z + 9t &= 39\end{aligned}$$

Considerem la matriu ampliada i esglaonem (són els mateixos passos d'abans, pel que tant sols escrivim la primera i última matriu):

$$\left(\begin{array}{cccc|c}1 & -1 & 2 & 3 & 21 \\ -1 & 2 & 1 & 5 & 26 \\ 3 & 1 & -2 & 1 & -9 \\ 3 & 2 & 1 & 9 & 39\end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c}1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 2 & 8 \\ 0 & 0 & 1 & 2 & 13 \\ 0 & 0 & 0 & 0 & 1\end{array}\right)$$

si som estrictes en el concepte de reduïda, falta utilitzar l'1 de l'última fila per a posar zeros a l'última columna. En qualsevol cas, veiem que la matriu ampliada té rang 4, mentre que l'associada té rang 3, pel que el sistema és incompatible.

Considerem ara el cas particular en que $B = \mathbf{0}_m$ (un vector format per zeros).

Definició 2.41. (a) Diem que el sistema d'equacions lineals $AX = B$ és *homogeni* si $B = \mathbf{0}_m$.

(b) Si $AX = B$ és un sistema, parlem de *sistema homogeni associat* al sistema $AX = \mathbf{0}_m$.

Tenim els resultats següents:

- (a) Si el sistema és homogeni, llavors $X = \mathbf{0}_n$ és una solució, per tant el sistema és compatible.
- (b) Si X és solució del sistema homogeni i $\lambda \in \mathbb{K}$, llavors λX també és solució del sistema: si ho pensem com a multiplicació de matrius, tenim la igualtat $A(\lambda X) = \lambda(AX) = \lambda \mathbf{0}_m = \mathbf{0}_m$, per tant λX també és solució.
- (c) Si X i Y són solucions d'un sistema homogeni, llavors $X + Y$ també és solució: tornem a escriure-ho en forma matricial: $A(X + Y) = AX + AY = \mathbf{0}_m + \mathbf{0}_m = \mathbf{0}_m$.
- (d) Si $AX = B$ és un sistema, amb X i Y una solucions, llavors $X - Y$ és una solució del sistema homogeni associat: $A(X - Y) = AX - AY = B - B = \mathbf{0}_m$.
- (e) Si $AX = B$ és un sistema i X una solució particular, qualsevol altre solució Y es pot escriure com $Y = X + Z$, amb Z solució del sistema homogeni associat: això es dedueix de l'apartat anterior: si X i Y són solucions, llavors $Z = Y - X$ és solució de l'homogeni. Si X és solució del sistema i Z de l'homogeni associat, llavors $AY = A(X + Z) = AX + AZ = B + \mathbf{0}_m = B$.

Exemple 2.42. Considerem el sistema d'equacions de l'Exemple 2.39:

$$\begin{aligned} x + y + 2z + 3t &= 21 \\ -x + 2y + z + 5t &= 26 \\ 3x + y - 2z + t &= -9 \\ 3x + 2y + z + 9t &= 38 \end{aligned}$$

Veiem que el sistema homogeni associat és

$$\begin{aligned} x + y + 2z + 3t &= 0 \\ -x + 2y + z + 5t &= 0 \\ 3x + y - 2z + t &= 0 \\ 3x + 2y + z + 9t &= 0 \end{aligned}$$

i té per solució

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \text{ amb } t \in \mathbb{K}.$$

Tenint en compte que una solució particular és $(x, y, z) = (3, 8, 13)$, podem escriure:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \\ 13 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \text{ amb } t \in \mathbb{K}.$$

Exercicis recomanats

Els exercicis que segueixen són útils per practicar el material presentat. La numeració és la de [1].

Secció 1.1: 6, 12, 16.

Secció 1.2: 10, 18.

Secció 1.3: 6, 8, 18, 20, 24, 28.

3 Espais vectorials i aplicacions lineals

El contingut d'aquesta secció també es pot trobar a [1, Secció 2.2] i [2, Temes 2, 3, 4].

3.1 Matrius com a aplicacions lineals

En aquest capítol considerarem funcions de l'espai \mathbb{K}^n a l'espai \mathbb{K}^m . Recordem que aquests espais els hem definit com $\mathbb{K}^n = M_{n \times 1}(\mathbb{K})$ (i anàlogament per \mathbb{K}^m), i els seus elements els anomenem *vectors*. Així, una funció f assigna a cada vector $\vec{x} \in \mathbb{K}^n$ un vector imatge $f(\vec{x}) \in \mathbb{K}^m$. Un primer exemple ens els donen les matrius:

Exemple 3.1. Considerem una matriu $A \in M_{m \times n}(\mathbb{K})$. Aleshores l'aplicació associada a A és f_A , definida com:

$$f_A(\vec{x}) = A\vec{x}.$$

Tal i com hem remarcat anteriorment, l'aplicació f_A assigna a cada vector \vec{x} una combinació lineal de les columnes d' A , on els coeficients venen donats per \vec{x} .

De les propietats del producte de matrius, sabem que f_A satisfà:

1. $f_A(\vec{x} + \vec{y}) = f_A(\vec{x}) + f_A(\vec{y})$ per a tot $\vec{x}, \vec{y} \in \mathbb{K}^n$, i
2. $f_A(\lambda\vec{x}) = \lambda f_A(\vec{x})$ per a tot $\lambda \in \mathbb{K}$ i tot $\vec{x} \in \mathbb{K}^n$.

Podem prendre l'aplicació f_A com a model de les aplicacions que volem considerar en aquest curs:

Definició 3.2. Diem que una aplicació $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ és *lineal* si satisfà

1. $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$ per a tot $\vec{x}, \vec{y} \in \mathbb{K}^n$, i
2. $f(\lambda\vec{x}) = \lambda f(\vec{x})$ per a tot $\lambda \in \mathbb{K}$ i tot $\vec{x} \in \mathbb{K}^n$.

Proposició 3.3. Una aplicació $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ és lineal si i només si existeix una matriu $A \in M_{m \times n}(\mathbb{K})$ tal que $f = f_A$.

Demostració. Ja hem vist que si $f = f_A$ aleshores f és lineal. Recíprocament, suposem que f és lineal. Aleshores, com que

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

el valor $f(\vec{x})$ es pot calcular com

$$f \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 f \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 f \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n f \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Definim la matriu A com

$$A = \begin{pmatrix} | & | & | & | \\ f(\vec{e}_1) & f(\vec{e}_2) & \cdots & f(\vec{e}_n) \\ | & | & | & | \end{pmatrix}$$

on escrivim e_i com el vector format per zeros excepte en la posició i , on hi ha un 1. Aleshores, per la definició de producte d'una matriu per un vector, tenim $f(\vec{x}) = A\vec{x}$ per a tot $\vec{x} \in \mathbb{K}^n$, com volíem veure. \square

És important remarcar com hem construït la matriu associada a una aplicació lineal f en la demostració anterior: donada $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$, la seva matriu associada és

$$A = \begin{pmatrix} | & | & | & | \\ f(\vec{e}_1) & f(\vec{e}_2) & \cdots & f(\vec{e}_n) \\ | & | & | & | \end{pmatrix}.$$

Els vectors \vec{e}_i sovint s'anomenen *vectors estàndard*:

$$\vec{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ (posició } i \text{)}$$

Com a primera aplicació, considerem ara dues aplicacions lineals

$$\mathbb{K}^n \xrightarrow{g} \mathbb{K}^r \xrightarrow{f} \mathbb{K}^m,$$

i suposem que f té matriu A i g té matriu B . Considerem la composició $f \circ g$, que és l'aplicació que envia $\vec{x} \in \mathbb{K}^n$ a $(f \circ g)(\vec{x}) = f(g(\vec{x}))$. Per definició, sabem que per a tot \vec{x} tenim

$$(f \circ g)(\vec{x}) = f(g(\vec{x})) = A(B\vec{x}) = (AB)\vec{x}.$$

Per tant, obtenim que $f \circ g$ és una aplicació lineal, amb matriu associada AB . Per tant, el producte de matrius “codifica” la composició d'aplicacions lineals.

Exemple 3.4. Donades tres aplicacions lineals

$$\mathbb{K}^t \xrightarrow{h} \mathbb{K}^n \xrightarrow{g} \mathbb{K}^r \xrightarrow{f} \mathbb{K}^m$$

amb matrius respectives $f = f_A$, $g = f_B$, $h = f_C$, el fet que la composició d'aplicacions és associativa (per definició) implica que el producte de matrius és associatiu, sense necessitat de fer cap càlcul: la matriu de la composició $f \circ g \circ h$ és $A(BC)$ i $(AB)C$ a la vegada, i per tant han de coincidir.

3.2 Aplicacions lineals i geometria

En aquesta § veurem algunes aplicacions lineals del pla \mathbb{R}^2 i aprendrem a reconèixer com actuen a partir de la forma de la seva matriu. Podem visualitzar aquestes transformacions per exemple a <https://shadanan.github.io/MatVis/>.

3.2.1 Homotècies

Considerem una matriu

$$\begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}, k \in \mathbb{R}.$$

Aquesta matriu envia el vector \vec{x} a $k\vec{x}$. Per tant, multiplica els vectors per un escalar. Si $|k| > 1$, es tracta d'una dilatació i, si $|k| < 1$ es tracta d'una contracció. Si k és negatiu, aleshores també canvia el signe (es fa primer una simetria respecte l'origen). En general, aquestes transformacions s'anomenen homotècies.

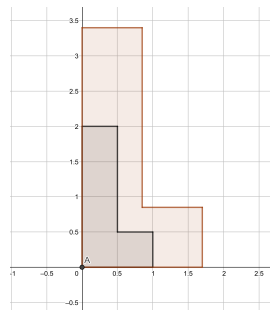


Figura 1: Una homotècia de factor 1.7.

3.2.2 Projections ortogonals

Considerem la matriu $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Aquesta matriu envia el vector $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ a $\begin{pmatrix} x \\ 0 \end{pmatrix}$. És a dir, projecta el vector sobre l'eix de les x .

Més en general, considerem una recta ℓ del pla, que talli l'origen. Qualsevol vector \vec{x} es pot escriure de manera única com $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$, on el vector \vec{x}^{\parallel} és sobre la recta ℓ , i \vec{x}^{\perp} és perpendicular a ℓ .

Definició 3.5. La *projecció ortogonal* respecte la recta ℓ és l'aplicació proj_{ℓ} definida com

$$\text{proj}_{\ell}(\vec{x}) = \vec{x}^{\parallel}.$$

Observem que proj_{ℓ} és una aplicació lineal (exercici).

Proposició 3.6. Sigui \vec{v} un vector paral·lel a la recta ℓ . Aleshores:

$$\text{proj}_{\ell}(\vec{x}) = \frac{\vec{x} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \vec{v}.$$

També podem trobar una fórmula per la matriu de la projecció ortogonal.

Proposició 3.7. Sigui $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ un vector paral·lel a la recta ℓ . Aleshores la matriu de proj_{ℓ} és

$$\frac{1}{v_1^2 + v_2^2} \begin{pmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{pmatrix}$$

3.2.3 Reflexions

Considerem com abans una recta ℓ que passi per l'origen. Volem definir la reflexió respecte la recta ℓ . És fàcil veure que

$$\text{ref}_\ell(\vec{x}) = \vec{x}^\parallel - \vec{x}^\perp.$$

Com que $\vec{x} = \vec{x}^\parallel + \vec{x}^\perp$, podem escriure

$$\vec{x}^\parallel - \vec{x}^\perp = 2\vec{x}^\parallel - \vec{x}.$$

Per tant,

$$\text{ref}_\ell = 2\text{proj}_\ell - \text{id}.$$

La matriu corresponent és fàcil de trobar a partir d'aquesta fórmula:

Proposició 3.8. *Si $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ un vector paral·lel a la recta ℓ . Aleshores la matriu de ref_ℓ és*

$$\frac{1}{v_1^2 + v_2^2} \begin{pmatrix} v_1^2 - v_2^2 & 2v_1v_2 \\ 2v_1v_2 & v_2^2 - v_1^2 \end{pmatrix}.$$

Fixem-nos que la matriu de ref_ℓ té la forma $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$, i que a més les entrades satisfan $a^2 + b^2 = 1$. De fet, es pot veure que qualsevol matriu d'aquesta forma és la matriu d'una reflexió.

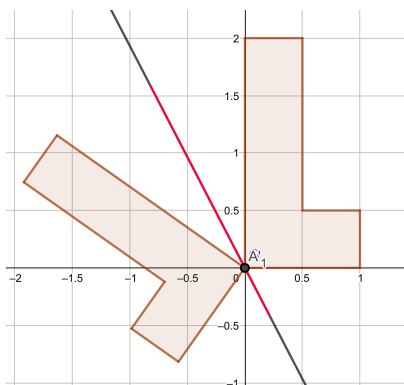


Figura 2: Una reflexió respecte una recta. També hi podem observar la projecció a la mateixa recta.

3.2.4 Rotacions

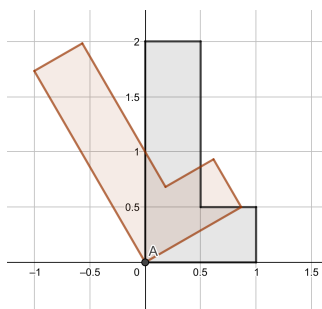
Les matrius de la forma

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

corresponen a una rotació d'angle θ . Observem que aquesta matriu és de la forma

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a^2 + b^2 = 1$$

i totes les matrius d'aquesta forma són rotacions.

Figura 3: Una rotació de 30° .

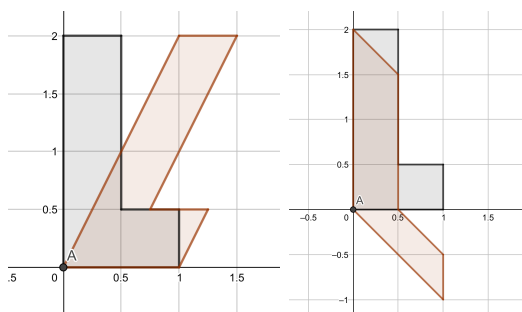
Exercici 3.1. Considereu la matriu de rotació d'angle θ_1 i la de rotació d'angle θ_2 . Descriviu la composició d'aquestes dues aplicacions i, fent servir el producte de matrius, redescobriu les fórmules per la duplicació d'angles que heu vist a secundària.

Proposició 3.9. Una matriu de la forma $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ és la composició d'una homotècia de magnitud $k = \sqrt{a^2 + b^2}$ amb una rotació d'angle $\arctan(b/a)$ (si a és positiu) o $\arctan(b/a) + \pi$ (si a és negatiu).

3.2.5 Lliscaments

Per últim estudiarem les matrius de la forma $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, on $k \in \mathbb{R}$. Aquestes matrius fixen el vector \vec{e}_1 , mentre envien el vector \vec{e}_2 a $\begin{pmatrix} k \\ 1 \end{pmatrix}$. Geomètricament, enviarien un quadrat en el pla a un rombe, on la base es manté fixa però els costats verticals s'esbiaixen. Direm que és un lliscament horitzontal.

De manera anàloga, la matriu $\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}$ representa un lliscament vertical.

Figura 4: Lliscament horitzontal de valor $\frac{1}{2}$ i lliscament vertical de valor -1 .

Exercici 3.2. Calculeu, en els casos que es pugui, les inverses de les matrius corresponents a homotècies, projeccions ortogonals, reflexions i lliscaments.

3.3 Subespais, generadors i bases

Tornem ara a considerar \mathbb{K}^n i les aplicacions lineals $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$.

Definició 3.10. Diem que un subconjunt $W \subset \mathbb{K}^n$ és un subespai vectorial si W compleix que:

- $\vec{0} \in W$,

- $\vec{u} + \vec{v} \in W$ per a tot \vec{u} i \vec{v} de W , i
- $\lambda\vec{u} \in W$ per a tot $\vec{u} \in W$ i $\lambda \in \mathbb{K}$.

Exemple 3.11. • A l' \mathbb{R} -espai vectorial \mathbb{R}^3 hi considerem $F = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$. Podem comprovar que és un subespai vectorial.

- Més en general, si tenim un sistema d'equacions lineals homogeni amb n incògnites, podem identificar una solució (x_1, \dots, x_n) (l'hem escrit com una fila per comoditat) amb un vector de \mathbb{K}^n i les solucions formen un subespai vectorial de \mathbb{K}^n .

Com a cas particular de subespai vectorial tenim el següent:

Definició 3.12. Si $\{\vec{v}_1, \dots, \vec{v}_m\} \subset \mathbb{K}^n$, definim el *subespai generat per* $\{\vec{v}_1, \dots, \vec{v}_m\}$ com els vectors de \mathbb{K}^n que es poden escriure com a combinació lineal de $\{\vec{v}_1, \dots, \vec{v}_m\}$. Denotem aquest subconjunt com $\langle \vec{v}_1, \dots, \vec{v}_m \rangle$, i per tant tenim:

$$\langle \vec{v}_1, \dots, \vec{v}_m \rangle = \{\lambda_1 \vec{v}_1 + \dots + \lambda_m \vec{v}_m \mid \lambda_i \in \mathbb{K}\}$$

Observació 3.13. Si $\{\vec{v}_1, \dots, \vec{v}_m\} \subset \mathbb{K}^n$, llavors $\langle \vec{v}_1, \dots, \vec{v}_m \rangle$ és un subespai vectorial.

Demostració. Per simplificar la notació, definim

$$W := \langle \vec{v}_1, \dots, \vec{v}_m \rangle.$$

Cal veure que si $\vec{u}, \vec{w} \in W$, llavors $\vec{u} + \vec{w} \in W$: per definició d' W , tenim que existeixen escalar $\lambda_1, \dots, \lambda_m$ i μ_1, \dots, μ_m tals que $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_m \vec{v}_m$ i $\vec{w} = \mu_1 \vec{v}_1 + \dots + \mu_m \vec{v}_m$. Llavors, tenim que:

$$\vec{u} + \vec{w} = (\lambda_1 + \mu_1) \vec{v}_1 + \dots + (\lambda_m + \mu_m) \vec{v}_m,$$

obtenint que $\vec{u} + \vec{w} \in W$.

Ara considerem $\lambda \in \mathbb{K}$ i $\vec{u} \in W$. Volem veure $\lambda\vec{u} \in W$: com que $\vec{u} \in W$, existeixen $\mu_1, \dots, \mu_m \in \mathbb{K}$ tals que $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_m \vec{v}_m$. Llavors:

$$\lambda\vec{u} = (\lambda\mu_1) \vec{v}_1 + \dots + (\lambda\mu_m) \vec{v}_m,$$

pel que $\lambda\vec{u} \in W$. □

Exemple 3.14. Considerem \mathbb{R}^3 com a \mathbb{R} -espai vectorial i els vectors (els escrivim per files) $\{(1, 0, 0), (0, 1, 0)\}$, tenim que $\langle (1, 0, 0), (0, 1, 0) \rangle$ són els vectors d' \mathbb{R}^3 amb tercera coordenada zero.

Definició 3.15. Sigui $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ una aplicació lineal.

- Definim el *nucli d' f* (i el denotem com $\text{Ker}(f)$) com el conjunt format pels vectors $\vec{v} \in \mathbb{K}^n$ tals que $f(\vec{v}) = \vec{0}$.
- Definim la *imatge d' f* (i la denotem com $\text{Im}(f)$) com el conjunt dels vectors $\vec{w} \in \mathbb{K}^m$ tals que existeix $\vec{v} \in \mathbb{K}^n$ tal que $f(\vec{v}) = \vec{w}$.

Proposició 3.16. Si $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ és una aplicació lineal, llavors:

- $\text{Ker}(f)$ és un subespai vectorial de \mathbb{K}^n .
- $\text{Im}(f)$ és un subespai vectorial de \mathbb{K}^m .

Demostració. Comprovem primer que $\text{Ker}(f)$ compleix les dues condicions de la Definició 3.10: Considerem $\vec{u}, \vec{v} \in \text{Ker}(f)$, per tant $f(\vec{u}) = f(\vec{v}) = \vec{0}$. Llavors:

$$f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}) = \vec{0} + \vec{0} = \vec{0},$$

per tant $\vec{u} + \vec{v} \in \text{Ker}(f)$.

També, si $\vec{u} \in \text{Ker}(f)$ i $\lambda \in \mathbb{K}$, llavors $f(\lambda\vec{u}) = \lambda f(\vec{u}) = \lambda\vec{0} = \vec{0}$, per tant $\lambda\vec{u} \in \text{Ker}(f)$.

Comprovem ara que $\text{Im}(f)$ també compleix les dues condicions: si $\vec{v}_1, \vec{v}_2 \in \text{Im}(f)$, llavors existeixen $\vec{u}_1, \vec{u}_2 \in \mathbb{K}^n$ tals que $f(\vec{u}_1) = \vec{v}_1$ i $f(\vec{u}_2) = \vec{v}_2$. Llavors es compleix que $f(\vec{u}_1 + \vec{u}_2) = f(\vec{u}_1) + f(\vec{u}_2) = \vec{v}_1 + \vec{v}_2$, pel que $\vec{v}_1 + \vec{v}_2 \in \text{Im}(f)$. També, si $\vec{v} \in \text{Im}(f)$ (per tant existeix $\vec{u} \in \mathbb{K}^n$ tal que $f(\vec{u}) = \vec{v}$) i $\lambda \in \mathbb{K}$, tenim que $f(\lambda\vec{u}) = \lambda f(\vec{u}) = \lambda\vec{v}$, pel que $\lambda\vec{v} \in \text{Im}(f)$. \square

Com a cas particular podem considerar les aplicacions $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ definides a la Subsecció 3.1 mitjançant una matriu $A \in M_{m \times n}(\mathbb{K})$. En aquest cas tenim que:

- $\text{Ker}(f_A)$ són les solucions del sistema homogeni que té per matriu associada A .
- $\text{Im}(f_A)$ és el subespai de \mathbb{K}^m generat per les columnes d' A pensades com a vectors de \mathbb{K}^m .

Exemple 3.17. Considerem $W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y - z = 0 \text{ i } x - y + 2z = 0\}$. Podem pensar que W és el nucli de f_A , on $A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}$.

Tot i que ja hi ha una definició d'independència lineal a la Definició 2.7, vegem-ne una d'equivalent:

Definició 3.18. Diem que $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$, són una família de vectors *linealment independents* si l'única manera d'escriure el vector zero com a combinació lineal de $\vec{v}_1, \dots, \vec{v}_n$ és amb tots els coeficients zero. O sigui, si:

$$\vec{0} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n \text{ amb } \lambda_i \in \mathbb{K} \implies \lambda_1 = \dots = \lambda_n = 0.$$

Observació 3.19. Les definicions 2.7 i 3.18 són equivalents.

Demostració. En realitat, veurem que negar una de les definicions és si i només si es nega l'altra:

Considerem $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ que no compleixen la definició de independència lineal de la Definició 2.7, llavors un dels vectors es pot escriure com a combinació lineal dels altres. Com que els podem reordenar, considerem que $\vec{v}_1 = \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n$. Llavors

$$\vec{0} = \vec{v}_1 - \lambda_2 \vec{v}_2 - \dots - \lambda_n \vec{v}_n$$

i el vector zero es pot escriure com a combinació lineal de $\{\vec{v}_1, \dots, \vec{v}_n\}$ amb, com a mínim, un coeficient no nul (el de \vec{v}_1).

Considerem $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ que no compleixen la definició de independència lineal de la Definició 3.18, llavors podem escriure el vector zero:

$$\vec{0} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n$$

amb algun $\lambda_i \neq 0$. Com que els podem reordenar, considerem que $\lambda_1 \neq 0$. Llavors:

$$\vec{v}_1 = -(\lambda_2/\lambda_1)\vec{v}_2 - \dots - (\lambda_n/\lambda_1)\vec{v}_n$$

i el vector \vec{v}_1 és combinació lineal de la resta. \square

Observació 3.20. Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$, és una família de vectors linealment independents, llavors qualsevol subfamília d'aquesta també ho serà. Per tant, treure vectors preserva la propietat de ser linealment independents.

Podem veure si una família de vectors és linealment independent mirant el nucli d'una aplicació lineal:

Proposició 3.21. *Siguin $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ una família de vectors i $A \in M_{m \times n}(\mathbb{K})$ la matriu que té per columnes els vectors \vec{v}_j . Llavors $\{\vec{v}_1, \dots, \vec{v}_n\}$ són linealment independents si, i només si, $\text{Ker}(f_A) = \{\vec{0}\}$.*

Demostració. Observem que donar una combinació lineal $x_1\vec{v}_1 + \dots + x_n\vec{v}_n$ és equivalent a multiplicar AX , on X és el vector columna que té per coeficients x_1, \dots, x_n .

llavors, que $\{\vec{v}_1, \dots, \vec{v}_n\}$ siguin linealment independents vol dir que la única manera d'obtenir $AX = \vec{0}$ és amb $X = \vec{0}$, o, el que és el mateix, $\text{Ker}(f_A) = \vec{0}$. \square

Podem veure quants vectors linealment independents hi ha dins d'una família de vectors mirant el rang de la matriu corresponent:

Lema 3.22. *Si $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{K}^m$ són vectors, llavors el nombre màxim de vectors linealment independents continguts a $\{\vec{v}_1, \dots, \vec{v}_n\}$ és el rang de la matriu A que té per columnes els vectors \vec{v}_i .*

Demostració. Anomenem k al nombre màxim de vectors linealment independents a $\{\vec{v}_1, \dots, \vec{v}_n\}$. Com que l'ordre no és important, reordenem els vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ tal que els k primers són linealment independents i que qualsevol altre \vec{v}_j amb $j \geq k$, $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}_j\}$ no són linealment independents. Sigui B la matriu que té per columnes els vectors $\vec{v}_1, \dots, \vec{v}_k$ i \overline{B}_j la que té per columnes els vectors $\vec{v}_1, \dots, \vec{v}_k, \vec{v}_j$. Pensem \overline{B}_j com la matriu ampliada d'un sistema d'equacions, amb B matriu associada. Com que \vec{v}_j és combinació lineal dels $\vec{v}_1, \dots, \vec{v}_k$, el sistema té solució i $\text{Rang}(B) = \text{Rang}(\overline{B}_j)$.

Podem iterar aquest procediment amb els altres vectors \vec{v}_j amb $j \geq k$ i obtenim que $\text{Rang}(B) = \text{Rang}(A)$.

Finalment, com que B té per columnes vectors linealment independents, el sistema d'equacions homogeni que té per matriu associada B , té solució única, i per tant $\text{Rang}(B) = k$ (el número d'incògnites). \square

D'aquí podem deduir una propietat que relaciona el producte de matrius amb el rang:

Corol·lari 3.23. *Si $A \in M_{m \times n}(\mathbb{K})$ i $B \in M_{n \times r}(\mathbb{K})$, llavors:*

$$\text{Rang}(AB) \leq \min\{\text{Rang}(A), \text{Rang}(B)\}$$

Demostració. Les columnes d' AB es poden pensar com a combinació lineal de columnes d' A , per tant no n'hi pot haver més de linealment independents, obtenint $\text{Rang}(AB) \leq \text{Rang}(A)$. Si ara transposem les matrius i fem el mateix raonament, tenim que

$$\text{Rang}(AB) = \text{Rang}((AB)^T) = \text{Rang}(B^T A^T) \leq \text{Rang}(B^T) = \text{Rang}(B)$$

Per tant $\text{Rang}(AB)$ és més petit que el rang d' A i que el rang de B . \square

Corol·lari 3.24. *Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ són linealment independents, llavors $n \leq m$.*

Demostració. Si $\{\vec{v}_1, \dots, \vec{v}_n\}$ són linealment independents, pel Lema 3.22, la matriu V , una matriu que té per columnes els vectors \vec{v}_j té rang n . Això només es pot tenir si hi ha com a mínim n files, i per tant $m \geq n$. \square

Vegem ara el concepte de *generadors*:

Definició 3.25. Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ són tals que $\mathbb{K}^m = \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, diem que $\{\vec{v}_1, \dots, \vec{v}_n\}$ són un sistema de generadors de \mathbb{K}^m .

Si $W \subset \mathbb{K}^m$ és un subespai vectorial i $\{\vec{v}_1, \dots, \vec{v}_n\} \subset W$ són tals que $W = \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, diem que $\{\vec{v}_1, \dots, \vec{v}_n\}$ són un sistema de generadors de W .

Observació 3.26. Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ és un sistema de generadors i hi afegim un vector \vec{u} , llavors $\{\vec{u}, \vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ és també un sistema de generadors. O sigui, afegir vectors conserva la propietat de ser sistema de generadors.

Observació 3.27. Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$, llavors $\{\vec{v}_1, \dots, \vec{v}_n\}$ són un sistema de generadors de $\langle \vec{v}_1, \dots, \vec{v}_n \rangle$ (que és un subespai vectorial).

Vegem algunes propietats dels sistemes de generadors:

Lema 3.28. Si $\vec{v}_1, \dots, \vec{v}_n, \vec{u}$ són vectors de \mathbb{K}^m , llavors:

- (a) $\vec{u} \in \langle \vec{v}_1, \dots, \vec{v}_n \rangle$ si i només si $\langle \vec{v}_1, \dots, \vec{v}_n \rangle = \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_n \rangle$.
- (b) Si $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$ i i és un índex tal que $\lambda_i \neq 0$, llavors:

$$\langle \vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_i, \vec{v}_{i+1}, \dots, \vec{v}_n \rangle = \langle \vec{v}_1, \dots, \vec{v}_{i-1}, \vec{u}, \vec{v}_{i+1}, \dots, \vec{v}_n \rangle.$$

Demostració. Comencem demostrant (a):

Suposem $\vec{u} \in \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, per tant existeixen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tals que $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$. Sempre tenim $\langle \vec{v}_1, \dots, \vec{v}_n \rangle \subset \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_n \rangle$, per tant cal comprovar l'altre inclusió. Sigui $\vec{w} \in \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_n \rangle$, per tant existeixen $\mu, \mu_1, \dots, \mu_n \in \mathbb{K}$ tals que $\vec{w} = \mu \vec{u} + \mu_1 \vec{v}_1 + \dots + \mu_n \vec{v}_n$. Utilitzant que $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$, tenim:

$$\vec{w} = (\mu \lambda_1 + \mu_1) \vec{v}_1 + \dots + (\mu \lambda_n + \mu_n) \vec{v}_n,$$

per tant $\vec{w} \in \langle \vec{v}_1, \dots, \vec{v}_n \rangle$.

En canvi, si $\vec{u} \notin \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, com que sempre tenim $\vec{u} \in \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_n \rangle$, obtenim que $\langle \vec{v}_1, \dots, \vec{v}_n \rangle \neq \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_n \rangle$.

Per demostrar (b) comencem observant que, com que $\lambda_i \neq 0$,

$$\vec{v}_i = (1/\lambda_i) \vec{u} - (\lambda_1/\lambda_i) \vec{v}_1 - \dots - (\lambda_{i-1}/\lambda_i) \vec{v}_{i-1} - (\lambda_{i+1}/\lambda_i) \vec{v}_{i+1} - \dots - (\lambda_n/\lambda_i) \vec{v}_n$$

pel que $\vec{v}_i \in \langle \vec{v}_1, \dots, \vec{v}_{i-1}, \vec{u}, \vec{v}_{i+1}, \dots, \vec{v}_n \rangle$. Apliquem ara dos cops (a) per obtenir (l'ordre en que s'escriuen els vectors en un sistema de generadors no importa):

$$\langle \vec{v}_1, \dots, \vec{v}_n \rangle = \langle \vec{u}, \vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_n \rangle = \langle \vec{v}_1, \dots, \vec{v}_{i-1}, \vec{u}, \vec{v}_{i+1}, \dots, \vec{v}_n \rangle.$$

□

Vegem ara quin és el nombre mínim de vectors generadors de \mathbb{K}^m :

Lema 3.29. Si $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ són un sistema de generadors, llavors $n \geq m$.

Demostració. Considerem els vectors estàndard $\vec{e}_j \in \mathbb{K}^m$ definits a la Subsecció 3.1. Com que $\langle \vec{v}_1, \dots, \vec{v}_n \rangle = \mathbb{K}^m$, tenim que $\vec{e}_j \in \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, per tant, existeixen x_{1j}, \dots, x_{nj} tals que $x_{1j} \vec{v}_1 + \dots + x_{nj} \vec{v}_n = \vec{e}_j$. Podem escriure aquesta igualtat de forma matricial:

$$AX = \mathbf{1}_m$$

on $A \in M_{m \times n}(\mathbb{K})$ és la matriu que té per columnes els vectors \vec{v}_j , $X \in M_{n \times m}(\mathbb{K})$ és la matriu que té per coeficients x_{ij} i $\mathbf{1}_m$ és la matriu que té per columnes els vectors \vec{e}_j (que és la identitat).

Ara utilitzem el rang: $m = \text{Rang}(AX) \leq \text{Rang}(A)$, per tant A té com a mínim m columnes i $n \geq m$. □

Definició 3.30. Diem que una família ordenada de vectors $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ amb $\vec{v}_j \in \mathbb{K}^m$, és una base de \mathbb{K}^m si $\{\vec{v}_1, \dots, \vec{v}_n\}$, són linealment independents i generadors de \mathbb{K}^m .

Observació 3.31. L'ordre dels vectors d'una base és important: de la definició es dedueix que si $[\vec{v}_1, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_n]$ és una base de \mathbb{K}^m , llavors $[\vec{v}_2, \vec{v}_1, \vec{v}_3, \dots, \vec{v}_n]$ també ho és, però considerarem que són bases diferents.

Teorema 3.32. *Si $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ una base de \mathbb{K}^m , llavors $m = n$. Diem que \mathbb{K}^m té dimensió m .*

Demostració. Com que han de ser linealment independents, del Corol·lari 3.24, obtenim que $n \leq m$. Com que han de ser generadors, del Lema 3.29, obtenim que $n \geq m$, per tant, $m = n$. \square

Exemple 3.33. • Els vectors (ho escrivim per files) $(1, 0)$ i $(0, 1)$ formen una base de \mathbb{R}^2 .

- Els vectors $(1, 0)$ i $(1, 1)$ també formen una base de \mathbb{R}^2 .
- Els vectors $(1, 0)$, $(0, 1)$ i $(1, 1)$ no formen una base de \mathbb{R}^2 (no són linealment independents).

Proposició 3.34. (a) *Si $\langle \vec{v}_1, \dots, \vec{v}_n \rangle = \mathbb{K}^m$, qualsevol subfamília maximal de vectors linealment independents de $\vec{v}_1, \dots, \vec{v}_n$ formaran una base de \mathbb{K}^m .*

(b) *Si $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{K}^m$ és una família de vectors linealment independents, podem ampliar la família fins a obtenir una base de \mathbb{K}^m .*

Demostració. Demostrem primer (a): en aquest cas, maximal vol dir que si afegim qualsevol vector, la família obtinguda deixa de ser linealment independent. Suposem que tenim una família maximal de vectors linealment independents $\vec{v}_{i_1}, \dots, \vec{v}_{i_k}$. Per demostrar que són base cal veure que són generadors i ho farem per contraposició lògica. Si no fossin generadors, tindríem:

$$\langle \vec{v}_{i_1}, \dots, \vec{v}_{i_k} \rangle \subsetneq \langle \vec{v}_1, \dots, \vec{v}_n \rangle$$

i per tant hi hauria un \vec{v}_j tal que $\vec{v}_j \notin \langle \vec{v}_{i_1}, \dots, \vec{v}_{i_k} \rangle$. Llavors $\vec{v}_{i_1}, \dots, \vec{v}_{i_k}, \vec{v}_j$ serien linealment independents (si $\lambda_1 \vec{v}_{i_1} + \dots + \lambda_k \vec{v}_{i_k} + \lambda \vec{v}_j = \vec{0}$, pot passar que $\lambda = 0$ o $\lambda \neq 0$; si $\lambda = 0$, com que $\vec{v}_{i_1}, \dots, \vec{v}_{i_k}$ són linealment independents, tots els $\lambda_i = 0$; si $\lambda \neq 0$, podríem aïllar \vec{v}_j en funció dels altres, contradient que $\vec{v}_j \notin \langle \vec{v}_{i_1}, \dots, \vec{v}_{i_k} \rangle$). Per tant, tindríem que $\vec{v}_{i_1}, \dots, \vec{v}_{i_k}$ no seria maximal, tenint una contradicció.

Per demostrar (b), considerem la família $\vec{v}_1, \dots, \vec{v}_n, \vec{e}_1, \dots, \vec{e}_m$ (on \vec{e}_i són els vectors estàndard) i en triem una subfamília maximal que contingui els primers k vectors. \square

El resultat de l'apartat (a) també es pot llegir com que si tenim un sistema de generadors de \mathbb{K}^n , podem anar eliminant vectors fins a quedar-nos amb una base.

Proposició 3.35. *Si $\vec{v}_1, \dots, \vec{v}_n$ és una família de vectors de \mathbb{K}^n , llavors són equivalents:*

- (a) $\vec{v}_1, \dots, \vec{v}_n$ formen una base.
- (b) $\vec{v}_1, \dots, \vec{v}_n$ són un sistema de generadors.
- (c) $\vec{v}_1, \dots, \vec{v}_n$ són linealment independents.

Demostració. Tenim que per definició de base (a) \Rightarrow (b) i (a) \Rightarrow (c).

Suposem que tenim (b). Si no fossin linealment independents, hi hauria una subfamília de vectors linealment independents $\vec{v}_{i_1}, \dots, \vec{v}_{i_k}$ maximals amb $k < n$, per tant tindríem una base

amb k elements en un espai de dimensió $n > k$, per tant contradicció. Per tant els vectors són linealment independents (i generadors), per tant tenim (a).

Suposem que tenim (c). Si $\vec{v}_1, \dots, \vec{v}_n$ no fossin generadors, podríem afegir vectors fins tenir una base, però aquesta nova base tindria més de n vectors, pel que contradiu que l'espai \mathbb{K}^n sigui de dimensió n . Per tant $\vec{v}_1, \dots, \vec{v}_n$ són generadors (i linealment independents), per tant tenim (a). \square

Encara podem donar una altra caracterització del concepte de base.

Teorema 3.36. *Una família de vectors $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ formen una base si i només si tot per a tot vector $\vec{v} \in \mathbb{K}^m$ existeixen uns únics $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tals que $\vec{v} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$.*

Demostració. Suposem $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$, llavors són generadors de \mathbb{K}^m si i només si per a tot vector $\vec{v} \in \mathbb{K}^m$ existeixen $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tals que $\vec{v} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$, per tant si i només si el sistema té solució.

Suposem $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$, llavors són linealment independents si i només si existeixen $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \mathbb{K}$ tals que $\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = \mu_1 \vec{v}_1 + \dots + \mu_n \vec{v}_n$ implica que $\lambda_j = \mu_j$ per a tot j , per tant, si i només si la solució és única. \square

Observació 3.37. Equivalentment al Teorema 3.36, podem dir que $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathbb{K}^m$ formen una base si i només si la matriu $A \in M_{m \times n}(\mathbb{K})$ formada pels vectors \vec{v}_i a les columnes és invertible (en particular, ha de ser quadrada i per tant $n = m$).

Demostració. Si són base, necessitem que el sistema $Ax = B$ tingui solució única per a tot $B \in M_{n \times 1}$. Anem considerant $B_j = \vec{e}_j$ (els vectors estàndard definits a la Secció 3.1) i resolent cada un dels sistemes amb solucions $X_j \in M_{n \times 1}$, tindrem una matriu X que compleix $AX = \mathbf{1}_n$, per tant A és invertible.

Si la matriu A és invertible, $x = A^{-1}B$, per tant el sistema és compatible determinat per a tot B . \square

Volem aprofitar les matrius per estudiar espais vectorials que no siguin del tipus \mathbb{K}^n . Com a cas particular, tenim els subespais. Comencem veient que el concepte de base també es pot estendre a subespais de \mathbb{K}^m .

Definició 3.38. Considerem $V \subset \mathbb{K}^m$ un subespai vectorial (Definició 3.10). Un conjunt ordenat de vectors $[\vec{v}_1, \dots, \vec{v}_n]$ amb $\vec{v}_i \in V$ són una base de V si són linealment independents i generen V .

Podem fer les observacions següents:

Proposició 3.39. *Si $[\vec{v}_1, \dots, \vec{v}_n]$ és una base d'un subespai $V \subset \mathbb{K}^m$, llavors:*

- (a) $n \leq m$ i
- (b) $n = m$ si i només si $V = \mathbb{K}^m$.

Demostració. (a) és conseqüència del Corol·lari 3.24: si $\{\vec{v}_1, \dots, \vec{v}_n\}$ són linealment independents a \mathbb{K}^m , llavors $n \leq m$.

(b) Suposem que $n = m$, per la Proposició 3.35, si són linealment independents són base de tot \mathbb{K}^m , per tant $V = \mathbb{K}^m$. I, si $V = \mathbb{K}^m$, llavors qualsevol base ha de tenir m vectors, i $n = m$. \square

Observació 3.40. Sigui V un subespai vectorial de \mathbb{K}^m , $\{\vec{u}_1, \dots, \vec{u}_k\} \subset V$ una família de vectors linealment independents i $\{\vec{v}_1, \dots, \vec{v}_n\} \subset V$ una família de vectors generadors de V . Llavors $k \leq n$.

Demostració. Com que $\vec{u}_1 \in V = \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, existeixen $\lambda_i \in \mathbb{K}$ tals que $\vec{u}_1 = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$. Com que $\vec{u}_1 \neq \vec{0}$ (el $\vec{0}$ no forma part de cap família de vectors linealment independents), existeix $\lambda_i \neq 0$. Si cal, reordenem els vectors $\vec{v}_1, \dots, \vec{v}_n$ i suposem $\lambda_1 \neq 0$. Aplicant ara el Lema 3.28, tenim:

$$V = \langle \vec{v}_1, \dots, \vec{v}_n \rangle = \langle \vec{u}_1, \vec{v}_2, \dots, \vec{v}_n \rangle.$$

Considerem ara \vec{u}_2 : com que $\vec{u}_2 \in V = \langle \vec{u}_1, \vec{v}_2, \dots, \vec{v}_n \rangle$, existeixen $\mu_1 \in \mathbb{K}$ i $\lambda_i \in \mathbb{K}$ ($i \geq 2$) tals que $\vec{u}_2 = \mu_1 \vec{u}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n$. Com que $\{\vec{u}_1, \vec{u}_2\}$ són linealment independents, existeix λ_i tal que $\lambda_i \neq 0$ amb $i \geq 2$. Si cal, podem reordenar els vectors $\{\vec{v}_2, \dots, \vec{v}_n\}$ i suposar que $\lambda_2 \neq 0$. Tornem a aplicar el Lema 3.28 i obtenim:

$$V = \langle \vec{u}_1, \vec{v}_2, \dots, \vec{v}_n \rangle = \langle \vec{u}_1, \vec{u}_2, \vec{v}_3, \dots, \vec{v}_n \rangle.$$

Iterem ara aquest procediment. Si $k > n$, llavors podem iterar el procediment amb els n primers vectors de $\{\vec{u}_1, \dots, \vec{u}_k\}$, obtenint:

$$V = \langle \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \rangle$$

i per tant $\vec{u}_{n+1} \in V = \langle \vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \rangle$, que contradiu el fet de que $\{\vec{u}_1, \dots, \vec{u}_n, \vec{u}_{n+1}\}$ siguin linealment independents. \square

Teorema 3.41. *Si $V \subset \mathbb{K}^m$ és un subespai vectorial, totes les bases de V tenen el mateix nombre d'elements i l'anomenem dimensió de V .*

Demostració. Suposem que tenim $\mathcal{B}_1 = [\vec{u}_1, \dots, \vec{u}_k]$ i $\mathcal{B}_2 = [\vec{v}_1, \dots, \vec{v}_n]$ dues bases de V . Com que els vectors de \mathcal{B}_1 són linealment independents i els de \mathcal{B}_2 generadors de V , $k \leq n$. Com que els vectors de \mathcal{B}_1 són generadors de V i \mathcal{B}_2 són linealment independents, $k \geq n$. Per tant $k = n$. \square

Veiem que aquests resultats són molt semblants als que obteníem per \mathbb{K}^n i els podem resumir a:

Teorema 3.42. *Si V és un subespai de \mathbb{K}^m de dimensió n , llavors:*

- (a) *n és el nombre màxim que pot tenir una família de vectors linealment independents de V .*
- (b) *n és el nombre mínim que pot tenir una família de vectors generadors de V .*
- (c) *n vectors de V són linealment independents si i només si són generadors de V , si i només si són base de V .*

Demostració. Només cal veure (c), i de (c) que $\vec{v}_1, \dots, \vec{v}_n$ són linealment independents si i només si generen V : si són linealment independents, i no generen V , podem trobar un vector $\vec{u} \notin \langle \vec{v}_1, \dots, \vec{v}_n \rangle$, per tant $\{\vec{u}, \vec{v}_1, \dots, \vec{v}_n\}$ és una família de vectors linealment independents que té $n+1$ vectors, contradient (a). Si $\vec{v}_1, \dots, \vec{v}_n$ són generadors i no són linealment independents, vol dir que un d'ells es pot escriure com ca combinació lineal dels altres, i per tant tindríem una família de vectors generadors amb $n-1$ vectors, que no pot ser. \square

Aprofitem tots aquests continguts per a veure com es pot trobar una base de la imatge d' $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$, una aplicació lineal:

Lema 3.43. *Si $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ és una aplicació lineal i A la matriu tal que $f = f_A$, llavors $\text{Im}(f) = \langle f(\vec{e}_1), \dots, f(\vec{e}_n) \rangle$ i $\dim(\text{Im}(f)) = \text{Rang}(A)$.*

Demostració. A és la matriu que té per columnes $f(\vec{e}_1), \dots, f(\vec{e}_n)$, i a la Proposició 3.3 varem veure que $\text{Im}(f) = \langle f(\vec{e}_1), \dots, f(\vec{e}_n) \rangle$.

Pel Lema 3.29, el rang d' A és el nombre màxim de vectors linealment independents que hi ha entre les seves columnes, per tant, també generen i formen una base, obtenint que és la dimensió de $\text{Im}(f)$. \square

Per trobar una base de la imatge, podem utilitzar el resultat següent:

Proposició 3.44. *Considerem $A \in M_{m \times n}(\mathbb{K})$ i la matriu reduïda equivalent $\text{rref}(A)$ (que l'hem obtingut amb operacions elementals per files). Siguin j_1, \dots, j_k les columnes de $\text{rref}(A)$ que contenen un pivot. Llavors $[f_A(\vec{e}_{j_1}), \dots, f_A(\vec{e}_{j_k})]$ és una base d' $\text{Im}(f_A)$.*

Demostració. Pel Lema 3.43 i el Teorema 3.42 (c), tant sols cal demostrar que $\{f_A(\vec{e}_{j_1}), \dots, f_A(\vec{e}_{j_k})\}$ són linealment independents: considerem B la submatriu d' A formada per les columnes j_1, \dots, j_k d' A i hi apliquem les mateixes transformacions que passen d' A a $\text{rref}(A)$. Obtindrem una matriu reduïda amb les columnes j_i, \dots, j_k de $\text{rref}(A)$, que són les que tenen un pivot, i que serà equivalent a B , per tant, per unicitat, tindrem $\text{rref}(B)$. Però $\text{rref}(B)$ té rang k , i per tant B també. Però k és el nombre de columnes de B , per tant, les columnes de B són linealment independents, que és el que volíem veure. \square

Teorema 3.45. *Sigui $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ l'aplicació lineal induïda per $A \in M_{m \times n}(\mathbb{K})$. Llavors:*

$$n = \dim(\text{Ker}(f_A)) + \dim(\text{Im}(f_A)).$$

Demostració. Es dedueix directament d'interpretar A com la matriu associada a un sistema d'equacions lineals homogeni:

- n és el número d'incògnites, que és igual a les que fan de paràmetre lliure més les que no.
- $\text{Ker}(f_A)$ són les solucions, per tant $\dim(\text{Ker}(f_A))$ són els graus de llibertat, el nombre de paràmetres lliures.
- $\dim(\text{Im}(f_A)) = \text{Rang}(A)$ són el nombre de columnes de $\text{rref}(A)$ que tenen un pivot i, per tant, les que no són paràmetres lliures.

\square

Observació 3.46. En general, els subespais es descriuen o bé mitjançant un sistema d'equacions lineals homogeni o bé a través de generadors. Vegem com podem passar d'una descripció a l'altra. Primer observem que si $V \leq \mathbb{K}^n$ és un subespai de dimensió k , necessitem, com a mínim, k vectors linealment independents per generar-lo, i, un sistema d'equacions lineal homogeni amb, com a mínim, $n - k$ equacions per definir-lo.

- Suposem que $V \subset \mathbb{K}^n$ és el subespai vectorial definit com les solucions d'un sistema d'equacions lineal homogeni amb matriu associada A (sense ampliar, ja que tots els termes independents són zero). En aquest cas, si $\dim(V) = k$, k són els graus de llibertat i $\text{Rang}(A) = n - k$. Resolem el sistema passant a $\text{rref}(A)$ i aïllant les variables corresponents a les columnes que no tenen pivot (termes independents) en funció de les que tenen pivot (k variables). Qualsevol solució es pot escriure d'aquesta manera (per tant els k vectors que sortiran generen V). Tenim que, per l'equació del Teorema 3.45, $k = \dim(\text{Ker}(A))$, obtenint que les solucions han de ser una base de V .

- Suposem que $V = \langle \vec{v}_1, \dots, \vec{v}_n \rangle \subset \mathbb{K}^m$ un subespai de dimensió k (per tant, esperem $m - k$ equacions). Una manera d'obtenir unes equacions que defineixin V és amb el procediment següent: escrivim els vectors \vec{v}_j en columna obtenint una matriu A :

$$A = \begin{pmatrix} | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & | & | \end{pmatrix}$$

i escrivim ara \overline{A} : una matriu nova formada per la matriu A amb una matriu identitat $\mathbf{1}_m$ a la dreta:

$$\overline{A} = (A \mid \mathbf{1}_m)$$

Esglaonem \overline{A} per files, obtenint, a les columnes formades per les \vec{v}_j , un total de k files no nul·les (on $k = \text{Rang}(A)$) i per tant $k = \dim(V)$, i $m - k$ files amb tots els valors zero a les n primeres columnes:

$$\overline{A} \sim \left(\begin{array}{c|c} T & P_1 \\ \hline 0 & P_2 \end{array} \right)$$

Amb $P_2 \in M_{(m-k) \times m}(\mathbb{K})$ complint que $P_2 A = 0$, pel que les columnes d' A (els vectors \vec{v}_j) són solució del sistema d'equacions homogeni determinat per P_2 . A més, P_2 és una submatriu d'una matriu quadrada de rang màxim, per tant les files són linealment independents i per tant les solucions de P_2 tenen dimensió k , obtenint que han de ser exactament V .

Exemple 3.47. Suposem que el subespai V està definit per les equacions següents i en volem calcular una base:

$$\begin{aligned} 3x_1 - 5x_2 + x_3 + x_4 &= 0 \\ x_1 - x_2 - x_3 + x_4 &= 0 \\ x_1 - 2x_2 + x_3 &= 0 \end{aligned}$$

La matriu associada al sistema és:

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

Que és equivalent a:

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

Amb solució $x_1 = 3x_3 - 2x_4$, $x_2 = 2x_3 - x_4$ i paràmetres lliures x_3 i x_4 . Per tant:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3x_3 - 2x_4 \\ 2x_3 - x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \end{pmatrix}$$

I una base de V és:

$$\left[\begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right]$$

Exemple 3.48. Suposem que V està generat per:

$$V = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \right\rangle,$$

i volem troba un sistema d'equacions lineals que descrigui V .

Considerem la matriu:

$$\left(\begin{array}{ccc|cccc} 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 & 0 \\ 1 & 3 & 4 & 0 & 0 & 1 & 0 \\ 1 & 4 & 5 & 0 & 0 & 0 & 1 \end{array} \right),$$

i l'esglaonem per files:

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

Per tant, V també es pot definir com el subespai que correspon a les equacions:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_1 - 3x_2 + x_4 &= 0. \end{aligned}$$

3.4 Suma i intersecció de subespais vectorials

Volem construir subespais vectorials a partir d'altres subespais: si $U, V \subset \mathbb{K}^n$ són subespais vectorials, en particular són conjunts, pel que té sentit considerar la intersecció $U \cap V$ i la unió $U \cup V$.

Veiem que:

- $\vec{0} \in U$ i $\vec{0} \in V$ (ja que són subespai), per tant, $\vec{0} \in U \cap V$,
- si $\vec{v}_1, \vec{v}_2 \in U \cap V$, llavors $\vec{v}_1 + \vec{v}_2 \in U \cap V$: \vec{v}_1 i \vec{v}_2 són vectors d' U i V . Com que U és subespai, $\vec{v}_1 + \vec{v}_2 \in U$. Com que V és subespai, $\vec{v}_1 + \vec{v}_2 \in V$. Obtenim, llavors $\vec{v}_1 + \vec{v}_2 \in U \cap V$ i
- si $\vec{v} \in U \cap V$ i $\lambda \in \mathbb{K}$, llavors si $\lambda\vec{v} \in U \cap V$: com que $\vec{v} \in U$ i U és subespai vectorial, $\lambda\vec{v} \in U$. El mateix argument implica $\lambda\vec{v} \in V$, per tant $\lambda\vec{v} \in U \cap V$.

Obtenint:

Proposició 3.49. Si $U, V \subset \mathbb{K}^n$ són subespais vectorials, llavors $U \cap V$ també és un subespai vectorial.

En canvi, si considerem la unió, si $\vec{v}_1, \vec{v}_2 \in U \cup V$, podem assegurar que $\vec{v}_1 + \vec{v}_2 \in U \cup V$? Vegem que no és cert: considerem $U = \langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \rangle \subset \mathbb{R}^2$ i $V = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle \subset \mathbb{R}^2$, llavors $U \cup V$ són els vectors que pertanyen als dos eixos. Llavors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in U \cup V$, però $\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \notin U \cup V$.

Per això, definim el concepte de suma:

Definició 3.50. Si $U, V \subset \mathbb{K}^m$ són subespais vectorials, definim la suma $U + V$ com:

$$U + V = \{ \vec{u} + \vec{v} \mid \vec{u} \in U \text{ i } \vec{v} \in V \}.$$

Lema 3.51. Si $U, V \subset \mathbb{K}^m$ són subespais vectorials, llavors $U + V$ és un subespai vectorial. Més concretament, $U + V$ és el subespai generat per U i V (podem escriure $U + V = \langle U, V \rangle$).

Demostració. Vegem que és $U + V$ un subespai:

- El vector $\vec{0}$ és d' $U + V$: com que U i V són subespais, $\vec{0}$ pertany a tots dos. A més, es pot escriure com $\vec{0} = \vec{0} + \vec{0}$, per tant és d' $U + V$.
- Si $\vec{w}_1, \vec{w}_2 \in U + V$, vol dir que existeixen $u_1, u_2 \in U$, i $\vec{v}_1, \vec{v}_2 \in V$ tals que $\vec{w}_1 = \vec{u}_1 + \vec{v}_1$ i $\vec{w}_2 = \vec{u}_2 + \vec{v}_2$. Llavors, $\vec{w}_1 + \vec{w}_2 = (\vec{u}_1 + \vec{u}_2) + (\vec{v}_1 + \vec{v}_2)$, amb $\vec{u}_1 + \vec{u}_2 \in U$ i $\vec{v}_1 + \vec{v}_2 \in V$ (ja que tant U com V són subespais vectorials).
- Si $\vec{w} \in U + V$ i $\lambda \in K$, llavors $\lambda\vec{w} \in U + V$: existeixen $\vec{u} \in U$ i $\vec{v} \in V$ tals que $\vec{w} = \vec{u} + \vec{v}$. Llavors $\lambda\vec{w} = \lambda\vec{u} + \lambda\vec{v}$, amb $\lambda\vec{u} \in U$ i $\lambda\vec{v} \in V$ (ja que U i V són subespais).

Ara tenim que $U \subset U + V$ i $V \subset U + V$, per tant, com que $U + V$ és subespai, $\langle U, V \rangle \subset U + V$. Per definició, els vectors d' $U + V$ estan generats pels vectors d' U i de V , per tant $U + V \leq \langle U, V \rangle$. \square

Observació 3.52. Podem calcular la intersecció i la suma de subespais vectorials amb la següent observació:

- Si $U \subset \mathbb{K}^n$ és el subespai vectorial definit com les solucions del sistema d'equacions lineals homogeni que té per matriu associada A , i $V \subset \mathbb{K}^n$ el que té per matriu associada B , llavors:

$$U \cap V \text{ és solució del sistema que té per matriu associada } \begin{pmatrix} A \\ B \end{pmatrix}.$$

Dit d'altra manera, si $U = \text{Ker}(f_A)$ i $V = \text{Ker}(f_B)$, aleshores $U \cap V = \text{Ker}(f_{\begin{pmatrix} A \\ B \end{pmatrix}})$.

- Si $U = \langle \vec{u}_1, \dots, \vec{u}_k \rangle \subset \mathbb{K}^n$ i $V = \langle \vec{v}_1, \dots, \vec{v}_l \rangle \subset \mathbb{K}^n$, llavors:

$$U + V = \langle \vec{u}_1, \dots, \vec{u}_k, \vec{v}_1, \dots, \vec{v}_l \rangle.$$

Dit d'altra manera, si $U = \text{Im}(f_A)$ i $V = \text{Im}(f_B)$, aleshores $U + V = \text{Im}(f_{(A|B)})$.

Exemple 3.53. Suposem que se'ns donen els següents dos subespais d' \mathbb{R}^4 :

$$U = \left\langle \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \\ -1 \end{pmatrix} \right\rangle, \quad V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : x_1 + x_2 + 2x_3 + x_4 = 0 \right\}.$$

Se'ns demana que donem una base d' $U \cap V$. Per resoldre aquest problema, observem que per calcular la intersecció ens cal tenir equacions per U i V . Per tant, ens cal realitzar els següents passos:

1. Trobar equacions per U ,
2. Definir equacions per $U \cap V$, i
3. Trobar una base d' $U \cap V$.

Per trobar equacions per U , esglaonem la matriu ampliada amb la identitat:

$$\left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array} \right) \sim \dots \sim \left(\begin{array}{cc|cccc} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 1 \\ 0 & 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -5 & 0 & 2 & 3 \end{array} \right).$$

Així, U ve definit per les equacions

$$\begin{aligned} -2x_1 + x_2 &= 0 \\ -5x_1 + 2x_3 + 3x_4 &= 0. \end{aligned}$$

Les equacions per $U \cap V$ són, doncs:

$$\begin{aligned} x_1 + x_2 + 2x_3 + x_4 &= 0 \\ -2x_1 + x_2 &= 0 \\ -5x_1 + 2x_3 + 3x_4 &= 0. \end{aligned}$$

Finalment, per trobar una base d' $U \cap V$ hem de resoldre el sistema anterior. Esglaonant la matriu corresponent, obtenim la matriu

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

Així, una base d' $U \cap V$ ve donada per

$$\left[\begin{pmatrix} -1 \\ 2 \\ -2 \\ 3 \end{pmatrix} \right].$$

Observació 3.54. Considerem $U, V \subset \mathbb{K}^n$ subespais vectorials. Suposem que $j = \dim(U \cap V)$, $k = \dim(U)$ i $l = \dim(V)$. Siguin:

- $\mathcal{B} = [\vec{w}_1, \dots, \vec{w}_j]$ una base d' $U \cap V$,
- $\mathcal{B}_U = [\vec{w}_1, \dots, \vec{w}_j, \vec{u}_{j+1}, \dots, \vec{u}_k]$ una base d' U i
- $\mathcal{B}_V = [\vec{w}_1, \dots, \vec{w}_j, \vec{v}_{j+1}, \dots, \vec{v}_l]$ una base de V .

Observem que hem començat amb una base d' $U \cap V$ que, per una banda hem ampliat fins a base d' U i per altra, fins a base de V .

Veiem que:

- $\{\vec{w}_1, \dots, \vec{w}_j, \vec{u}_{j+1}, \dots, \vec{u}_k, \vec{v}_{j+1}, \dots, \vec{v}_l\}$ generen $U + V$: si $\vec{w} \in U + V$, existeixen $\vec{u} \in U$ i $\vec{v} \in V$ tals que $\vec{w} = \vec{u} + \vec{v}$. Com que \mathcal{B}_U i \mathcal{B}_V són bases d' U i V respectivament, llavors, podem escriure:

$$\begin{aligned} \vec{u} &= \lambda_1 \vec{w}_1 + \dots + \lambda_j \vec{w}_j + \lambda_{j+1} \vec{u}_{j+1} + \dots + \lambda_k \vec{u}_k \\ \vec{v} &= \mu_1 \vec{w}_1 + \dots + \mu_j \vec{w}_j + \mu_{j+1} \vec{v}_{j+1} + \dots + \mu_l \vec{v}_l \end{aligned}$$

i per tant:

$$\vec{w} = (\lambda_1 + \mu_1) \vec{w}_1 + \dots + (\lambda_j + \mu_j) \vec{w}_j + \lambda_{j+1} \vec{u}_{j+1} + \dots + \lambda_k \vec{u}_k + \mu_{j+1} \vec{v}_{j+1} + \dots + \mu_l \vec{v}_l.$$

- $\{\vec{w}_1, \dots, \vec{w}_j, \vec{u}_{j+1}, \dots, \vec{u}_k, \vec{v}_{j+1}, \dots, \vec{v}_l\}$ són linealment independents: si

$$\vec{0} = \lambda_1 \vec{w}_1 + \dots + \lambda_j \vec{w}_j + \lambda_{j+1} \vec{u}_{j+1} + \dots + \lambda_k \vec{u}_k + \mu_{j+1} \vec{v}_{j+1} + \dots + \mu_l \vec{v}_l, \quad (2)$$

llavors:

$$\vec{w} := \lambda_1 \vec{w}_1 + \dots + \lambda_j \vec{w}_j + \lambda_{j+1} \vec{u}_{j+1} + \dots + \lambda_k \vec{u}_k = -(\mu_{j+1} \vec{v}_{j+1} + \dots + \mu_l \vec{v}_l)$$

és un vector d' U i V , per tant d' $U \cap V$. Llavors, es pot escriure de manera única com:

$$\vec{w} = \gamma_1 \vec{w}_1 + \cdots \gamma_j \vec{w}_j,$$

i per tant $\lambda_i = \gamma_i$ per a $i \leq j$ i $\lambda_{j+1} = \cdots \lambda_k = 0$. Llavors, de l'Equació (2) ens queda la igualtat:

$$\vec{0} = \lambda_1 \vec{w}_1 + \cdots \lambda_j \vec{w}_j + \mu_{j+1} \vec{v}_{j+1} + \cdots \mu_k \vec{v}_k$$

i tots els coeficients són zero ja que \mathcal{B}_V és una base de V .

Per tant $[\vec{w}_1, \dots, \vec{w}_j, \vec{u}_{j+1}, \dots, \vec{u}_k, \vec{v}_{j+1}, \dots, \vec{v}_l]$ és una base d' $U + V$.

Si ara comptem quants elements té cadascuna de les bases, tenim:

Teorema 3.55. *Si $U, V \subset \mathbb{K}^n$ són subespais vectorials, llavors:*

$$\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V).$$

Definició 3.56. Si $U, V, W \subset \mathbb{K}^n$ són subespais vectorials, diem que W és la *suma directa* d' U i V , i escrivim $W = U \oplus V$, si $W = U + V$ i $U \cap V = \{\vec{0}\}$.

Proposició 3.57. *Si $U, V, W \subset \mathbb{K}^n$ són subespais vectorials, llavors $W = U \oplus V$ si, i només si, tot vector de W es pot escriure de forma única com a suma d'un vector d' U més un vector de V .*

Demostració. Si $W = U \oplus V$ llavors, com que $W = U + V$, per a tot $\vec{w} \in W$, existeixen $\vec{u} \in U$ i $\vec{v} \in V$ tals que $\vec{w} = \vec{u} + \vec{v}$. Falta veure que \vec{u} i \vec{v} són únics: suposem que també podem escriure $\vec{w} = \vec{u}' + \vec{v}'$, amb $\vec{u}' \in U$ i $\vec{v}' \in V$. Tenim:

$$\vec{0} = \vec{w} - \vec{w} = (\vec{u} - \vec{u}') + (\vec{v} - \vec{v}')$$

i per tant:

$$\vec{u} - \vec{u}' = -(\vec{v} - \vec{v}') \in U \cap V$$

però com que U i V estan en suma directa, $U \cap V = \{\vec{0}\}$, obtenint que $\vec{u} = \vec{u}'$ i $\vec{v} = \vec{v}'$.

Suposem que tot vector $\vec{w} \in W$ es pot escriure de forma única com $\vec{w} = \vec{u} + \vec{v}$, amb $\vec{u} \in U$ i $\vec{v} \in V$. Com que es pot escriure, tenim $W = U + V$. Sigui $\vec{w} \in U \cap V$. Podem escriure \vec{w} com:

$$\vec{w} = \vec{w} + \vec{0} = \vec{0} + \vec{w}$$

on a cada part de la igualtat el primer vector és d' U i el segon de V . Per unicitat, les dues expressions han de coincidir i per tant $\vec{w} = \vec{0}$. \square

Exemple 3.58. A \mathbb{R}^3 , considerem el pla $U = \{x = 0\}$ i el pla $V = \{y = 0\}$. Aleshores

$\mathbb{R}^3 = U + V$, ja que to vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ es pot escriure (per exemple) com

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} + \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix},$$

on el primer sumand pertany a U i el segon a V . La suma, però, no és directa:

$$U \cap V = \{x = y = 0\} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$$

En canvi, si en comptes de V considerem la recta definida per $V' = \{y = z = 0\}$, aleshores $\mathbb{R}^3 = U \oplus V'$.

3.5 Aplicacions injectives, exhaustives i bijectives

Recordem unes definicions que tenen sentit per qualssevol aplicacions entre conjunts: si A i B són conjunts i $f: A \rightarrow B$ és una aplicació entre A i B (això vol dir que assignem a cada element $a \in A$ un element $f(a) \in B$, i anomenem a $f(a)$ *imatge de l'element a per f*), diem que:

- f és *injectiva* si quan $a \neq a'$, llavors $f(a) \neq f(a')$ (envia elements diferents a elements diferents).
- f és *exhaustiva* si per a tot $b \in B$, existeix algun $a \in A$ tal que $f(a) = b$ (tot element té antiimatge).
- f és *bijectiva* si és injectiva i exhaustiva. Equivalentment, per a tot $b \in B$ existeix un únic $a \in A$ tal que $f(a) = b$.
- Anomenem $\text{Im}(f) \subset B$ a la *imatge de l'aplicació f* . Per tant, f és exhaustiva si i només si $\text{Im}(f) = B$.
- Si tenim $f: A \rightarrow B$ i $g: B \rightarrow C$ aplicacions de conjunts, definim la composició $g \circ f: A \rightarrow C$ com l'aplicació $(g \circ f)(a) = g(f(a))$.

Exemple 3.59. • $A = \{1, 2, 3\}$, $B = \{a, b, c, d\}$ i definim $f: A \rightarrow B$ com $f(1) = a$, $f(2) = b$ i $f(3) = a$. Aquesta aplicació no és ni injectiva ($f(1) = f(3)$ quan $1 \neq 3$), ni exhaustiva ($c \notin \text{Im}(f)$).

- $A = \{1, 2, 3\}$, $B = \{a, b, c, d\}$ i definim $g: A \rightarrow B$ com $g(1) = a$, $g(2) = b$ i $g(3) = d$. Aquesta aplicació és injectiva, però no exhaustiva ($c \notin \text{Im}(g)$).
- $A = \{1, 2, 3\}$, $B = \{a, b, c, d\}$ i definim $h: B \rightarrow A$ com $h(a) = 1$, $h(b) = 2$, $h(c) = 1$ i $h(d) = 3$. Aquesta aplicació no és injectiva ($h(c) = h(a)$), però sí que és exhaustiva.
- Veiem que $h \circ g: A \rightarrow A$ és l'aplicació que envia cada element a ell mateix (l'anomenem aplicació identitat d' A i la denotem per Id_A).

Les aplicacions lineals entre espais vectorials són un cas particular d'aplicacions de conjunts, per tant té sentit de parlar d'aplicacions lineals injectives, exhaustives i bijectives. Tenim una caracterització de les aplicacions lineals injectives:

Lema 3.60. $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ una aplicació lineal. Llavors f és injectiva si i només si $\ker(f) = \{\vec{0}\}$.

Demostració. Suposem que f és injectiva. Com que és lineal $f(\vec{0}) = \vec{0}$, per tant $\{\vec{0}\} \subset \ker(f)$. Si $\ker(f)$ tingués algun altre vector $\vec{v} \neq \vec{0}$, llavors f no seria injectiva (tindria dos vectors amb la mateixa imatge).

Suposem que f no és injectiva: llavors existeixen $\vec{u} \neq \vec{v}$ tals que $f(\vec{u}) = f(\vec{v})$. Però llavors $f(\vec{u} - \vec{v}) = f(\vec{u}) - f(\vec{v}) = \vec{0}$, amb $\vec{u} - \vec{v} \neq \vec{0}$, i tenim que $\ker(f) \neq \{\vec{0}\}$. \square

Exercici 3.3. Considerem $A \in M_{m \times n}(\mathbb{K})$ i l'aplicació lineal $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^m$ definida a l'Exemple 3.1. Demostreu que:

- f_A és injectiva si i només si $\text{Rang}(A) = n$.
- f_A és exhaustiva si i només si $\text{Rang}(A) = m$.
- f_A és bijectiva si i només si $n = m$ i A és invertible (en aquest cas diem que f_A és un isomorfisme).

3.6 Coordenades de vectors

Definició 3.61. Suposem que V és un subespai vectorial de \mathbb{K}^m i $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ una base de V . Si $\vec{w} \in V$, definim les *coordenades de \vec{w} en la base \mathcal{B}* com el vector $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tal que

$$\vec{w} = \sum_{i=1}^n \lambda_i \vec{v}_i.$$

Si hi pot haver confusió, escriurem $[\vec{w}]_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)$ per parlar de les coordenades d'un vector en la base \mathcal{B} .

Cal veure que aquesta definició és correcta:

Teorema 3.62. *Suposem V és un subespai vectorial de \mathbb{K}^m i $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ una base de V . Llavors tot vector $\vec{w} \in V$ té coordenades i són úniques.*

Demostració. És el que diu el Teorema 3.36. □

Observació 3.63. Podem veure que si fixem una base \mathcal{B} d'un subespai $V \subset \mathbb{K}^m$ i $[\vec{v}_1]_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)$, $[\vec{v}_2]_{\mathcal{B}} = (\mu_1, \dots, \mu_n)$ i $\gamma \in \mathbb{K}$, tenim que:

- $[\vec{v}_1 + \vec{v}_2]_{\mathcal{B}} = (\lambda_1 + \mu_1, \dots, \lambda_n + \mu_n)$ i
- $[\gamma \vec{v}_1]_{\mathcal{B}} = (\gamma \lambda_1, \dots, \gamma \lambda_n)$.

Aquesta observació ens permet identificar qualsevol subespai vectorial de \mathbb{K}^m amb una base fixada \mathcal{B} amb n vectors amb l'espai vectorial \mathbb{K}^n . Les aplicacions lineals entre \mathbb{K}^n i \mathbb{K}^m les hem estudiat a la Proposició 3.3, obtenint que es corresponen amb matrius $A \in M_{m \times n}(\mathbb{K})$. Això ens permet fer la definició següent:

Definició 3.64. Si $f: V \rightarrow W'$ és una aplicació lineal entre subespais vectorials de \mathbb{K}^m i \mathbb{K}^r respectivament amb bases $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ i $\mathcal{B}' = [\vec{w}_1, \dots, \vec{w}_s]$ respectivament, definim la *matriu associada a f en les bases \mathcal{B} i \mathcal{B}'* (escriurem $[f]_{\mathcal{B}}$ si l'espai de sortida és el mateix d'arribada i utilitzem la mateixa base \mathcal{B}) i l'escriuim com $[f]_{\mathcal{B}, \mathcal{B}'}$ com la matriu A que té per coeficients a_{ij} complint:

$$f(\vec{v}_j) = \sum_{i=1}^s a_{ij} \vec{w}_i.$$

O sigui,

$$[f]_{\mathcal{B}, \mathcal{B}'} = \begin{pmatrix} | & | & & | \\ [f(\vec{v}_1)]_{\mathcal{B}} & [f(\vec{v}_2)]_{\mathcal{B}} & \cdots & [f(\vec{v}_n)]_{\mathcal{B}} \\ | & | & & | \end{pmatrix}$$

Del càlcul fet a la Proposició 3.3 es dedueix que aquesta matriu compleix que si $[\vec{v}]_{\mathcal{B}} = (\lambda_1, \dots, \lambda_n)$, llavors $[f(v)]_{\mathcal{B}'} = (\mu_1, \dots, \mu_s)$, on:

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_s \end{pmatrix} = \begin{pmatrix} | & | & & | \\ [f(\vec{v}_1)]_{\mathcal{B}} & [f(\vec{v}_2)]_{\mathcal{B}} & \cdots & [f(\vec{v}_n)]_{\mathcal{B}} \\ | & | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \quad (3)$$

Vegem un cas particular d'aquest argument, quan $V = \mathbb{K}^n$ i $W = \mathbb{K}^m$. De la definició de matriu associada i que la composició d'aplicacions es correspon amb la multiplicació de matrius, es dedueixen els resultats següents:

Proposició 3.65. *Fixem les notacions següents:*

- $\mathcal{S}_n = [\vec{e}_1, \dots, \vec{e}_n]$ és la base de \mathbb{K}^n formada pels vectors estàndard,
- $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ és una base (qualsevol) de \mathbb{K}^n , \mathcal{B}' de \mathbb{K}^m i \mathcal{B}'' de \mathbb{K}^r ,
- $A \in M_{m \times n}(\mathbb{K})$ i $f_A: \mathbb{K}^m \rightarrow \mathbb{K}^n$ l'aplicació lineal induïda per A tal i com hem definit a l'Exemple 3.1 i
- $\text{Id}_n: \mathbb{K}^n \rightarrow \mathbb{K}^n$ és l'aplicació lineal definida per $\text{Id}(\vec{v}) = \vec{v}$ (que també es pot pensar com $f_{\mathbf{1}_n}$)

Llavors:

(a) $[\text{Id}_n]_{\mathcal{B}} = \mathbf{1}_n$ (per a qualsevol base \mathcal{B} de \mathbb{K}^n),

(b)

$$[\text{Id}_n]_{\mathcal{B}, \mathcal{S}_n} = \begin{pmatrix} | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & | & | \end{pmatrix},$$

la matriu que té per columnes els vectors de \mathcal{B} .

(c) $[\text{Id}_n]_{\mathcal{S}_n, \mathcal{B}} = [\text{Id}_n]_{\mathcal{B}, \mathcal{S}_n}^{-1}$.

(d) $[g \circ f]_{\mathcal{B}, \mathcal{B}''} = [g]_{\mathcal{B}', \mathcal{B}''} \cdot [f]_{\mathcal{B}, \mathcal{B}'}$ per a totes $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ i $g: \mathbb{K}^m \rightarrow \mathbb{K}^r$ aplicacions lineals.

(e) $A = [f_A]_{\mathcal{S}_n, \mathcal{S}_m}$.

(f) $[f_A]_{\mathcal{B}, \mathcal{B}'} = [\text{Id}_m]_{\mathcal{B}', \mathcal{S}_m}^{-1} \cdot A \cdot [\text{Id}_n]_{\mathcal{B}, \mathcal{S}_n}$.

Demostració. (a), (b) i (e) es dedueixen de la definició. Si demostrem (d), es deduiran la resta de resultats, ja que utilitzen la relació entre composició d'aplicacions lineals i el producte de matrius en les bases corresponents.

Per demostrar (d) utilitzem directament la Definició 3.64 amb l'expressió de l'Equació (3):

si $\vec{v} \in \mathbb{K}^n$ i $[\vec{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$, llavors, les coordenades en la base \mathcal{B}' d' $f(\vec{v})$ són $[f(\vec{v})]_{\mathcal{B}'} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}$,

on:

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = [f]_{\mathcal{B}, \mathcal{B}'} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

I les coordenades de $g(f(v))$ en la base \mathcal{B}'' són $[g(f(v))]_{\mathcal{B}''} = (\gamma_1, \dots, \gamma_r)$, on:

$$\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_r \end{pmatrix} = [g]_{\mathcal{B}', \mathcal{B}''} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = [g]_{\mathcal{B}', \mathcal{B}''} [f]_{\mathcal{B}, \mathcal{B}'} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

I per tant tenim (d): aquesta propietat caracteritza la matriu $[g \circ f]_{\mathcal{B}, \mathcal{B}''}$. \square

Una manera de trobar aquestes igualtats és escriure les aplicacions en forma de diagrama, composant-les (posem les bases com a subíndex, amb \mathcal{S} sense especificar la dimensió, que es dedueix de l'espai):

$$\begin{array}{ccc} (\mathbb{K}^n)_{\mathcal{S}} & \xrightarrow{f_A} & (\mathbb{K}^m)_{\mathcal{S}} \\ \text{Id} \updownarrow & & \updownarrow \text{Id} \\ (\mathbb{K}^n)_{\mathcal{B}} & \xrightarrow{f_A} & (\mathbb{K}^m)_{\mathcal{B}'} \end{array}$$

Veiem que per anar de $(\mathbb{K}^n)_{\mathcal{B}}$ a $(\mathbb{K}^m)_{\mathcal{B}'}$ podem seguir dos camins, amb el mateix resultat (diem que el diagrama és commutatiu). Si ara ho escrivim en matrius (de les fletxes verticals tant sols hem deixat la que dóna la matriu):

$$\begin{array}{ccc} (\mathbb{K}^n)_{\mathcal{S}} & \xrightarrow{A} & (\mathbb{K}^m)_{\mathcal{S}} \\ \uparrow B & & \uparrow B' \\ (\mathbb{K}^n)_{\mathcal{B}} & \xrightarrow{[f_A]_{\mathcal{B},\mathcal{B}'}} & (\mathbb{K}^m)_{\mathcal{B}'} \end{array}$$

on $B = [Id_n]_{\mathcal{B},\mathcal{S}_n}$ i $B' = [Id_m]_{\mathcal{B}',\mathcal{S}_m}$ (les matrius que tenen per columnes les bases \mathcal{B} i \mathcal{B}' respectivament). Si ara tenim en compte que les un vector en les coordenades de $(\mathbb{K}^n)_{\mathcal{B}}$ pot seguir els dos camins per arribar a $(\mathbb{K}^m)_{\mathcal{S}}$, obtenim una igualtat equivalent a l'apartat (f) de la Proposició 3.65.

Considerem ara el cas particular $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ ($n = m$ al que hem fet fins ara). Llavors podem identificar el \mathbb{K}_n de l'espai de sortida amb el d'arribada, i, si considerem \mathcal{B} una base de \mathbb{K}^n , la podem considerar com a base de l'espai de sortida i el d'arribada. De l'apartat (f) de la Proposició 3.65, obtenim la igualtat de matrius:

$$[f]_{\mathcal{B},\mathcal{B}} = [f]_{\mathcal{B}} = [Id_n]_{\mathcal{B},\mathcal{S}_n}^{-1} [f]_{\mathcal{S}_n,\mathcal{S}_n} [Id_n]_{\mathcal{B},\mathcal{S}_n}$$

Per tant si anomenem $B = [Id_n]_{\mathcal{B},\mathcal{S}_n}$, $A = [f]_{\mathcal{S}_n,\mathcal{S}_n}$ i $A' = [f]_{\mathcal{B},\mathcal{B}}$ tenim que:

$$A' = B^{-1}AB.$$

Això ens porta a la definició següent:

Definició 3.66. Diem que A i A' , dues matrius quadrades $n \times n$, són *similars* si existeix una matriu invertible $B \in M_{n \times n}(\mathbb{K})$ tal que

$$A' = B^{-1}AB.$$

O sigui, si corresponen a una mateixa aplicació lineal, però amb bases diferents.

Proposició 3.67. La relació “ser similar a” és una relació d'equivalència, o sigui, compleix:

- A és similar a A (reflexiva),
- si A és similar a A' , llavors A' és similar a A (simètrica),
- si A és similar a A' i A' és similar a A'' , llavors A és similar a A'' (transitiva).

Demostració. Per demostrar la propietat reflexiva, observem que $A = \mathbf{1}_n^{-1}A\mathbf{1}_n$.

Per demostrar la propietat simètrica, observem que si $A' = B^{-1}AB$, llavors $A = BA'B^{-1}$, per tant considerem B^{-1} a la definició.

Per demostrar la propietat transitiva: l'enunciat ens diu que existeixen matrius invertibles B i B' tals que:

$$A' = B^{-1}AB \text{ i } A'' = (B')^{-1}A'B'$$

llavors tenim:

$$A'' = (B')^{-1}A'B' = (B')^{-1}(B^{-1}AB)B' = (BB')^{-1}A(BB')$$

per tant, A'' i A' són similars. □

Exemple 3.68. Són les matrius $A = \begin{pmatrix} -16 & -45 \\ 6 & 17 \end{pmatrix}$ i $A' = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$ similars?

Per respondre la pregunta necessitem calcular la B tal que $A' = B^{-1}AB$, o bé demostrar que no existeix aquesta B . Intentem calcular-la:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

i ha de complir que $BA' = AB$ (així ens estalviem d'escriure la inversa), obtenint que:

$$\begin{pmatrix} -b_{11} & 2b_{12} \\ -b_{21} & 2b_{22} \end{pmatrix} = \begin{pmatrix} -16b_{11} - 45b_{21} & -16b_{12} - 45b_{22} \\ 6b_{11} + 17b_{21} & 6b_{12} + 17b_{22} \end{pmatrix}$$

Igualant coeficient a coeficient arribem a un sistema d'equacions lineal amb 4 incògnites i 4 equacions, que si el resollem té 2 paràmetres lliures i una possible solució particular és

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

i per tant A i A' són similars.

Exemple 3.69. Són les matrius $A = \begin{pmatrix} -16 & -45 \\ 6 & 17 \end{pmatrix}$ i $A' = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ similars? (així aprofitem els càlculs d'abans).

Intentem calcular B definida com:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

i ha de complir que és invertible i $BA' = AB$, obtenint que:

$$\begin{pmatrix} b_{11} & 2b_{12} \\ b_{21} & 2b_{22} \end{pmatrix} = \begin{pmatrix} -16b_{11} - 45b_{21} & -16b_{12} - 45b_{22} \\ 6b_{11} + 17b_{21} & 6b_{12} + 17b_{22} \end{pmatrix}$$

Igualant coeficient a coeficient arribem a un sistema d'equacions lineal amb 4 incògnites i 4 equacions, que si el resollem té 1 paràmetre lliure i les solucions són

$$B = \begin{pmatrix} 0 & -5\lambda \\ 0 & 2\lambda \end{pmatrix}$$

Observem que aquesta matriu mai pot ser invertible (té una columna tot zeros) i per tant A i A' no són similars.

Veiem que hem hagut de fer molts càlculs per comprovar si són similars o no. Més endavant, veurem resultats que ens permetrà determinar si són similars amb menys càlculs.

Exercici 3.4. Determineu per a quins valors de $\lambda_1, \lambda_2, \mu_1$ i μ_2 les matrius

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ i } A' = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}$$

són similars.

3.7 Espais vectorials

Fins ara hem estat treballant amb \mathbb{K}^n , però tan sols estàvem utilitzant algunes de les propietats que té \mathbb{K}^n . Quan un té uns resultats, és habitual intentar veure si es poden generalitzar a altres objectes.

Exemple 3.70. Considerem els polinomis en una variable amb coeficients en \mathbb{K} de grau més petit o igual que 2. Per tant, un element serà $p(x) = a_0 + a_1x + a_2x^2$ amb $a_i \in \mathbb{K}$. Anomenem $\mathbb{K}[x]_{\leq 2}$ a aquest conjunt. Observem que a $\mathbb{K}[x]_{\leq 2}$ tenim:

- Una operació suma: donats dos polinomis de grau més petit o igual a 2, els podem sumar, i tenim un polinomi de grau més petit o igual a 2. Aquesta suma és commutativa, associativa, té un element neutre (el polinomi 0) i per a qualsevol $p(x) \in \mathbb{K}[x]_{\leq 2}$, existeix un altre $q(x) \in \mathbb{K}[x]_{\leq 2}$ tal que $p(x) + q(x) = 0$.
- Una operació producte per escalar: donat un $\lambda \in \mathbb{K}$ i un polinomi $p(x) \in \mathbb{K}[x]_{\leq 2}$, podem fer la multiplicació $\lambda p(x)$ i, com que no ha crescut el grau, $\lambda p(x) \in \mathbb{K}[x]_{\leq 2}$. Aquesta operació compleix que multiplicar per 1 és no fer res, és distributiva respecte la suma d'escalars i polinomis, i que $\lambda(\mu p(x)) = (\lambda\mu)p(x)$.

Aquestes propietats impliquen que podem fer tot el que fèiem a \mathbb{K}^n , canviant les coordenades pels coeficients a cada grau. Més concretament, si considerem l'aplicació:

$$\begin{aligned} f: \mathbb{K}[x]_{\leq 2} &\longrightarrow \mathbb{K}^3 \\ a_0 + a_1x + a_2x^2 &\mapsto (a_0, a_1, a_2) \end{aligned}$$

compleix que:

- És lineal: $f(p(x) + q(x)) = f(p(x)) + f(q(x))$ i $f(\lambda p(x)) = \lambda f(p(x))$ per a tots $p(x), q(x) \in \mathbb{K}[x]_{\leq 2}$ i $\lambda \in \mathbb{K}$.
- És bijectiva (i per tant un isomorfisme).

I això voldrà dir que tot el que hem fet a \mathbb{K}^3 i té a veure amb aquesta estructura (suma de vectors i multiplicació per escalar) també es podrà fer a $\mathbb{K}[x]_{\leq 2}$.

Definició 3.71. Un \mathbb{K} -espai vectorial E és un conjunt E amb dues operacions, suma i producte per escalar:

$$\begin{aligned} +: E \times E &\longrightarrow E & \text{i} & \quad \cdot: \mathbb{K} \times E \longrightarrow E \\ (u, v) &\mapsto u + v & & \quad (\lambda, v) \mapsto \lambda \cdot v \end{aligned}$$

tals que:

1. La suma és commutativa: $u + v = v + u$ per a tots $u, v \in E$,
2. associativa $(u + v) + w = u + (v + w)$ per a tots $u, v, w \in E$,
3. té element neutre: existeix $0 = 0_E \in E$ tal que $0 + u = u$ per a tot $u \in E$ i
4. tot $u \in E$ té un element invers v complint que $u + v = 0$ (escrivim $v = -u$).
5. El producte per escalar compleix que $1 \cdot u = u$ per a tot $u \in E$,
6. $(\lambda\mu) \cdot u = \lambda \cdot (\mu \cdot u)$ per a tot $\lambda, \mu \in \mathbb{K}$ i $u \in E$,
7. $(\lambda + \mu) \cdot u = (\lambda \cdot u) + (\mu \cdot u)$ per a tot $\lambda, \mu \in \mathbb{K}$ i $u \in E$ i
8. $\lambda \cdot (u + v) = (\lambda \cdot u) + (\lambda \cdot v)$ per a tot $\lambda \in \mathbb{K}$ i $u, v \in E$.

Exemple 3.72. • Els conjunts \mathbb{K}^n amb les operacions que s'estan utilitzant a aquest curs són un \mathbb{K} -espai vectorial.

- Les matrius $M_{m \times n}(\mathbb{K})$, amb la suma coeficient a coeficient i la multiplicació per escalar (que afecta a tots els coeficients) també és un \mathbb{K} -espai vectorial.
- Els polinomis de $\mathbb{K}[x]$ (no limitem el grau!).
- Si A és un conjunt i \mathbb{K} un cos, considerem $\text{Map}(A, \mathbb{K})$ el conjunt de les aplicacions de conjunts $f: A \rightarrow \mathbb{K}$. Si $f, g \in \text{Map}(A, \mathbb{K})$, definim $(f + g)(x) = f(x) + g(x)$ i $(\lambda f)(x) = \lambda f(x)$ (per a tot $\lambda \in \mathbb{K}$), llavors $\text{Map}(A, \mathbb{K})$ és un \mathbb{K} -espai vectorial (veiem que poder sumar funcions i multiplicar-les per escalar depèn de que ho puguem fer al cos \mathbb{K}). El vector zero és la funció constant $f(a) = 0$ per a tot $a \in A$.
- Si $A = \mathbb{N}$, podem identificar $f \in \text{Map}(\mathbb{N}, \mathbb{K})$ com la successió d'escalars (x_0, x_1, \dots) amb $x_i = f(i) \in \mathbb{K}$.
- Si considerem les funcions contínues $f: \mathbb{R} \rightarrow \mathbb{R}$ amb l'estructura de l'apartat anterior, són un espai vectorial (ja que la suma de contínues és contínua i la multiplicació d'un escalar per una funció contínua és una funció contínua).
- L'espai d'equacions lineals amb n incògnites:

$$\{a_1x_1 + a_2x_2 + \dots + a_nx_n = b : a_i, b \in \mathbb{K}\},$$

amb la suma d'equacions variable a variable (i terme independent més terme independent), i producte per escalar que multiplica tota l'equació. El vector zero seria l'equació $0 = 0$.

La resta d'aquesta secció la dedicarem a generalitzar conceptes que ja hem vist per \mathbb{K}^n , a espais vectorials abstractes.

Definició 3.73. Un subconjunt $W \subseteq V$ d'un \mathbb{K} -espai vectorial V és un *subespai* de V si:

1. El vector 0_V pertany a W .
2. Si w_1 i w_2 són elements de W , aleshores la seva suma $w_1 + w_2$ també és a W (direm que W és tancat per la suma).
3. Si w és de W i $\lambda \in \mathbb{K}$, aleshores λw també pertany a W (direm que W és tancat pel producte per escalars).

Observació 3.74. Notem que un subespai W d'un \mathbb{K} -espai vectorial V és un \mathbb{K} -espai vectorial per si sol (amb la suma i producte per escalars heretat de V).

Definició 3.75. Considerem un subconjunt F d'un \mathbb{K} -espai vectorial V .

1. Diem que F genera V si tot $v \in V$ es pot escriure com a combinació lineal d'alguns elements v_1, \dots, v_n de F .
2. Diem que F és un conjunt *linealment dependent* si algun $v \in F$ es pot escriure com a combinació lineal d'altres. Diem que F és *linealment independent* si no és linealment dependent. De manera equivalent, el conjunt F és linealment independent si l'única manera d'obtenir $\vec{0}$ com a combinació lineal $\lambda_1 v_1 + \dots + \lambda_n v_n$ amb $v_1, \dots, v_n \in F$ és amb $\lambda_1 = \dots = \lambda_n = 0$.
3. El conjunt F és una *base* si genera V i és linealment independent.

Observació 3.76. Hi ha espais vectorials que tenen bases amb infinits elements. Per exemple, l'espai vectorial $\mathbb{K}[x]$ format pels polinomis amb coeficients a \mathbb{K} té base

$$[1, x, x^2, x^3, x^4, \dots].$$

Definició 3.77. Direm que un \mathbb{K} -espai vectorial V té *dimensió finita* si té una base $F = [v_1, \dots, v_n]$ formada per un nombre finit de vectors. En cas contrari, diem que V té *dimensió infinita*.

De la mateixa manera que hem demostrat el Teorema 3.41, si V té dimensió finita aleshores totes les bases tenen el mateix nombre d'elements. En aquest cas, direm que la *dimensió de V* és el nombre d'elements d'una base qualsevol.

Exemple 3.78. Considerem l'espai vectorial $\mathcal{C}^\infty(\mathbb{R})$ format per les funcions diferenciables. Considerem el subconjunt:

$$V = \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid f'' + f = 0\}.$$

Demostrem primer que $V \subset \mathcal{C}^\infty(\mathbb{R})$ és un subespai vectorial:

- La funció $f = 0$ compleix l'equació, ja que $f = f' = f'' = 0$ i per tant $f'' + f = 0$.
- Si $f \in V$ i $g \in V$, tenim $f'' + f = 0$ i $g'' + g = 0$. Tenim que $(f + g)'' = f'' + g''$, i per tant:

$$(f + g)'' + (f + g) = f'' + g'' + f + g = f'' + f + g'' + g = 0 + 0 = 0,$$

per tant $f + g \in V$.

- Si $f \in V$ i $\lambda \in \mathbb{R}$, tenim $(\lambda f)'' = \lambda f''$, per tant:

$$(\lambda f)'' + \lambda f = \lambda f'' + \lambda f = \lambda(f'' + f) = \lambda \cdot 0 = 0,$$

i per tant $\lambda f \in V$.

Intentem calcular la dimensió de V : del curs de càlcul sabem que les funcions sin i cos compleixen aquesta equació, i com que no són múltiple una de l'altra, $\dim(V) \geq 2$.

Suposem que tenim tres funcions $f, g, h \in V$ i considerem la matriu (on els coeficients són funcions!):

$$\begin{pmatrix} f & g & h \\ f' & g' & h' \\ f'' & g'' & h'' \end{pmatrix}$$

I, si apliquem la condició $f'' = -f$, $g'' = -g$, $h'' = -h$ i en calculem el rang tenim:

$$\text{Rang} \begin{pmatrix} f & g & h \\ f' & g' & h' \\ f'' & g'' & h'' \end{pmatrix} = \text{Rang} \begin{pmatrix} f & g & h \\ f' & g' & h' \\ -f & -g & -h \end{pmatrix} = \text{Rang} \begin{pmatrix} f & g & h \\ f' & g' & h' \\ 0 & 0 & 0 \end{pmatrix} \leq 2,$$

on a la segona igualtat hem sumat la primera fila a la tercera.

Per tant hi ha una combinació lineal entre els vectors $\begin{pmatrix} f \\ f' \\ f'' \end{pmatrix}$, $\begin{pmatrix} g \\ g' \\ g'' \end{pmatrix}$ i $\begin{pmatrix} h \\ h' \\ h'' \end{pmatrix}$, en particular entre f , g i h , pel que no poden ser linealment independents i $\dim(V) < 3$.

Per tant, hem demostrat que $\dim(V) = 2$ i una possible base és $[\sin, \cos]$.

Una demostració més explícita del fet que sin i cos generen V és la següent: demostrem primer que, si $g \in V$, aleshores

$$C_g = g^2 + (g')^2$$

és una funció constant. Això es veu calculant la derivada, i veient que és zero:

$$\frac{d}{dx}(g^2 + (g')^2) = 2gg' + 2g'g'' = 2gg' - 2g'g = 0.$$

Per tant, $C_g = g(0)^2 + g'(0)^2$, per exemple. Donada $f \in V$, apliquem aquest resultat a la funció

$$x \mapsto g(x) = f(x) - f(0) \cos(x) - f'(0) \sin(x),$$

que és de V si assumim que f ho és. Aleshores $C_g = g(0)^2 + g'(0)^2 = 0 + 0 = 0$, i per tant $(g(x))^2 + (g'(x))^2 = 0$ per tot x , i com que és la suma de dos quadrats a \mathbb{R} , obtenim que tant $g(x) = 0$ com $g'(x) = 0$ per tot $x \in \mathbb{R}$. D'aquí es dedueix que $f(x) = f(0) \cos(x) + f'(0) \sin(x)$.

Definició 3.79. Sigui V un \mathbb{K} -espai vectorial de dimensió finita n , i sigui $\mathcal{B} = [v_1, \dots, v_n]$ una base ordenada. Aleshores tot element $v \in V$ es pot escriure de manera única com

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Definim les *coordenades de v en la base \mathcal{B}* com

$$[v]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

A continuació introduïrem el concepte d'aplicació lineal entre espais vectorials sobre el mateix cos \mathbb{K} .

Definició 3.80. Una aplicació $f: V \rightarrow W$ entre dos \mathbb{K} -espais vectorials V i W és una *aplicació lineal* si:

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$ per a tot $v_1, v_2 \in V$, i
2. $f(\lambda v) = \lambda f(v)$ per a tot $v \in V$ i tot $\lambda \in \mathbb{K}$.

Exemple 3.81. Si V és un \mathbb{K} -espai vectorial de dimensió finita i \mathcal{B} és una base, l'aplicació $[\cdot]_{\mathcal{B}}: V \rightarrow \mathbb{K}^n$ és una aplicació lineal.

Definició 3.82. Si $f: V \rightarrow W$ és una aplicació lineal entre \mathbb{K} -espais vectorials V i W , definim el *nucli* i la *imatge* com els subespais

$$\ker(f) = \{v \in V : f(v) = \vec{0}\} \subset V,$$

$$\operatorname{Im}(f) = \{f(v) : v \in V\} \subset W.$$

Exemple 3.83. Considerem l'espai vectorial $\mathcal{C}^\infty(\mathbb{R})$ format per les funcions diferenciables. L'aplicació $D: \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$, definida com $D(f) = f'$ (la derivada) és una aplicació lineal. El nucli de D és

$$\ker(D) = \{f : f' = 0\} = \{\text{funcions constants}\},$$

que és un subespai de dimensió 1. Donada una funció $g \in \mathcal{C}^\infty(\mathbb{R})$, el teorema fonamental del càlcul ens diu que podem trobar una funció G (una primitiva de g) tal que $G' = g$. Per tant $\operatorname{Im}(D) = \mathcal{C}^\infty(\mathbb{R})$.

Si $f: V \rightarrow W$ és bijectiva (equivalentment, $\ker f = \{\vec{0}\}$ i $\operatorname{Im}(f) = W$), direm que f és invertible, o que és un *isomorfisme*. També direm que els espais V i W són *isomòrfics*, i escriurem $V \cong W$.

Proposició 3.85. *Siguin V i W dos \mathbb{K} -espais vectorials de dimensió finita. Aleshores:*

- Observació 3.86.* La hipòtesi de dimensió finita és necessària: a l'Exemple 3.83 hem vist com la derivada no és un isomorfisme (perquè té nucli no zero), però en canvi és exhaustiu.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ [\cdot]_{\mathcal{B}} \downarrow \cong & & \cong \downarrow [\cdot]_{\mathcal{C}} \\ \mathbb{K}^m & \dashrightarrow & \mathbb{K}^n \end{array} \quad (4)$$

Fixem-nos que l'aplicació resultant de resseguir el diagrama d'esquerra a dreta és una aplicació lineal de \mathbb{K}^m a \mathbb{K}^n i, per tant, té una matriu associada. La columna i -èssima d'aquesta matriu és $[f(v_i)]_{\mathcal{C}}$.

$$[f]_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} | & \cdots & | \\ [f(v_1)]_{\mathcal{C}} & \cdots & [f(v_m)]_{\mathcal{C}} \\ | & \cdots & | \end{pmatrix}$$

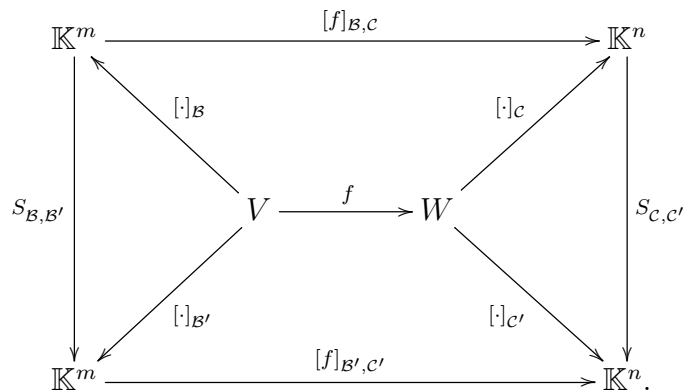
Si l'espai de sortida és el mateix que el d'arribada, i escrivim la mateixa base tant per la sortida com l'arribada, escrivim $[f]_{\mathcal{B}}$ enlloc de $[f]_{\mathcal{B}, \mathcal{B}}$.

Un cas particular interessant es dona quan considerem l'aplicació identitat:

$$S_{\mathcal{B}, \mathcal{C}} = [\text{Id}]_{\mathcal{B}, \mathcal{C}}.$$

El cas més general que podem considerar és el següent. Siguin V i W com fins ara, i prenem bases \mathcal{B} i \mathcal{B}' de V , i bases \mathcal{C} i \mathcal{C}' de W . Aleshores enganxant quatre còpies del Diagrama (4)

obtenim el diagrama següent:



Observem que les quatre fletxes exteriors tenen matrius associades, però que les cinc fletxes interiors no en tenen pas. En tot cas, les fletxes exteriors ens donen la següent fórmula, que ens relaciona la matriu de f en les bases \mathcal{B}, \mathcal{C} i la matriu de f en les bases $\mathcal{B}', \mathcal{C}'$.

$$[f]_{\mathcal{B}', \mathcal{C}'} S_{\mathcal{B}, \mathcal{B}'} = S_{\mathcal{C}, \mathcal{C}'} [f]_{\mathcal{B}, \mathcal{C}}.$$

Exemple 3.89. Considerem $\mathbb{R}^3[x]_{\leq 2}$, els polinomis de grau més petit o igual que 2 com a \mathbb{R} -espai vectorial amb la base $\mathcal{B} = [1, x, x^2]$.

Per $a \in \mathbb{R}$, considerem l'aplicació lineal:

$$\begin{aligned} f_a: \mathbb{R}[x]_{\leq 2} &\longrightarrow \mathbb{R}^3 \\ p(x) &\mapsto \begin{pmatrix} p(a) \\ p'(a) \\ p''(a) \end{pmatrix} \end{aligned}$$

Considerem les imatges dels elements de \mathcal{B} :

$$f_a(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, f_a(x) = \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} \text{ i } f_a(x^2) = \begin{pmatrix} a^2 \\ 2a \\ 2 \end{pmatrix}.$$

Per tant, si $\mathcal{C} = [\vec{e}_1, \vec{e}_2, \vec{e}_3]$ és la base d' \mathbb{R}^3 formada pels vectors estàndard, tenim:

$$[f_a]_{\mathcal{B}, \mathcal{C}} = \begin{pmatrix} 1 & a & a^2 \\ 0 & 1 & 2a \\ 0 & 0 & 2 \end{pmatrix}.$$

En particular, deduïm que f_a és un isomorfisme (la matriu $[f_a]_{\mathcal{B}, \mathcal{C}}$ és invertible).

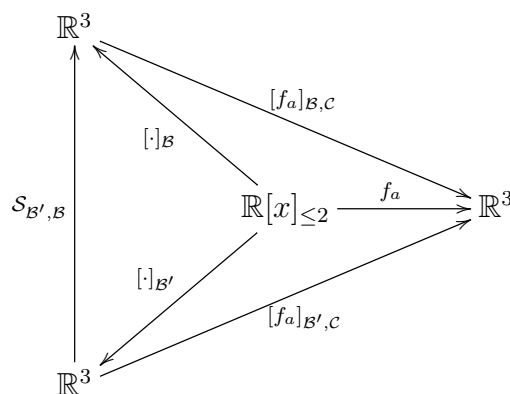
Considerem ara la base de $\mathbb{R}[x]_{\leq 2}$ formada pels vectors $\mathcal{B}' = [1, (x-a), (x-a)^2]$. En aquest cas:

$$f_a(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, f_a(x-a) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ i } f_a((x-a)^2) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

I per tant:

$$[f_a]_{\mathcal{B}', \mathcal{C}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Llavors, el diagrama que tenim és:



On

$$\mathcal{S}_{\mathcal{B}', \mathcal{B}} = \begin{pmatrix} 1 & -a & a^2 \\ 0 & 1 & -2a \\ 0 & 0 & 1 \end{pmatrix}.$$

Podem comprovar que es compleix:

$$[f_a]_{\mathcal{B}', \mathcal{C}} = [f_a]_{\mathcal{B}, \mathcal{C}} \mathcal{S}_{\mathcal{B}', \mathcal{B}},$$

és a dir,

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

Exercicis recomanats

Els exercicis que segueixen són útils per practicar el material presentat. La numeració és la de [1].

Secció 2.1: 6, 8, 24, 26, 28, 30.

Secció 2.2: 2, 6, 10, 20.

Secció 2.3: 4, 12, 20.

Secció 2.4: 4, 20, 28.

Secció 3.1: 10, 16, 22.

Secció 3.2: 6, 18, 24, 28.

Secció 3.3: 16, 22, 32.

Secció 3.4: 8, 22, 38.

Secció 4.1: 1-5, 20, 26.

Secció 4.2: 6, 10, 58, 66, 68, 74.

Secció 4.3: 4, 14, 52, 54, 60.

4 Diagonalització

El contingut d'aquesta secció és pot trobar a [1, Temes 6, 7] i a [2, Tema 4].

4.1 Motivació

Considerem les matrius següents:

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ i } B = \begin{pmatrix} 5 & -15 & -21 \\ -3 & 9 & 13 \\ 3 & -9 & -13 \end{pmatrix}$$

i suposem que volem calcular A^5 , $\text{Rang}(A)$, $\text{Ker}(f_A)$ o una base de $\text{Im}(f_A)$; i exactament els mateixos càlculs per B .

Per A , aquests càlculs són quasi immediats:

$$A^5 = \begin{pmatrix} (-1)^5 & 0 & 0 \\ 0 & 2^5 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\text{Rang}(A) = 2$, $\text{Ker}(f_A) = \langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle$ i una base de $\text{Im}(f_A)$ pot ser $\mathcal{B} = (\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix})$.

Per la matriu B , podem fer B^5 , però ens porta més càlculs, igual que calcular-ne el rang, el nucli de f_B o una base de la imatge de f_B .

Observem ara que hi ha una relació entre A i B :

$$B = SAS^{-1}$$

on

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

i això permet aprofitar els càlculs que hem fet per A per deduir els de B :

$$B^5 = SAS^{-1}SAS^{-1} \dots SAS^{-1} = SA^5S^{-1},$$

tenim la igualtat $\text{Rang}(B) = \text{Rang}(A) = 2$, $\text{Ker}(f_B) = S \text{Ker}(f_A)$ i una base de $\text{Im}(f_B)$ s'aconsegueix amb la base de $\text{Im}(f_A)$ multiplicada (per l'esquerra) per la matriu S^{-1} .

A aquest exemple, el que hem vist és que encara que les matrius A i B corresponen a una mateixa transformació lineal f expressada en dues bases diferents, els càlculs han quedat molt més fàcils amb la matriu A pel simple fet de ser una matriu diagonal. A aquest capítol, l'objectiu principal és: donada una aplicació lineal $f: E \rightarrow E$ (amb E un espai vectorial de dimensió finita), trobar una base \mathcal{B} d' E tal que $[f]_{\mathcal{B}}$ sigui una matriu diagonal. També veurem que, a vegades, no existeix cap base amb aquesta propietat.

4.2 Determinants

Probablement ja coneixem el determinant de matrius 2×2 i 3×3 definits directament com:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (5)$$

i

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \quad (6)$$

Una de les propietats principals és la següent: el determinant val zero si i només si els vectors columna de la matriu són linealment dependents.

Si pensem les columnes de les matrius com a vectors de \mathbb{K}^n , podem considerar el determinant com una aplicació:

$$\begin{aligned} \det: M_n(\mathbb{K}) = (\mathbb{K}^n)^n &\longrightarrow \mathbb{K} \\ (\vec{v}_1, \dots, \vec{v}_n) &\mapsto \det(\vec{v}_1, \dots, \vec{v}_n) \end{aligned}$$

i habitualment escriurem:

$$\det(\vec{v}_1, \dots, \vec{v}_n) = \begin{vmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{vmatrix}$$

Demanarem que el determinant compleixi les propietats següents:

D1. $\det(\vec{v}_1, \dots, \lambda \vec{v}_j, \dots, \vec{v}_n) = \lambda \det(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_n)$ per a tots $\vec{v}_j \in \mathbb{K}^n$ i $\lambda \in \mathbb{K}$.

D2. $\det(\vec{v}_1, \dots, \vec{v}_{j-1}, \vec{v}_j + \vec{w}, \vec{v}_{j+1}, \dots, \vec{v}_n) = \det(\vec{v}_1, \dots, \vec{v}_{j-1}, \vec{v}_j, \vec{v}_{j+1}, \dots, \vec{v}_n) + \det(\vec{v}_1, \dots, \vec{v}_{j-1}, \vec{w}, \vec{v}_{j+1}, \dots, \vec{v}_n)$ per a tots \vec{v}_k i $\vec{w} \in \mathbb{K}^n$.

D3. $\det(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_k, \dots, \vec{v}_n) = 0$ si $\vec{v}_j = \vec{v}_k$ amb $j \neq k$.

D4. $\det(\mathbf{1}_n) = \det(\vec{e}_1, \dots, \vec{e}_n) = 1$, on \vec{e}_j són els vectors estàndard.

Exercici 4.1. Demostreu que els determinants de matrius 2×2 i 3×3 definits a les Equacions (5) i (6) respectivament compleixen aquestes propietats i que són les úniques aplicacions de $M_2(\mathbb{K}) \rightarrow \mathbb{K}$ i $M_3(\mathbb{K}) \rightarrow \mathbb{K}$ respectivament que les compleixen.

Observació 4.1. Observem que **D1** implica que $\det(\vec{v}_1, \dots, \vec{0}, \dots, \vec{v}_n) = 0$: com que $\vec{0} = 0 \cdot \vec{0}$:

$$\begin{aligned} \det(\vec{v}_1, \dots, \vec{0}, \dots, \vec{v}_n) &= \det(\vec{v}_1, \dots, 0 \cdot \vec{0}, \dots, \vec{v}_n) = \\ &= 0 \det(\vec{v}_1, \dots, \vec{0}, \dots, \vec{v}_n) = 0 \end{aligned}$$

Lema 4.2. Si considerem les transformacions elementals per columnes corresponents a les transformacions **T1**, **T2** i **T3** definides a la Subsecció 2.2, el determinant es modifica com:

- **T1:** Si B és la matriu resultant de multiplicar una columna per λ a la matriu A :

$$\det(B) = \lambda \det(A).$$

- **T2:** Si B és la matriu resultant de sumar a una columna d' A μ vegades una altra columna d' A :

$$\det(B) = \det(A).$$

- **T3:** Si B és la matriu resultant d'intercanviar dues columnes diferents d' A , llavors:

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

Demostració. **D1** ens diu com es transforma \det per la transformació elemental **T1**, obtenint el resultat de l'enunciat.

Estudiem ara el canvi **T2**: sumem a una columna μ vegades una altra:

$$\begin{aligned} \det(\dots, \vec{v}_j, \dots, \vec{v}_k + \mu \vec{v}_j, \dots) &= \det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots) + \det(\dots, \vec{v}_j, \dots, \mu \vec{v}_j, \dots) = \\ &= \det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots) + \mu \det(\dots, \vec{v}_j, \dots, \vec{v}_j, \dots) = \\ &= \det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots) + 0 \end{aligned}$$

on primer hem aplicat **D2** (separar la suma de vectors), llavors **D1** (treure el μ fora del determinant) i finalment **D3** per veure que un dels determinants és zero.

Mirem com es modifica el determinant per la transformació **T3**: vegem que si intercanviem dues columnes $j \neq k$, llavors

$$\det(\vec{v}_1, \dots, \vec{v}_j, \dots, \vec{v}_k, \dots, \vec{v}_n) = -\det(\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_j, \dots, \vec{v}_n).$$

Per tal de simplificar la notació escrivim $\det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots)$:

$$\begin{aligned} 0 &= \det(\dots, \vec{v}_j + \vec{v}_k, \dots, \vec{v}_j + \vec{v}_k, \dots) = \\ &= \det(\dots, \vec{v}_j + \vec{v}_k, \dots, \vec{v}_j, \dots) + \det(\dots, \vec{v}_j + \vec{v}_k, \dots, \vec{v}_k, \dots) = \\ &= \det(\dots, \vec{v}_j, \dots, \vec{v}_j, \dots) + \det(\dots, \vec{v}_k, \dots, \vec{v}_j, \dots) + \\ &\quad + \det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots) + \det(\dots, \vec{v}_k, \dots, \vec{v}_k, \dots) = \\ &= 0 + \det(\dots, \vec{v}_k, \dots, \vec{v}_j, \dots) + \det(\dots, \vec{v}_j, \dots, \vec{v}_k, \dots) + 0 \end{aligned}$$

□

Teorema 4.3. Si tenim dues aplicacions $\det: (\mathbb{K}^n)^n \rightarrow \mathbb{K}$ i $\det': (\mathbb{K}^n)^n \rightarrow \mathbb{K}$ complint **D1**, **D2**, **D3** i **D4**, llavors $\det = \det'$.

Demostració. Volem veure que si $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{K}^n$, llavors $\det(\vec{v}_1, \dots, \vec{v}_n) = \det'(\vec{v}_1, \dots, \vec{v}_n)$: si considerem A la matriu formada pels vectors $\vec{v}_1, \dots, \vec{v}_n$ per columna, podem aplicar transformacions elementals per columnes fins a tenir, o bé una matriu identitat, o bé una matriu amb l'última columna tot zeros. Pel Lema 4.2, veiem com es modifica qualsevol aplicació (tant \det , com \det') que compleixi els axiomes **D1**, **D2**, **D3** i **D4**, obtenint que $\det(A) = \lambda \det(\text{rcef}(A))$ (*reduced column echelon form*) i $\det'(A) = \lambda \det'(\text{rcef}(A))$ (amb el mateix λ), però o bé $\text{rcef}(A) = \mathbf{1}_n$ (i tenim $\det(\text{rcef}(A)) = 1 = \det'(\text{rcef}(A))$ per **D4**), o bé $\text{rcef}(A)$ té una columna tot zeros i per la Observació 4.1, $\det(A) = 0 = \det'(A)$.

Per tant, com que \det i \det' estan determinats pels canvis elementals per columnes i pel seu valor a la identitat o a una matriu amb una columna tot zeros, han de valer el mateix. □

Exemple 4.4. Calculem el determinant d' A , on:

$$A = \begin{pmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 5 & 6 & 0 \end{pmatrix}$$

fent transformacions elementals:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 5 & 6 & 0 \end{vmatrix} &= \begin{vmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 0 & -4 & -30 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 6 \\ 0 & 1 & 8 \\ 0 & -4 & -30 \end{vmatrix} = \\ &= - \begin{vmatrix} 1 & 0 & 6 \\ 0 & 1 & 8 \\ 0 & 0 & 2 \end{vmatrix} = -2 \begin{vmatrix} 1 & 0 & 6 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{vmatrix} = -2 \det(\mathbf{1}_3) = -2. \end{aligned}$$

Amb tot això, el que no hem demostrat és que existeixi una aplicació \det que tingui les propietats **D1**, **D2**, **D3** i **D4**.

Per demostrar la existència, el que farem és construir-la explícitament.

Considerem $A \in M_n(\mathbb{K})$ i el producte de n coeficients de la matriu $a_{i_1 j_1} \cdots a_{i_n j_n}$ tals que no hi hagi dos coeficients d'una mateixa columna, ni d'una mateixa fila. Dit d'una altra manera, com que hi ha n files i n columnes, considerem n parelles $\{(i_1, j_1), \dots, (i_n, j_n)\}$ que compleixin la igualtat de conjunts:

$$\{i_1, \dots, i_n\} = \{1, \dots, n\} = \{j_1, \dots, j_n\}. \quad (7)$$

Definició 4.5. Anomenem un *patró* d' n elements a un conjunt $\mathcal{P} = \{(i_1, j_1), \dots, (i_n, j_n)\} \subset \{1, \dots, n\}^2$ que compleixi que $i_k \neq i_l$ i $j_k \neq j_l$ si $k \neq l$, o, equivalentment, que compleixi l'Equació (7).

Donat un patró d' n elements \mathcal{P} i una matriu $A \in M_n(\mathbb{K})$, definim l'element $a_{\mathcal{P}}$ com el producte:

$$a_{\mathcal{P}} = \prod_{(i,j) \in \mathcal{P}} a_{ij}.$$

Exemple 4.6. En el cas de les matrius 3×3 , tenim 6 patrons possibles i ens donen els productes:

$$\begin{aligned} \begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & \boxed{a_{22}} & a_{23} \\ a_{31} & a_{32} & \boxed{a_{33}} \end{pmatrix} &\iff \mathcal{P} = \{(1, 1), (2, 2), (3, 3)\} \iff a_{\mathcal{P}} = a_{11}a_{22}a_{33} \\ \begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & a_{22} & \boxed{a_{23}} \\ a_{31} & \boxed{a_{32}} & a_{33} \end{pmatrix} &\iff \mathcal{P} = \{(1, 1), (2, 3), (3, 2)\} \iff a_{\mathcal{P}} = a_{11}a_{23}a_{32} \\ \begin{pmatrix} \boxed{a_{11}} & \boxed{a_{12}} & a_{13} \\ \boxed{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & \boxed{a_{33}} \end{pmatrix} &\iff \mathcal{P} = \{(1, 2), (2, 1), (3, 3)\} \iff a_{\mathcal{P}} = a_{12}a_{21}a_{33} \\ \begin{pmatrix} \boxed{a_{11}} & \boxed{a_{12}} & a_{13} \\ a_{21} & a_{22} & \boxed{a_{23}} \\ \boxed{a_{31}} & a_{32} & a_{33} \end{pmatrix} &\iff \mathcal{P} = \{(1, 2), (2, 3), (3, 1)\} \iff a_{\mathcal{P}} = a_{12}a_{23}a_{31} \\ \begin{pmatrix} \boxed{a_{11}} & a_{12} & \boxed{a_{13}} \\ \boxed{a_{21}} & a_{22} & a_{23} \\ a_{31} & \boxed{a_{32}} & \boxed{a_{33}} \end{pmatrix} &\iff \mathcal{P} = \{(1, 3), (2, 1), (3, 2)\} \iff a_{\mathcal{P}} = a_{13}a_{21}a_{32} \\ \begin{pmatrix} \boxed{a_{11}} & a_{12} & \boxed{a_{13}} \\ a_{21} & \boxed{a_{22}} & a_{23} \\ \boxed{a_{31}} & a_{32} & a_{33} \end{pmatrix} &\iff \mathcal{P} = \{(1, 3), (2, 2), (3, 1)\} \iff a_{\mathcal{P}} = a_{13}a_{22}a_{31} \end{aligned}$$

I podeu veure que coincideixen amb els sumands de l'Equació (6). Ara tant sols cal decidir si cada element suma o resta, i per això necessitem el *signe d'un patró*:

Definició 4.7. Fixat un enter positiu n i un patró $\mathcal{P} = \{(i_1, j_1), \dots, (i_n, j_n)\} \subset \{1, \dots, n\}^2$, definim $\text{Sign}(\mathcal{P})$, el *signe de \mathcal{P}* , com $(-1)^\epsilon$, on ϵ és el nombre de parelles $[(i, j), (i', j')] \in \mathcal{P}$ tals que $i < i'$ i $j > j'$.

Exemple 4.8. En el cas 3×3 , els patrons tenen el signe següent:

Patró	Parelles	ϵ	Signe
$\{(1, 1), (2, 2), (3, 3)\}$		0	$(-1)^0 = 1$
$\{(1, 1), (2, 3), (3, 2)\}$	$\{[(2, 3), (3, 2)]\}$	1	$(-1)^1 = -1$
$\{(1, 2), (2, 1), (3, 3)\}$	$\{[(1, 2), (2, 1)]\}$	1	$(-1)^1 = -1$
$\{(1, 2), (2, 3), (3, 1)\}$	$\{[(1, 2), (3, 1)], [(2, 3), (3, 1)]\}$	2	$(-1)^2 = 1$
$\{(1, 3), (2, 1), (3, 2)\}$	$\{[(1, 3), (2, 1)], [(1, 3), (3, 2)]\}$	2	$(-1)^2 = 1$
$\{(1, 3), (2, 2), (3, 1)\}$	$\{[(1, 3), (2, 2)], [(1, 3), (3, 1)], [(2, 2), (3, 1)]\}$	3	$(-1)^3 = -1$

Els patrons tenen les propietats següents:

Lema 4.9. Considerem $A \in M_n(\mathbb{K})$, amb n fixada.

- (a) Cada patró \mathcal{P} es correspon amb una aplicació bijectiva $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ (s'anomena permutació).

(b) Hi ha $n!$ (factorial d' n) patrons diferents a A .

(c) Definim el signe d'una permutació $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ com

$$\text{Sign}(\sigma) = \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Si σ és la permutació corresponent a \mathcal{P} , llavors $\text{Sign}(\sigma) = \text{Sign}(\mathcal{P})$.

(d) Si $\text{Id}: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ és l'aplicació $\text{Id}(i) = i$ (identitat), llavors $\text{Sign}(\text{Id}) = 1$.

(e) Si σ i τ són dues permutacions de $\{1, \dots, n\}$, llavors

$$\text{Sign}(\sigma \circ \tau) = \text{Sign}(\sigma) \text{Sign}(\tau)$$

(f) $\text{Sign}(\sigma) = \text{Sign}(\sigma^{-1})$.

(g) Si $\tau_{k\ell}$ correspon a la permutació:

$$\tau_{k\ell}(m) = \begin{cases} m & \text{si } m \notin \{k, \ell\} \\ \ell & \text{si } m = k \\ k & \text{si } m = \ell \end{cases}$$

amb $k \neq \ell$, llavors $\text{Sign}(\tau_{k\ell}) = -1$.

Demostració. Els patró \mathcal{P} té n parelles (i, j) on, si mirem tant sols la primera coordenada, hi ha exactament un 1, un 2, ... i un n . Definim $\sigma(i) = j$ si $(i, j) \in \mathcal{P}$. Aquesta aplicació és injectiva i exhaustiva perquè a la segona coordenada de les parelles de \mathcal{P} també hi ha un únic 1, un únic 2, ... i un únic n .

Si tenim una aplicació bijectiva $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$, definim els patró corresponent com $\mathcal{P} = \{(1, \sigma(1)), \dots, (n, \sigma(n))\}$.

Per demostrar (b) comptem quantes aplicacions bijectives σ hi ha de $\{1, 2, \dots, n\}$ en ell mateix: $\sigma(1)$ pot ser qualsevol element de $\{1, 2, \dots, n\}$, per tant en podem escollir n ; $\sigma(2)$ pot ser qualsevol element de $\{1, 2, \dots, n\}$ excepte $\sigma(1)$, per tant en podem escollir $n - 1$; iterant aquest procediment, tindrem $n(n - 1)(n - 2) \cdots 2 \cdot 1$ aplicacions bijectives, i aquesta és la definició de factorial d' n .

A l'expressió de (c) veiem que tant al denominador hi ha el producte $(2 - 1)(3 - 1) \cdots (n - 1)(3 - 2) \cdots$, mentre que al numerador hi ha els mateixos factors, on es canvien de signe els que compleixen $\sigma(i) > \sigma(j)$ amb $i < j$, pel que el quocient serà ± 1 , i el signe ve determinat pel nombre de vegades que passa $\sigma(i) > \sigma(j)$ amb $i < j$, que és la definició de signe d'un patró.

Calculem el signe de la identitat per demostrar (d):

$$\text{Sign}(\text{Id}) = \prod_{i < j} \frac{\text{Id}(j) - \text{Id}(i)}{j - i} = \prod_{i < j} \frac{j - i}{j - i} = 1.$$

La demostració d'(e) ve de considerar:

$$\begin{aligned} \text{Sign}(\sigma \circ \tau) &= \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} = \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \frac{\tau(j) - \tau(i)}{j - i} = \\ &= \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \prod_{i < j} \frac{\tau(j) - \tau(i)}{j - i} = \text{Sign}(\sigma) \text{Sign}(\tau), \end{aligned}$$

on observem que:

$$\prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i} = \prod_{i < j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}$$

ja que hi ha els mateixos factors al numerador i denominador, i si un factor del numerador ha canviat de signe, el corresponent factor del denominador també.

La demostració d'(f) es dedueix de que $1 = \text{Sign}(\text{Id}) = \text{Sign}(\sigma \circ \sigma^{-1}) = \text{Sign}(\sigma) \text{Sign}(\sigma^{-1})$, i que el signe de qualsevol permutació és ± 1 .

Finalment, la demostració de (g) és molt semblant a la de (d): primer considerem $k < \ell$ (com que $\tau_{k\ell} = \tau_\ell$, si cal, les intercanviem). Llavors el signe de $\tau_{k\ell}$ serà com el de Id , però hi haurà un factor diferent a quan $(i, j) = (k, \ell)$, que sortirà $\frac{k-\ell}{\ell-k} = -1$, per tant $\text{Sign}(\tau_{k\ell}) = -1$. \square

Ara ja podem definir el determinant d'una matriu:

Definició 4.10. Sigui $A \in M_n(\mathbb{K})$ una matriu quadrada. Considerem P_n el conjunt de tots els patrons d' n elements (P_n té $n!$ elements) i per cada $\mathcal{P} \in P_n$, considerem $a_{\mathcal{P}}$ com a la Definició 4.5. Definim el *determinant* d' A com:

$$\det(A) = \sum_{\mathcal{P} \in P_n} \text{Sign}(\mathcal{P}) a_{\mathcal{P}}. \quad (8)$$

Teorema 4.11. El determinant de la Definició 4.10 compleix les propietats **D1**, **D2**, **D3** i **D4**.

Demostració. Fixem $A \in M_n(\mathbb{K})$ i a_{ij} els seus coeficients.

D1: si fixem una columna j , cada sumand de l'Equació (8) té exactament un coeficient a_{ij} . Si el substituïm per λa_{ij} , tots els sumands de l'Equació (8) queden multiplicats per λ , pel que es compleix **D1**.

D2: si la columna j es pot escriure com la suma de dues columnes: $a_{ij} = a'_{ij} + a''_{ij}$, com que a l'Equació (8) cada sumand té exactament un d'aquests coeficients, podem separar la suma, obtenint **D2**.

D3: Si tenim la columna j i la columna k d' A que són iguals (amb $j \neq k$), tenim que cada sumand de l'Equació (8) apareix dues vegades, una pel patró \mathcal{P}_1 i \mathcal{P}_2 . Hem de veure que apareix amb signe diferent, i llavors la suma serà zero. Si σ_1 és la permutació que correspon al patró \mathcal{P}_1 i σ_2 la que correspon al patró \mathcal{P}_2 , llavors $\sigma_1 = \tau_{jk} \circ \sigma_2$ i, pel Lema 4.9, tenen signe diferent.

D4: l'únic patró \mathcal{P} tal que $(\mathbf{1}_n)_{\mathcal{P}} \neq 0$ és el patró $\{(1, 1), (2, 2), \dots, (n, n)\}$, i és un producte d'uns amb signe positiu. \square

Ara ja podem demostrar una de les propietats principals del determinant:

Vegem ara més propietats dels determinants:

Proposició 4.12. Si $A, B \in M_{n \times n}(\mathbb{K})$, llavors:

- (a) A és invertible si i només si $\det(A) \neq 0$.
- (b) $\det(A) \neq 0$ si i només si les files (i les columnes) d' A són linealment independents.
- (c) $\det(A) = \det(A^T)$, on A^T és la transposada d' A .
- (d) $\det(AB) = \det(A) \det(B)$.

Demostració. Per demostrar (a), recordem que una matriu és invertible si i només si podem aconseguir la identitat mitjançant transformacions elementals. En aquest cas, cada transformació modifica el determinant canviant-li el signe o multiplicant per un $\lambda \neq 0$. Llavors:

Si A és invertible, $\det(A) = \lambda \det(\mathbf{1}_n) = \lambda \neq 0$.

Si A no és invertible, $\det(A) = \lambda \det(B)$, on B és una matriu amb l'última fila tot zeros, per tant $\det(B) = 0$ (cada sumand del determinant a l'Equació (8) té un coeficient de l'última fila), d'on es dedueix que $\det(A) = 0$.

Per demostrar (b), el raonament de files (o columnes) linealment independents és el mateix, tenint en compte que les files (o columnes) són linealment independents si i només si la matriu A és equivalent a la identitat.

Per demostrar (c), observem que si $B = A^T$, per l'Equació (8), $\det(A)$ i $\det(B)$ tenen els mateixos sumands, pel que cal veure que els signes dels patrons \mathcal{P} i \mathcal{P}' corresponents a coeficients iguals $a_{\mathcal{P}}$ i $b_{\mathcal{P}'}$, són els mateixos: observem que si $a_{\mathcal{P}}$ és el coeficient corresponent a \mathcal{P} , que és un patró que correspon a una permutació σ (veure Lema 4.9), llavors, si \mathcal{P}' és el patró corresponent a σ^{-1} , tenim que $a_{\mathcal{P}} = b_{\mathcal{P}'}$, i amb el mateix signe, ja que pel Lema 4.9, $\text{Sign}(\mathcal{P}) = \text{Sign}(\sigma) = \text{Sign}(\sigma^{-1}) = \text{Sign}(\mathcal{P}')$.

Falta demostrar la fórmula del producte de determinants $\det(AB) = \det(A) \det(B)$: comencem amb la matriu A i li fem les transformacions elementals necessàries per obtenir $\text{rref}(A)$. Aquestes transformacions elementals aniran modificant el determinant, i obtindrem $\det(A) = \lambda \det(\text{rref}(A))$, on:

- Aquestes transformacions elementals es corresponen amb multiplicar per una matriu invertible P : $\text{rref}(A) = PA$.
- λ només depèn de les transformacions elementals que hem fet, per tant, si agafem $C \in M_n(\mathbb{K})$ i hi fem les mateixes transformacions que hem fet a A i obtenim $C' = PC$, tindrem $\det(C) = \lambda \det(C')$.
- Si A és invertible, llavors $\text{rref}(A) = PA = \mathbf{1}_n$ i $\det(A) = \lambda \det(\mathbf{1}_n) = \lambda$.

Considerem primer el cas en que A no és invertible: llavors $\det(A) = 0$ i les n columnes d' A no són linealment independents. Com que les columnes d' AB són combinació lineal de les de A i també en té n , tenim que les columnes de AB tampoc són linealment independents i per tant $\det(AB) = 0$ i en particular $\det(AB) = \det(A) \det(B)$.

Suposem ara que A és invertible, per tant $\text{rref}(A) = PA = \mathbf{1}_n$. Considerem ara $C = AB$ i fem les mateixes transformacions elementals que hem fet a A per obtenir $\text{rref}(A)$. Amb els raonaments que hem vist, tenim:

$$\det(AB) = \det(C) = \lambda \det(C') = \lambda \det(PAB) = \lambda \det(\mathbf{1}_n B) = \det(A) \det(B).$$

□

Ara veurem una manera inductiva de calcular el determinant. També es podia haver definit com veurem ara, i s'hauria de demostrar que compleix les propietats **D1**, **D2**, **D3** i **D4**. Per això, necessitem una notació que utilitzarem a aquesta secció (es pot confondre amb una de les notacions utilitzades a les seccions anteriors, fixeuvos amb la c del superíndex).

Notació 4.13. Si $A \in M_n(\mathbb{K})$, notem per $A_{ij}^c \in M_{(n-1)}(\mathbb{K})$ la matriu que resulta d'eliminar la fila i i la columna j d' A .

Proposició 4.14. Si $A \in M_{n \times n}(\mathbb{K})$, amb $n \geq 2$, podem desenvolupar el determinant per qualsevol fila o columna segons les fórmules següents:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}^c) \quad (\text{si desenvolupem per la fila } i),$$

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}^c) \quad (\text{si desenvolupem per la columna } j).$$

Demostració. Per a demostrar això, mirem com són tots els sumands i els comparem amb els de l'Equació (8): l'hem calculat de manera recursiva, i cada cop que fem una iteració anem esborrant la fila i columna corresponent a aquell coeficient. Per tant, hi haurà un coeficient de la primera fila, un altre de la segona, ..., de tal manera que cada cop agafem una columna diferent, i per tant tindrem el sumand:

$$(-1)^\epsilon a_{1j_1} a_{2j_2} \cdots a_{nj_n} \quad (9)$$

amb $j_k \neq j_l$ si $k \neq l$, i el signe ve determinat per la paritat de ϵ , que es pot veure que és la mateixa que la de la permutació, obtenint el mateix que a l'Equació (8)

Aquesta expressió no depèn de per quina fila o columna desenvolupem, pel que el resultat final serà el mateix. \square

Exemple 4.15. Calculem el determinant d' A , on:

$$A = \begin{pmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 5 & 6 & 0 \end{pmatrix}$$

desenvolupant per la primera fila:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 6 \\ 0 & -1 & -8 \\ 5 & 6 & 0 \end{vmatrix} &= 1 \begin{vmatrix} -1 & -8 \\ 6 & 0 \end{vmatrix} - 2 \begin{vmatrix} 0 & -8 \\ 5 & 0 \end{vmatrix} + 6 \begin{vmatrix} 0 & -1 \\ 5 & 6 \end{vmatrix} = \\ &= (0 - (-48)) - 2(40) + 6(5) = 48 - 80 + 30 = -2. \end{aligned}$$

Exercici 4.2. Si \mathcal{P} és un patró amb n elements i $A_{\mathcal{P}} \in M_n(\mathbb{K})$ és una matriu que té tots els coeficients zero, excepte els elements $a_{ij} = 1$, per a $(i, j) \in \mathcal{P}$, llavors $\det(A_{\mathcal{P}}) = \text{Sign}(\mathcal{P})$.

Exercici 4.3. Si $A \in M_n(\mathbb{K})$ és una matriu triangular superior (o inferior), llavors $\det(A)$ és el producte d'elements de la diagonal.

4.3 Polinomi característic. Valors i vectors propis

A la motivació d'aquest capítol (§4.1) hem vist una aplicació lineal $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ que escrita en una base \mathcal{B} o en una altra \mathcal{C} tenia les expressions següents:

$$[f]_{\mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{i} \quad [f]_{\mathcal{C}} = \begin{pmatrix} 5 & -15 & -21 \\ -3 & 9 & 13 \\ 3 & -9 & -13 \end{pmatrix}$$

Veiem que la primera és diagonal, mentre que la segona no. Això vol dir que l'aplicació lineal envia les rectes generades per cada vector de la base \mathcal{B} a elles mateixes, mentre que les rectes generades pels vectors de la base \mathcal{C} es converteixen en altres rectes. Si fixem uns eixos de coordenades amb la base \mathcal{B} , podrem veure com f modifica els vectors d' \mathbb{R}^3 molt més fàcilment que si fixem els eixos a la base \mathcal{C} . Per tant, estem dient que per la mateixa aplicació lineal $f: E \rightarrow E$, aprofitant que els coeficients de $[f]_{\mathcal{B}}$ depenen de la base \mathcal{B} , buscarem una base \mathcal{B} tal que la matriu $[f]_{\mathcal{B}}$ sigui el més “senzilla” possible.

Definició 4.16. Sigui $A \in M_n(\mathbb{K})$ i $f_A: \mathbb{K}^n \rightarrow \mathbb{K}^n$ l'aplicació lineal induïda. Diem que A és *diagonalitzable* si es compleix una de les condicions equivalents següents:

(a) A és similar a una matriu diagonal.

(b) Existeix una base \mathcal{B} de \mathbb{K}^n tal que $[f_A]_{\mathcal{B}}$ és diagonal.

(c) Existeix una matriu invertible $S \in M_n(\mathbb{K})$ tal que el producte $S^{-1}AS$ és diagonal.

Exemple 4.17. La matriu $B = \begin{pmatrix} 5 & -15 & -21 \\ -3 & 9 & 13 \\ 3 & -9 & -13 \end{pmatrix}$ és diagonalitzable (veure apartat 4.1).

Observació 4.18. Si una matriu A és diagonalitzable i $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ és una base en que $[f_A]_{\mathcal{B}}$ és diagonal amb coeficients a la diagonal $\lambda_1, \dots, \lambda_n$, llavors es compleix que $A\vec{v}_j = \lambda_j\vec{v}_j$. Per tant, els vectors \vec{v}_j de la base \mathcal{B} compleixen que $A\vec{v}_j$ és un múltiple de \vec{v}_j .

Utilitzem aquest fet per veure que no totes les matrius són diagonalitzables:

Exemple 4.19. La matriu $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$, que correspon a una rotació d'angle $\pi/2$, no és diagonalitzable: suposem que sí, i que \vec{v} és un vector d'una base \mathcal{B} tal que $[f_A]_{\mathcal{B}}$ és diagonal, llavors es complirà que $A\vec{v} = \lambda\vec{v}$ per a cert $\lambda \in \mathbb{R}$, però \vec{v} i $A\vec{v}$ són perpendiculars per a tot $\vec{v} \in \mathbb{R}^2$, pel que no pot ser. Una altra manera de veure-ho és que si $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ s'hauria de complir:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$$

I queda el sistema homogeni:

$$\begin{aligned} -\lambda x - y &= 0 \\ x - \lambda y &= 0 \end{aligned}$$

Com que estem treballant a \mathbb{R} , aquest sistema té rang 2 per qualsevol $\lambda \in \mathbb{R}$, i per tant l'única solució és $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, que no pot formar part de cap base.

Veiem, doncs, que per saber si una matriu $A \in M_n(\mathbb{K})$ és diagonalitzable, hem de veure si existeix una base $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ i escalars $\lambda_j \in \mathbb{K}$ tals que $A\vec{v}_j = \lambda_j\vec{v}_j$. Posem nom a aquests escalars i vectors:

Definició 4.20. Donada una matriu $A \in M_n(\mathbb{K})$, diem que un vector no nul $\vec{v} \in \mathbb{K}^n$ és un *vector propi* de *valor propi* $\lambda \in \mathbb{K}$ si $A\vec{v} = \lambda\vec{v}$.

Els elements del conjunt format pels $\lambda \in \mathbb{K}$ tals que existeix un vector $\vec{v} \in \mathbb{K}^n$ no nul tal que $A\vec{v} = \lambda\vec{v}$ s'anomenen *valors propis* d' A .

Exemple 4.21. Analitzem què és un vector propi de valor propi 0 d'una matriu $A \in M_n(\mathbb{K})$: serà $\vec{v} \neq \vec{0}$ tal que $A\vec{v} = \vec{0}$, per tant, serà un vector de $\text{Ker}(f_A)$ (les solucions del sistema homogeni amb matriu associada A). Deduïm que 0 és un valor propi d' A si i només si $\text{Ker}(f_A) \neq \{\vec{0}\}$, si i només si $\det(A) = 0$.

Exemple 4.22. Quins són els valors propis i els vectors propis de $\mathbf{1}_n$? Per a tot $\vec{v} \in \mathbb{K}^n$ es compleix que $\mathbf{1}_n\vec{v} = \vec{v}$, per tant, tot vector no nul és vector propi de valor propi 1.

Exemple 4.23. Considerem la matriu de la reflexió a \mathbb{R}^2 respecte la recta r que passa per l'origen amb vector director $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Segons la Proposició 3.8, correspon a la matriu:

$$\text{Refl} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Per definició de la reflexió, els vectors \vec{v} de la recta r compliran $A\vec{v} = \vec{v}$, per tant són valors propis de vector propi 1: $A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

En canvi, els vectors \vec{w} perpendiculars a r (els generats per $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$), compleixen que $A\vec{w} = -\vec{w}$, per tant, són vectors propis de valor propi -1 .

Això vol dir que, per a $\mathcal{B} = [\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}]$,

$$[\text{Refl}]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Mirant l'Exemple 4.22, podem deduir quan un valor $\lambda \in \mathbb{K}$ és un valor propi d'una matriu:

Teorema 4.24. *Fixada una matriu $A \in M_n(\mathbb{K})$, $\lambda \in \mathbb{K}$ és un valor propi de A si, i només si, $\det(A - \lambda \mathbf{1}_n) = 0$.*

Demostració. λ és un valor propi d' A si i només si existeix $\vec{v} \neq \vec{0}$ tal que $A\vec{v} = \lambda\vec{v}$, i aquesta igualtat és equivalent a que $(A - \lambda \mathbf{1}_n)\vec{v} = \vec{0}$. Per tant, és equivalent a que el sistema homogeni donat per la matriu $A - \lambda \mathbf{1}_n$ tingui solució diferent de $\vec{0}$, i això és equivalent a que $\det(A - \lambda \mathbf{1}_n) = 0$. \square

Això ens porta a la definició següent:

Definició 4.25. Si $A \in M_n(\mathbb{K})$, $\det(A - x \mathbf{1}_n)$ és un polinomi $p_A(x)$ de grau n en la variable x que s'anomena *polinomi característic d' A* .

El que hem vist al Teorema 4.24 és que λ és un valor propi d' A si i només si $p_A(\lambda) = 0$, on $p_A(x)$ és el polinomi característic d' A en la variable x .

Exemple 4.26. Calculem els valors propis de la matriu $B = \begin{pmatrix} 5 & -15 & -21 \\ -3 & 9 & 13 \\ 3 & -9 & -13 \end{pmatrix}$: hem de fer el determinant:

$$0 = \begin{vmatrix} 5-x & -15 & -21 \\ -3 & 9-x & 13 \\ 3 & -9 & -13-x \end{vmatrix} = -x^3 - x^2 + 2x = -x(x+1)(x-2)$$

i per tant els valors propis són $\{0, -1, 2\}$ (tal i com hem vist a l'apartat 4.1).

Observació 4.27. El Teorema 4.24 redueix el problema de trobar els valors propis d'una matriu a un de trobar les arrels d'un polinomi de grau n . Per a $n \leq 4$ existeixen fórmules explícites per trobar expressions d'aquestes arrels, mentre que per a $n \geq 5$ es pot demostrar que no n'hi ha (de fórmula explícita general).

Exercici 4.4. Calculeu el polinomi característic d'una matriu 2×2 general $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Hi ha casos, però en que és fàcil calcular les arrels del polinomi característic:

Lema 4.28. *Si $A \in M_n(\mathbb{K})$ és una matriu triangular superior (o inferior), els valors propis d' A són els valors de la diagonal d' A (els elements a_{ii}).*

Demostració. Podem calcular l'expressió $\det(A - x \mathbf{1}_n)$ desenvolupant per l'última fila i anar tirant enrera, i tindrem:

$$\det(A - x \mathbf{1}_n) = (a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x).$$

I aquesta expressió s'anul·la pels valors $x = a_{ii}$. \square

Exemple 4.29. Estudiem la diagonalització de la matriu

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

Pel Lema 4.28, com que A és triangular superior, els únics possibles valors propis són els elements 2 i 1.

Un vector propi de valor propi 2 és un vector $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ tal que:

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

I per tant queda només l'equació $y = 0$. Per tant la solució és $\vec{v} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Un vector propi de valor propi 1 és un vector $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ tal que:

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

i queda l'equació $x + y = 0$, per tant la solució és $\vec{v} = x \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Per tant, a la base $\mathcal{B} = [\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}]$, tenim:

$$[f_A]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Exemple 4.30. Estudiem la diagonalització de la matriu

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

Pel Lema 4.28, com que A és triangular superior, l'únic valor propi possible és l'1.

Si fos diagonalitzable, A hauria de ser similar a $\mathbf{1}_2$, la matriu identitat 2×2 , per tant, hauria d'existir S una matriu invertible tal que $A = S \cdot \mathbf{1}_2 \cdot S^{-1}$, però $S \cdot \mathbf{1}_2 \cdot S^{-1} = \mathbf{1}_2$, pel que A no és diagonalitzable.

El fet de que els valors propis d' A siguin les arrels del polinomi característic d' A dona una limitació del nombre de valors propis diferents que pot tenir A :

Lema 4.31. Si $A \in M_n(\mathbb{K})$, el nombre de valors propis d' A és com a molt n .

Demostració. Els valors propis d' A són els zeros de $p_A(x)$, que és un polinomi de grau n . Utilitzem ara que un polinomi de grau n sobre un cos \mathbb{K} pot tenir com a molt n arrels diferents. \square

També es poden calcular alguns coeficients del polinomi característic:

Lema 4.32. Si $A \in M_n(\mathbb{K})$,

$$p_A(x) = (-1)^n x^n + (-1)^{n-1} \text{Tr}(A) x^{n-1} + \dots + \det(A),$$

on $\text{Tr}(A)$ s'anomena la traça d' A i és la suma dels elements de la diagonal; i els termes que no estan escrits corresponen a x , x^2 , ... i x^{n-2} .

Demostració. Quan calculem $p_A(x) = \det(A - x\mathbf{1}_n)$ a partir de l'Equació (8), l'únic sumand que conté n o $n-1$ factors que contenen una x és el corresponent al patró $(1, 1), (2, 2), \dots, (n, n)$ (si fem el producte d'un patró a $A - x\mathbf{1}_n$ que té un element de fora de la diagonal, com a mínim en té dos fora de la diagonal, pel que el grau en x és menor o igual a $n-2$), i per tant, tenim que:

$$p_A(x) = (a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x) + \text{polinomi de grau } n-2 \text{ en } x,$$

i d'aquí veiem que el coeficient de x^n és $(-1)^n$ i el de x^{n-1} és $(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$.

Per veure que el terme independent de $p_A(x)$ és el determinant d' A , cal utilitzar que el terme independent de $p_A(x)$ és $p_A(0)$, i per tant és $\det(A - 0 \cdot \mathbf{1}_n) = \det(A)$. \square

Demostrem ara que el polinomi característic és el mateix per matrius similars:

Proposició 4.33. Si A i $B \in M_n(\mathbb{K})$ són matrius similars, llavors $p_A(x) = p_B(x)$.

Demostració. Si A i B són matrius similars, existeix $S \in M_n(\mathbb{K})$ una matriu invertible tal que $S^{-1}AS = B$. Llavors:

$$\begin{aligned} p_B(x) &= \det(B - x\mathbf{1}_n) = \det(S^{-1}AS - x\mathbf{1}_n) = \det(S^{-1}AS - S^{-1}(x\mathbf{1})_n S) = \\ &= \det(S^{-1}(A - x\mathbf{1}_n)S) = \det(S^{-1}) \det(A - x\mathbf{1}_n) \det(S) = \\ &= \det(S)^{-1} \det(S) \det(A - x\mathbf{1}_n) = p_A(x). \end{aligned}$$

On a l'última igualtat hem utilitzat que $\det(S^{-1}) = \det(S)^{-1}$. \square

Exercici 4.5. Demostreu que el recíproc no és cert: trobeu dues matrius A i B amb $p_A(x) = p_B(x)$ però tals que A i B no siguin similars.

4.4 Vectors propis associats a un valor propi

L'objectiu d'aquesta secció és estudiar els vectors propis d'un valor propi donat. Aquestes tenen un nom:

Definició 4.34. Fixat $\lambda \in \mathbb{K}$, un valor propi d'una matriu $A \in M_n(\mathbb{K})$, el subespai de vectors propis de valor propi λ (més el vector $\vec{0}$) s'anomena *subespai propi d'A associat a λ* i el denotem per E_λ .

En altres paraules:

$$E_\lambda = \text{Ker}(f_{A-\lambda\mathbf{1}_n}),$$

quedant demostrat que és un subespai vectorial.

Vegem que els subespais propis només s'intersequen al vector $\vec{0}$:

Lema 4.35. Fixem $A \in M_n(\mathbb{K})$ i $\lambda_1 \neq \lambda_2$ dos valors propis diferents d' A . Llavors

$$E_{\lambda_1} \cap E_{\lambda_2} = \{\vec{0}\}.$$

Més en general, vectors propis de valor propi diferent són linealment independents.

Demostració. Sigui $\vec{v} \in E_{\lambda_1} \cap E_{\lambda_2}$, llavors $A\vec{v} = \lambda_1\vec{v}$ i $A\vec{v} = \lambda_2\vec{v}$, amb $\lambda_1 \neq \lambda_2$, i això només pot passar si $\vec{v} = \vec{0}$.

Vegem ara el cas general: siguin $\{\vec{v}_1, \dots, \vec{v}_k\}$ vectors propis de valors propis $\{\lambda_1, \dots, \lambda_k\}$ respectivament, amb tots els λ_i diferents. Considerem una subconjunt de vectors de $\{\vec{v}_1, \dots, \vec{v}_k\}$ que siguin linealment independents i que sigui maximal. Aquest subconjunt tindrà ℓ vectors i podem considerar que són els primers (si és necessari, els reordenem). Per tant $\{\vec{v}_1, \dots, \vec{v}_\ell\} \subset \{\vec{v}_1, \dots, \vec{v}_k\}$ amb els ℓ primers linealment independents. Volem veure $\ell = k$. Suposem que no, llavors $\ell < k$ i existeixen uns únics $\mu_1, \dots, \mu_\ell \in \mathbb{K}$ tals que:

$$\vec{v}_{\ell+1} = \mu_1\vec{v}_1 + \dots + \mu_\ell\vec{v}_\ell. \quad (10)$$

També podem considerar que hi ha dues $\mu_i \neq 0$: com que $\vec{v}_{\ell+1} \neq \vec{0}$, com a mínim n'hi ha una. Si només n'hi hagués una, llavors $E_{\lambda_i} \cap E_{\lambda_{\ell+1}} \neq \{\vec{0}\}$, contradient el primer apartat d'aquest lema. Suposem doncs (si cal, reordenem), $\mu_1 \neq 0 \neq \mu_2$.

Resumint, la situació és la següent: si $\ell < k$, podem considerar $\vec{v}_{\ell+1}$ és un vector propi de valor propi $\lambda_{\ell+1}$ que es pot posar com $\vec{v}_{\ell+1} = \mu_1\vec{v}_1 + \dots + \mu_\ell\vec{v}_\ell$ amb $\mu_1 \neq 0 \neq \mu_2$ i \vec{v}_1 i \vec{v}_2 són vectors propis de valor propi λ_1 i λ_2 respectivament, amb $\lambda_1 \neq \lambda_2$. Apliquem A a l'Equació (10):

$$\begin{aligned} \lambda_{\ell+1}\vec{v}_{\ell+1} &= A\vec{v}_{\ell+1} = A(\mu_1\vec{v}_1 + \dots + \mu_\ell\vec{v}_\ell) = \\ &= \lambda_1\mu_1\vec{v}_1 + \lambda_2\mu_2\vec{v}_2 + \dots + \lambda_\ell\mu_\ell\vec{v}_\ell \end{aligned}$$

I també:

$$\lambda_{\ell+1}\vec{v}_{\ell+1} = \lambda_{\ell+1}\mu_1\vec{v}_1 + \lambda_{\ell+1}\mu_2\vec{v}_2 + \cdots + \lambda_{\ell+1}\mu_\ell\vec{v}_\ell$$

Per tant, com que $\{\vec{v}_1, \dots, \vec{v}_\ell\}$ són linealment independents, tenim:

$$\lambda_{\ell+1}\mu_1 = \lambda_1\mu_1 \text{ i } \lambda_{\ell+1}\mu_2 = \lambda_2\mu_2$$

amb $\mu_1 \neq 0 \neq \mu_2$, per tant $\lambda_1 = \lambda_{\ell+1} = \lambda_2$, contradient que $\lambda_1 \neq \lambda_2$.

Per tant, la contradicció ve de suposar $\ell < k$, el que implica $\ell = k$ i els k vectors són linealment independents. \square

Exemple 4.36. Si agafem la matriu

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix},$$

als càlculs de l'Exemple 4.29 hem vist que els subespais propis són:

$$E_2 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \text{ i } E_1 = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle.$$

A més, com que tenim dos vectors propis linealment independents en un espai de dimensió 2, tenim una base de vectors propis i la matriu A diagonalitza.

Exemple 4.37. Si agafem ara la matriu

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

als càlculs de l'Exemple 4.30 hem vist que l'únic subespai propi és:

$$E_1 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.$$

En aquest cas, el subespai format pels vectors propis té dimensió 1, pel que no tenim una base de vectors propis i la matriu A no diagonalitza.

Més en general, per saber si una matriu diagonalitza, cal estudiar el polinomi característic i els subespais propis. Per això, ens convé la definició següent:

Definició 4.38. Considerem $A \in M_n(\mathbb{K})$ i λ un valor propi d' A .

- Definim la *multiplicitat algebraica de λ com a valor propi d' A* com el valor $m \geq 1$ tal que

$$p_A(x) = (x - \lambda)^m q(x)$$

amb $q(x)$ un polinomi de grau $n - m$ tal que $q(\lambda) \neq 0$. Escriurem $\text{MultAlg}_A(\lambda)$.

- Definim la *multiplicitat geomètrica de λ com a valor propi d' A* com la dimensió de E_λ . Escriurem $\text{MultGeom}_A(\lambda)$.

Observació 4.39. A la definició de $\text{MultAlg}_A(\lambda)$ utilitzem que si λ és un valor propi, llavors, $p(\lambda) = 0$. Si ara fem la divisió de polinomis $p(x)$ dividit per $(x - \lambda)$, existeixen polinomis $q_1(x)$ (quocient) i $r_1(x)$ (la resta, que com que ha de ser de grau menor a 1, ha de ser una constant, i per tant escrivim r_1) tals que:

$$p(x) = (x - \lambda)q_1(x) + r.$$

Si avaluem a λ ens queda:

$$0 = p(\lambda) = (\lambda - \lambda)q_1(\lambda) + r_1 = r_1,$$

i per tant $r_1 = 0$, obtenint que:

$$p(x) = (x - \lambda)q_1(x).$$

Aquest procediment es pot fer un altre cop si $q_1(\lambda) = 0$, i tindriem que $p(x) = (x - \lambda)^2 q_2(x)$.

Iterem el procediment fins que $q_m(\lambda) \neq 0$, no podem continuar el procediment, i definim la multiplicitat algebraica d'aquesta manera.

Mirem ara un cas particular:

Proposició 4.40. *Si $A \in M_n(\mathbb{K})$ té n valor propis diferents, llavors:*

- (a) *Per a cada λ valor propi d' A , $\text{MultAlg}_A(\lambda) = \text{MultGeom}_A(\lambda) = 1$.*
- (b) *La matriu A diagonalitza.*

Demostració. Com que tenim n valors propis diferents, com a mínim tenim n vectors propis de valor propi diferents, i pel Lema 4.35, seran linealment independents a \mathbb{K}^n , per tant una base. D'aquí obtenim que A diagonalitza.

Les n multiplicitats $\text{MultAlg}_A(\lambda)$ i $\text{MultGeom}_A(\lambda)$ són com a mínim 1 i, sumades, com a molt n (la suma de les MultAlg és, com a molt, el grau de $p_A(x)$; la suma de les MultGeom és, com a molt, el nombre màxim d'una família de vectors linealment independents a \mathbb{K}^n), per tant, han de ser 1. \square

En general, la situació no serà tant bona i el que tenim és:

Lema 4.41. *Si $A \in M_n(\mathbb{K})$ i λ és un valor propi d' A , llavors:*

- (a) *Si λ és un valor propi d' A , llavors $\text{MultGeom}(\lambda) \leq \text{MultAlg}(\lambda)$.*
- (b) *A diagonalitza si i només si*

- *$p_A(x)$ es pot escriure com a producte de polinomis de grau 1*

$$p_A(x) = (-1)^n (x - \lambda_1)^{\text{MultAlg}_A(\lambda_1)} \cdots (x - \lambda_k)^{\text{MultAlg}_A(\lambda_k)} \quad (11)$$

amb λ_i valors propis diferents i

- *$\text{MultGeom}_A(\lambda_i) = \text{MultAlg}_A(\lambda_i)$ per a tot $i = 1, \dots, k$.*

Demostració. Vegem primer (a): sigui λ un valor propi d' A i considerem $m = \text{MultGeom}(\lambda)$, i per tant $\vec{v}_1, \dots, \vec{v}_m$ vectors linealment independents de $\text{Ker}(A - \lambda \mathbf{1}_n)$. Ampliem la família $\vec{v}_1, \dots, \vec{v}_m$ fins a tenir una base \mathcal{B} de \mathbb{K}^n . En aquesta base, la matriu $[f_A]_{\mathcal{B}}$ tindrà a les primeres m columnes tots els coeficients zero, excepte la diagonal, que valdrà λ , per tant, $p_A(x) = p_{[f_A]_{\mathcal{B}}}(x) = (-1)^n (x - \lambda)^m q(x)$ i per tant $m = \text{MultGeom}(A) \leq \text{MultAlg}(A)$.

Vegem ara (b): suposem primer que A diagonalitza, llavors, la suma de les multiplicitats geomètriques d' A serà n :

$$\text{MultGeom}_A(\lambda_1) + \cdots + \text{MultGeom}_A(\lambda_k) = n$$

amb λ_i valors propis diferents. Però com que $\text{MultGeom}(\lambda) \leq \text{MultAlg}(\lambda)$,

$$n \leq \text{MultAlg}_A(\lambda_1) + \cdots + \text{MultAlg}_A(\lambda_n) \leq n$$

on la segona desigualtat és perquè $p_A(x)$ té grau n , per tant, les multiplicitats geomètriques i algebraiques són iguals i el polinomi $p_A(x)$ es pot escriure com a producte de polinomis de grau 1.

El recíproc es fa desfent els arguments previs: si $p_A(x)$ factoritza com a producte de factors de grau 1 com a l'Equació (11), per a que diagonalitzi, s'ha de complir que trobem una base de vectors propis. Com que per cada λ valor propi, per hipòtesis, $\text{MultGeom}(\lambda)_A = \text{MultAlg}_A(\lambda)$, tindrem n vector propis linealment independents, i per tant una base en que A diagonalitza. \square

Exemple 4.42. Volem estudiar si la matriu A següent diagonalitza i, si ho fa, en quina base:

$$A = \begin{pmatrix} 55 & 91 & -29 & 50 \\ -27 & -47 & 15 & -24 \\ -33 & -67 & 21 & -26 \\ -33 & -57 & 18 & -29 \end{pmatrix}.$$

Calculem primer el polinomi característic:

$$p_A(x) = \begin{vmatrix} 55-x & 91 & -29 & 50 \\ -27 & -47-x & 15 & -24 \\ -33 & -67 & 21-x & -26 \\ -33 & -57 & 18 & -29-x \end{vmatrix} = x^4 - 3x^2 + 2x = x(x-1)^2(x+2).$$

Lavors hem de calcular $E_0 = \text{Ker}(f_A)$, $E_1 = \text{Ker}(f_{A-1_4})$ i $E_{-2} = \text{Ker}(f_{A+2 \cdot 1_4})$. Sense calcular-los explícitament, per comprovar si diagonalitza només cal calcular $\dim(E_1)$. Si $\dim(E_1) = 2$, llavors diagonalitza, però si $\dim(E_1) = 1$, no diagonalitza. En aquest cas surt $\dim(E_1) = 2$ i per tant diagonalitza. Fem els càlculs:

$$E_0 = \left\langle \begin{pmatrix} 4 \\ -3 \\ -7 \\ -3 \end{pmatrix} \right\rangle, E_1 = \left\langle \begin{pmatrix} 4 \\ 2 \\ -4 \\ -3 \end{pmatrix}, \begin{pmatrix} 11 \\ -5 \\ -9 \\ -8 \end{pmatrix} \right\rangle \text{ i } E_{-2} = \left\langle \begin{pmatrix} 15 \\ -9 \\ -16 \\ -10 \end{pmatrix} \right\rangle$$

Pel que podem confirmar que diagonalitza (hi ha 4 vectors propis linealment independents). Lavors es pot comprovar que:

$$D = S^{-1}AS$$

on

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \text{ i } S = \begin{pmatrix} 4 & 4 & 11 & 15 \\ -3 & 2 & -5 & -9 \\ -7 & -4 & -9 & -16 \\ -3 & -3 & -8 & -10 \end{pmatrix}.$$

Exemple 4.43. Considerem ara la matriu:

$$A = \begin{pmatrix} 107 & 151 & -49 & 106 \\ -53 & -77 & 25 & -52 \\ -85 & -127 & 41 & -82 \\ -72 & -102 & 33 & -71 \end{pmatrix}$$

Podem fer els càlculs d'abans, obtenint el mateix polinomi característic:

$$p_A(x) = \begin{vmatrix} 107-x & 151 & -49 & 106 \\ -53 & -77-x & 25 & -52 \\ -85 & -127 & 41-x & -82 \\ -72 & -102 & 33 & -71-x \end{vmatrix} = x^4 - 3x^2 + 2x = x(x-1)^2(x+2).$$

En aquest cas, però, tenim que (després de fer els càlculs) $\dim(E_1) = 1 < 2 = \text{MultAlg}_A(1)$, per tant no diagonalitza.

Exemple 4.44 (La successió de Fibonacci). Es defineix la successió de Fibonacci com:

$$a_1 = a_2 = 1 \text{ i } a_n = a_{n-1} + a_{n-2} \text{ per a } n \geq 2.$$

Podem calcular-ne els seus primers termes:

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$$

aquesta definició fa que per calcula el terme n , haguem de calcular tots els anteriors. L'objectiu és trobar una fórmula que $f(n)$ tal que $a_n = f(n)$.

Considerem els vectors $\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}$, i veiem que es compleix la relació següent:

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} a_{n-2} \\ a_{n-3} \end{pmatrix} = \dots = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-2} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Per tant, ens convé calcular:

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

Si aconseguim escriure $A = SDS^{-1}$ (D matriu diagonal, per tant, la diagonal de D està formada pels valors propis de A), tenim que $A^{n-2} = SD^{n-2}S^{-1}$, i, si $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, llavors:

$$D^{n-2} = \begin{pmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{pmatrix}$$

Calculem els valor propis d' A . Calculem el polinomi característic:

$$p(x) = \det(A - x\mathbf{1}_2) = \begin{vmatrix} 1-x & 1 \\ 1 & -x \end{vmatrix} = x^2 - x - 1$$

I els valors propis són les solucions de l'equació $p(x) = 0$, per tant:

$$\lambda_i = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

Per tant, com que tenim dos valors propis diferents, ja sabem que diagonalitza i tenim:

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

Per calcular els vectors propis corresponents a cada valor propi, hem de resoldre els sistemes homogenis:

$$\left. \begin{matrix} (1 - \lambda_i)x + y = 0 \\ x - \lambda_i y = 0 \end{matrix} \right\} \iff \begin{pmatrix} x \\ y \end{pmatrix} \in \left\langle \begin{pmatrix} \lambda_i \\ 1 \end{pmatrix} \right\rangle$$

Llavors, tenim que:

$$S = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \text{ i } S^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}$$

I per tant:

$$A^{n-2} = SD^{n-2}S^{-1} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^{n-2} & 0 \\ 0 & \lambda_2^{n-2} \end{pmatrix} \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}$$

I si fem els càlculs:

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1^{n-1} - \lambda_2^{n-1} & \lambda_1^{n-2} - \lambda_2^{n-2} \\ \lambda_1^{n-2} - \lambda_2^{n-2} & \lambda_1^{n-3} - \lambda_2^{n-3} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Llavors, després de simplificar:

$$a_n = \frac{1}{\lambda_1 - \lambda_2} (\lambda_1^n - \lambda_2^n) = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$

4.5 Matrius sobre \mathbb{R}

En aquest apartat, suposarem donada una matriu $A \in M_n(\mathbb{R})$. Una conseqüència del teorema fonamental de l'àlgebra és que $p_A(x)$ descomposa en producte de factors de grau 1 i 2. Si hi ha algun factor de grau 2, aleshores A no diagonalitza sobre \mathbb{R} , tal i com hem vist.

Exemple 4.45. Considerem la matriu $R_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, on $a, b \in \mathbb{R}$ i $b \neq 0$. El seu polinomi característic és

$$p_A(x) = x^2 - 2ax + a^2 + b^2 = (x - (a + ib))(x - (a - ib)).$$

Escrivim $\lambda^+ = a + bi$, i $\lambda^- = a - bi$. Podem calcular

$$E_{\lambda^+} = \left\langle \begin{pmatrix} i \\ 1 \end{pmatrix} \right\rangle, \quad E_{\lambda^-} = \left\langle \begin{pmatrix} -i \\ 1 \end{pmatrix} \right\rangle.$$

Per tant, obtenim

$$\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} a + bi & 0 \\ 0 & a - bi \end{pmatrix}.$$

Exemple 4.46. Suposem que $A \in M_2(\mathbb{R})$ té valors propis $a \pm bi$, amb $b \neq 0$. Com que els valors propis són diferents, la matriu A és diagonalitzable i, per tant, és similar a

$$D = \begin{pmatrix} a + bi & 0 \\ 0 & a - bi \end{pmatrix}.$$

Com que la matriu D és també similar a

$$R_{a,b} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

la matriu A és similar a $R_{a,b}$. De fet, tenim:

$$D = S^{-1}AS,$$

on S té per columnes \vec{u} i $\vec{\bar{u}}$, on $\vec{u} = \vec{v} + i\vec{w}$, amb $\vec{v}, \vec{w} \in \mathbb{R}^2$. Substituint $D = T^{-1}R_{a,b}T$ amb $T = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$, tenim

$$T^{-1}R_{a,b}T = S^{-1}AS \implies R_{a,b} = TS^{-1}AST^{-1}.$$

Calculem directament que

$$ST^{-1} = \begin{pmatrix} | & | \\ \vec{w} & \vec{v} \\ | & | \end{pmatrix}.$$

Exercicis recomanats

Els exercicis que segueixen són útils per practicar el material presentat. La numeració és la de [1].

Secció 6.2: 4, 18.

Secció 6.3: 2, 6.

Secció 7.1: 10, 18, 34, 36.

Secció 7.2: 2, 10, 20.

Secció 7.3: 14, 22, 24.

Secció 7.5: 14, 24, 26.

5 Ortogonalitat

Podeu trobar el contingut de la part d'ortogonalitat a [1, Tema 5], i de la part de formes quadràtiques a ([1, Tema 8],[2, Tema 4]).

5.1 Ortogonalitat a \mathbb{R}^n

Considerem l'espai vectorial \mathbb{R}^n i recordem el producte escalar definit com: si $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ i $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ són vectors d' \mathbb{R}^n , llavors:

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = \sum_{i=1}^n u_i v_i.$$

Definició 5.1. • Diem que dos vectors \vec{u} i \vec{v} són *ortogonals* si $\vec{u} \cdot \vec{v} = 0$.

- Definim la *longitud* d'un vector:

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{\sum_{i=1}^n u_i^2}.$$

- Diem que un vector $\vec{u} \in \mathbb{R}^n$ és *unitari* si $\|\vec{u}\| = 1$.
- Diem que els vectors $\vec{u}_1, \dots, \vec{u}_k$ de \mathbb{R}^n són *ortogonals* si són ortogonals dos a dos, o sigui, si $\vec{u}_i \cdot \vec{u}_j = 0$ per a tot $i \neq j$.
- Diem que els vectors $\vec{u}_1, \dots, \vec{u}_k$ de \mathbb{R}^n són *ortonormals* si són unitaris i ortogonals dos a dos, o sigui, si

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Observació 5.2. Podem passar de vectors no nuls $\vec{u}_1, \dots, \vec{u}_k$ de \mathbb{R}^n ortogonals a ortonormals dividint cadascun per la seva norma: si $\vec{u}_1, \dots, \vec{u}_k$ de \mathbb{R}^n són no nuls, llavors $\frac{\vec{u}_1}{\|\vec{u}_1\|}, \dots, \frac{\vec{u}_k}{\|\vec{u}_k\|}$ són unitaris, i si $\vec{u}_i \cdot \vec{u}_j = 0$, llavors, $\frac{\vec{u}_i}{\|\vec{u}_i\|} \cdot \frac{\vec{u}_j}{\|\vec{u}_j\|} = \frac{1}{\|\vec{u}_i\|\|\vec{u}_j\|}(\vec{u}_i \cdot \vec{u}_j) = 0$.

Exemple 5.3. La base estàndard d' \mathbb{R}^n formada pels vectors \vec{e}_i amb un 1 a la posició i i 0 a la resta de posicions és una base ortonormal.

Lema 5.4. Si els vectors no nuls $\vec{u}_1, \dots, \vec{u}_k$ de \mathbb{R}^n són ortogonals, llavors són linealment independents.

En particular, si tenim n vectors $\vec{u}_1, \dots, \vec{u}_n$ de \mathbb{R}^n ortonormals, formen una base.

Demostració. Suposem que tenim una combinació lineal dels vectors $\vec{u}_1, \dots, \vec{u}_k$ que dona zero:

$$0 = \lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n.$$

Llavors, fent el producte escalar per \vec{u}_i tenim:

$$\begin{aligned} 0 &= (\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n) \cdot \vec{u}_i = \\ &= \lambda_1 (\vec{u}_1 \cdot \vec{u}_i) + \dots + \lambda_n (\vec{u}_n \cdot \vec{u}_i) = \\ &= \lambda_i \|\vec{u}_i\|^2 \quad (\text{tots els altres productes escalars són zero}) \end{aligned}$$

Per tant, com que $\vec{u}_i \neq \vec{0}$, $\lambda_i = 0$ i són linealment independents.

En el cas de tenir n vectors ortonormals, són n vectors ortogonals no nuls, per tant linealment independents i com que la dimensió de \mathbb{R}^n és n , són base (Teorema 3.42). \square

Veurem més endavant com trobar una base ortonormal d'un subespai V de \mathbb{R}^n . Suposem però que tenim $[\vec{u}_1, \dots, \vec{u}_k]$ una base ortonormal de V , i considerem la *projecció ortogonal* $\text{Proj}_V: \mathbb{R}^n \rightarrow \mathbb{R}^n$, que envia un vector $\vec{v} \in \mathbb{R}^n$ al vector $\text{Proj}_V(\vec{v}) = \vec{v}^\parallel$, on

$$\vec{v}^\parallel = (\vec{v} \cdot \vec{u}_1)\vec{u}_1 + \dots + (\vec{v} \cdot \vec{u}_k)\vec{u}_k \in V.$$

Observem que $\text{Proj}_V: \mathbb{R}^n \rightarrow \mathbb{R}^n$ és una aplicació lineal (per les propietats de linealitat del producte escalar). Com que la seva imatge està continguda a V i $\text{Proj}_V(\vec{u}_i) = \vec{u}_i$, en deduïm que $\text{Im}(\text{Proj}_V) = V$. A més,

$$\begin{aligned} \text{Ker}(\text{Proj}_V) &= \{\vec{w} \in \mathbb{R}^n \mid \vec{w}^\parallel = 0\} \\ &= \{\vec{w} \in \mathbb{R}^n \mid \vec{w} \cdot \vec{u}_i = 0, \quad i = 1, \dots, k\} \\ &= \{\vec{w} \in \mathbb{R}^n \mid \vec{w} \cdot \vec{v} = 0 \quad \text{per a tot } \vec{v} \in V\}. \end{aligned}$$

Donarem un nom al nucli de Proj_V :

Definició 5.5. Donat un subespai $V \subseteq \mathbb{R}^n$, el *complement ortogonal* de V , que escriurem V^\perp , és el conjunt

$$V^\perp = \{\vec{w} \in \mathbb{R}^n \mid \vec{w} \cdot \vec{v} = 0 \text{ per a tot } \vec{v} \in V\}.$$

Proposició 5.6. Sigui $V \subseteq \mathbb{R}^n$ un subespai d' \mathbb{R}^n . Aleshores:

- (a) V^\perp també és un subespai d' \mathbb{R}^n ,
- (b) $V \cap V^\perp = \{\vec{0}\}$,
- (c) $\dim(V) + \dim(V^\perp) = n$,
- (d) $(V^\perp)^\perp = V$.

Demostració. La primera afirmació es comprova de manera directa, fent servir les propietats del producte escalar.

Per veure la segona, sigui $\vec{w} \in V \cap V^\perp$. Per tant, $\vec{w} \cdot \vec{w} = \|\vec{w}\|^2 = 0$. Però l'únic vector de longitud 0 és el vector nul.

Seguidament, observem que $V = \text{Im}(\text{Proj}_V)$ i que $V^\perp = \text{Ker}(\text{Proj}_V)$. Per tant, per la fórmula del nucli-imatge (Teorema 3.45) tenim

$$\dim(\text{Ker Proj}_V) + \dim(\text{Im Proj}_V) = \dim(\mathbb{R}^n) = n,$$

com volíem veure.

Finalment, per veure que $(V^\perp)^\perp = V$, denotem $W = V^\perp$. Aleshores, si $\vec{v} \in V$, i $\vec{w} \in W$, es té que $\vec{v} \cdot \vec{w} = 0$, i per tant $\vec{v} \in W^\perp$. Concloem doncs que $V \subseteq (V^\perp)^\perp$. Però ara observem que $\dim((V^\perp)^\perp) = n - \dim(V^\perp) = n - (n - \dim(V)) = \dim(V)$. Veïem que $(V^\perp)^\perp$ és un subespai que conté V i que té la mateixa dimensió que V i, per tant, ha de coincidir amb V . \square

Corol·lari 5.7. Si $V \subseteq \mathbb{R}^n$ és un subespai d' \mathbb{R}^n , aleshores $\mathbb{R}^n = V \oplus V^\perp$.

Demostració. Per l'apartat (b) de la proposició anterior, la intersecció és buida. Per l'apartat (c), la suma de les dimensions és n . Per tant, si prenem una base \mathcal{B} de V i una base \mathcal{C} de V^\perp , el conjunt $\mathcal{B} \cup \mathcal{C}$ serà linealment independent i tindrà n elements. Per tant, serà base de \mathbb{R}^n , que és el que volíem veure. \square

D'aquest corol·lari es dedueix que, donat V un subespai d' \mathbb{R}^n , tot vector $\vec{w} \in \mathbb{R}^n$ es pot escriure de forma única com:

$$\vec{w} = \vec{w}^\parallel + \vec{w}^\perp$$

amb $\vec{w}^\parallel \in V$ i $\vec{w}^\perp \in V^\perp$. A més, tenim que $\vec{w}^\parallel = \text{Proj}_{V^\perp}(\vec{w})$.

5.2 El teorema de Pitàgores

El conegut teorema de Pitàgores es pot formular a \mathbb{R}^n . Fixem-nos que el teorema té dues implicacions.

Teorema 5.8. *Si \vec{v} i \vec{w} són dos vectors d' \mathbb{R}^n , aleshores:*

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 \iff \vec{v} \perp \vec{w}.$$

Demostració. Calculem:

$$\|\vec{v} + \vec{w}\|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \vec{w} \cdot \vec{w} = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2(\vec{v} \cdot \vec{w}).$$

Per tant, la igualtat es compleix si i només si $\vec{v} \cdot \vec{w} = 0$. □

Corol·lari 5.9. *Si $V \subseteq \mathbb{R}^n$ un subespai. Aleshores*

$$\|\text{Proj}_V(\vec{w})\| \leq \|\vec{w}\| \text{ per a tot } \vec{w} \in \mathbb{R}^n,$$

i la desigualtat és una igualtat si i només si $\vec{w} \in V$.

Demostració. Apliquem el teorema anterior a \vec{w}^\parallel i \vec{w}^\perp (notem $\vec{w} = \vec{w}^\parallel + \vec{w}^\perp$).

$$\|\vec{w}\|^2 = \|\vec{w}^\parallel\|^2 + \|\vec{w}^\perp\|^2,$$

i per tant

$$\|\text{Proj}_V(\vec{w})\|^2 \leq \|\vec{w}\|^2,$$

amb igualtat si i només si $\|\vec{w}^\perp\|^2 = 0$, és a dir si i només si $\vec{w}^\perp = 0$, que és equivalent a $\vec{w} \in V$. □

Una aplicació d'aquests resultats és una desigualtat molt famosa, coneguda com la *desigualtat de Cauchy-Schwarz*:

Teorema 5.10. *Si \vec{v} i \vec{w} són vectors de \mathbb{R}^n , aleshores*

$$|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \|\vec{w}\|,$$

amb igualtat si i només si \vec{v} i \vec{w} són paral·lels.

Demostració. Si $\vec{v} = \vec{0}$, aleshores la desigualtat es compleix trivialment. En cas contrari, considerem el subespai $V = \langle \vec{v} \rangle$, i apliquem el resultat anterior. Ja havíem vist la fórmula

$$\text{Proj}_V(\vec{w}) = \frac{\vec{w} \cdot \vec{v}}{\|\vec{v}\|^2} \vec{v},$$

i per tant

$$\|\text{Proj}_V(\vec{w})\| = \frac{|\vec{w} \cdot \vec{v}|}{\|\vec{v}\|^2} \|\vec{v}\|,$$

i obtenim

$$\frac{|\vec{w} \cdot \vec{v}|}{\|\vec{v}\|} \leq \|\vec{w}\|,$$

que és la desigualtat que busquem un cop passem $\|\vec{v}\|$ a l'altra banda (i utilitzem que $\vec{w} \cdot \vec{v} = \vec{v} \cdot \vec{w}$). La igualtat es dona si i només si $\vec{w} \in V$, que és equivalent a \vec{v} essent paral·lel a \vec{w} . □

Observem que si \vec{v} i \vec{w} no són zero, aleshores tenim les desigualtats

$$-1 \leq \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \|\vec{w}\|} \leq 1.$$

Definició 5.11. Donats dos vectors no-nuls \vec{v} i \vec{w} d' \mathbb{R}^n , l'angle entre \vec{v} i \vec{w} és

$$\theta = \arccos \left(\frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \|\vec{w}\|} \right).$$

5.3 Mètode de Gram-Schmidt

Aquest mètode permet calcular una base ortonormal d'un subespai de \mathbb{R}^n a partir de projeccions ortogonals. Fixem les notacions següents:

- $V \subset \mathbb{R}^n$ un subespai de dimensió k fixat,
- $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_k]$ una base de V ,
- $V_i = \langle \vec{v}_1, \dots, \vec{v}_i \rangle \subset V$ el subespai (de dimensió i) generat pels i primers vectors de \mathcal{B} . Tenim que $V_1 \subset V_2 \subset \dots \subset V_k = V$.
- Si $\vec{v}_i \in \mathcal{B}$, denotarem per \vec{v}_i^{\parallel} la projecció ortogonal de \vec{v}_i a V_{i-1} i per \vec{v}_i^{\perp} el vector $\vec{v}_i - \vec{v}_i^{\parallel}$, que és ortogonal a V_{i-1} .

El mètode de Gram-Schmidt construeix de forma iterativa una base ortonormal a cada V_i :

- Considerem primer $V_1 = \langle \vec{v}_1 \rangle$, i una base és $\mathcal{B}_1 = [\vec{v}_1]$. En aquest cas, definim

$$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1$$

i tenim que $\mathcal{C}_1 = [\vec{u}_1]$ és una base ortonormal de V_1 .

- Considerem ara $V_2 = \langle \vec{v}_1, \vec{v}_2 \rangle = \langle \vec{u}_1, \vec{v}_2 \rangle$. Pel Corol·lari 5.7, \vec{v}_2 es pot escriure de forma única com:

$$\vec{v}_2 = \vec{v}_2^{\parallel} + \vec{v}_2^{\perp}$$

amb $\vec{v}_2^{\parallel} \in V_1$ i $\vec{v}_2^{\perp} \perp V_1$. A més, sabem com calcular-los:

$$\vec{v}_2^{\parallel} = \text{Proj}_{V_1}(\vec{v}_2) = (\vec{v}_2 \cdot \vec{u}_1) \vec{u}_1$$

i per tant

$$\vec{v}_2^{\perp} = \vec{v}_2 - \vec{v}_2^{\parallel} = \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1) \vec{u}_1.$$

Definim $\vec{u}_2 = \frac{1}{\|\vec{v}_2^{\perp}\|} \vec{v}_2^{\perp}$, i tenim que $V_2 = \langle \vec{u}_1, \vec{u}_2 \rangle$, amb $\mathcal{C}_2 = [\vec{u}_1, \vec{u}_2]$ una base ortonormal.

- Suposem ara que ja tenim $\mathcal{C}_{i-1} = [\vec{u}_1, \dots, \vec{u}_{i-1}]$ una base ortonormal de V_{i-1} . Considerem $V_i = \langle \vec{v}_1, \dots, \vec{v}_{i-1}, \vec{v}_i \rangle = \langle \vec{u}_1, \dots, \vec{u}_{i-1}, \vec{v}_i \rangle$, i escrivim:

$$\vec{v}_i = \vec{v}_i^{\parallel} + \vec{v}_i^{\perp}$$

amb $\vec{v}_i^{\parallel} \in V_{i-1}$ i $\vec{v}_i^{\perp} \perp V_{i-1}$. Com que els vectors \vec{u}_i són ortonormals, el càlcul és:

$$\vec{v}_i^{\parallel} = \text{Proj}_{V_{i-1}}(\vec{v}_i) = (\vec{v}_i \cdot \vec{u}_1) \vec{u}_1 + \dots + (\vec{v}_i \cdot \vec{u}_{i-1}) \vec{u}_{i-1}$$

i per tant

$$\vec{v}_i^{\perp} = \vec{v}_i - \vec{v}_i^{\parallel} = \vec{v}_i - (\vec{v}_i \cdot \vec{u}_1) \vec{u}_1 - \dots - (\vec{v}_i \cdot \vec{u}_{i-1}) \vec{u}_{i-1}.$$

Definim $\vec{u}_i = \frac{1}{\|\vec{v}_i^{\perp}\|} \vec{v}_i^{\perp}$, i tenim que $V_i = \langle \vec{u}_1, \dots, \vec{u}_i \rangle$, amb $\mathcal{C}_i = [\vec{u}_1, \dots, \vec{u}_i]$ una base ortonormal.

Exemple 5.12. Considerem $V = \langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle \subset \mathbb{R}^3$ i volem calcular una base ortonormal.

Comencem amb la base $\mathcal{B} = [\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}]$ i, seguint el procediment, $V_1 = \langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \rangle$ i tant sols hem de fer-lo unitari: $\|\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\| = 2$ i per tant $V_1 = \langle \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix} \rangle$.

Calculem ara el segon vector:

$$\vec{v}_2^\perp = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix} \right) \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 3/2 \\ -1/2 \end{pmatrix}.$$

$$\text{Llavors } \vec{u}_2 = \frac{1}{\|\vec{v}_2^\perp\|} \vec{v}_2^\perp = \begin{pmatrix} -1/2 \\ 3/2 \\ -1/2 \end{pmatrix}.$$

Tenim, doncs, que $\mathcal{C} = [\begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 3/2 \\ -1/2 \end{pmatrix}]$ és una base ortonormal de V .

A més, la matriu del canvi de base és:

$$S_{\mathcal{B},\mathcal{C}} = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}$$

L'algorisme de Gram-Schmidt ens porta a la factorització QR d'una matriu A (amb les seves columnes linealment independents).

Teorema 5.13. Donada una matriu $A \in M_{m \times n}(\mathbb{R})$ amb les seves columnes linealment independents, existeixen una matriu $Q \in M_{m \times n}(\mathbb{R})$ i $R \in M_{n \times n}$ tals que:

- $A = QR$,
- $Q^T Q = \mathbf{1}_n$ (diem que Q és una matriu ortogonal),
- R és una matriu triangular superior amb els coeficients de la diagonal positius.

A més, aquesta factorització és única.

Demostració. La demostració és constructiva: anirem calculant els coeficients de les columnes Q i R aplicant el mètode de Gram-Schmidt. Fixem abans les notacions següents per a les columnes d' A , de Q i els coeficients de R :

$$A = \begin{pmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{pmatrix}, Q = \begin{pmatrix} | & | & \cdots & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & \cdots & | \end{pmatrix} \text{ i } R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}.$$

I les matrius formades per les primeres j columnes d' A , Q i la submatriu quadrada $j \times j$ d' R :

$$A_j = \begin{pmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_j \\ | & | & \cdots & | \end{pmatrix}, Q_j = \begin{pmatrix} | & | & \cdots & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_j \\ | & | & \cdots & | \end{pmatrix} \text{ i } R_j = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ 0 & r_{22} & \cdots & r_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{jj} \end{pmatrix}.$$

La construcció comença amb $j = 1$, $j = 2$ i veurem com es fa el cas j a partir del $j - 1$:

- Cas $j = 1$: llavors $A_1 = \vec{v}_1$: calculem el coeficient $R_1 = (r_{11}) = (\|\vec{v}_1\|)$ i la primera columna de Q com $Q_1 = \vec{u}_1 = \frac{1}{r_{11}}\vec{v}_1$. Veiem que, de moment, es compleix $A_1 = Q_1 R_1$. A més, $Q_1^T Q_1 = \|\vec{u}_1\| = 1$ i R_1 és triangular superior i amb la diagonal positiva (és un mòdul).
- Cas $j = 2$: considerem $r_{12} = \vec{u}_1 \cdot \vec{v}_2$, el vector $\vec{v}_2^\perp = \vec{v}_2 - r_{12}\vec{u}_1$, $r_{22} = \|\vec{v}_2^\perp\|$ i $\vec{u}_2 = \frac{1}{r_{22}}\vec{v}_2^\perp$. La primera columna de $Q_2 R_2$ és la mateixa que la de $Q_1 R_1 = A_1$. La segona és

$$r_{12}\vec{u}_1 + r_{22}\vec{u}_2 = (u_1 \cdot v_2)\vec{u}_1 + \|v_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1\| \frac{1}{\|\vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1\|} (\vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1) = \vec{v}_2$$

i per tant tenim la segona columna de A , llavors $A_2 = Q_2 R_2$. També veiem que $Q_2^T Q_2 = \mathbf{1}_2$ (u_2 és perpendicular a \vec{u}_1 i tots dos són unitaris) i R_2 és triangular superior i la diagonal positiva (són mòduls de vectors).

- Suposem que tenim les $j-1$ primeres columnes de Q , i les de R tal que $A_{j-1} = Q_{j-1} R_{j-1}$, $Q_{j-1}^T Q_{j-1} = \mathbf{1}_{j-1}$ i R_{j-1} triangular superior i amb la diagonal positiva. Considerem $r_{ij} = \vec{u}_i \cdot \vec{v}_j$ per a $1 \leq i < j$, $\vec{v}_j^\perp = \vec{v}_j - r_{1j}\vec{u}_1 - \dots - r_{(j-1)j}\vec{u}_{j-1}$, $r_{jj} = \|\vec{v}_j^\perp\|$ i $\vec{u}_j = \frac{1}{r_{jj}}\vec{v}_j^\perp$. Si fem $Q_j R_j$, les primeres $j-1$ columnes són les de $Q_{j-1} R_{j-1}$, per tant són A_{j-1} . Si calculem la columna j -èssima de $Q_j R_j$:

$$\begin{aligned} r_{1j}\vec{u}_1 + \dots + r_{jj}\vec{u}_j &= (\vec{u}_1 \cdot \vec{v}_j)\vec{u}_1 + \dots + (\vec{u}_{j-1} \cdot \vec{v}_j)\vec{u}_{j-1} + \|\vec{v}_j^\perp\| \frac{1}{\|\vec{v}_j^\perp\|} \vec{v}_j^\perp = \\ &= (\vec{u}_1 \cdot \vec{v}_j)\vec{u}_1 + \dots + (\vec{u}_{j-1} \cdot \vec{v}_j)\vec{u}_{j-1} + \vec{v}_j^\perp = \vec{v}_j. \end{aligned}$$

I per tant $A_j = Q_j R_j$. A més, la primera caixa $(j-1) \times (j-1)$ superior esquerra de $Q_j^T Q_j$ és una $\mathbf{1}_{j-1}$. L'última columna i fila són $\vec{u}_i \cdot \vec{u}_j$, i per tant, com que \vec{u}_j és perpendicular als altres \vec{u}_i i unitari, tenim que $Q_j^T Q_j = \mathbf{1}_j$. Finalment, R_j és triangular superior per construcció, i la diagonal formada per mòduls, per tant positius.

Falta veure la unicitat, també per inducció sobre j :

- Cas $j = 1$: com que R és triangular superior, $v_1 = r_{11}\vec{u}_1$ amb $\|\vec{u}_1\| = 1$, per tant $\vec{u}_1 = \frac{\pm 1}{\|v_1\|}v_1$, però com que r_{11} és positiva, $\vec{u}_1 = \frac{1}{\|v_1\|}v_1$ i $r_{11} = \|\vec{v}_1\|$.
- Suposem que les $j-1$ primeres columnes de Q i de R ja estan fixades, volem veure que tant sols hi ha una elecció per la j -èssima: el vector \vec{u}_j és un vector perpendicular a l'espai $V_{j-1} = \langle \vec{v}_1, \dots, \vec{v}_{j-1} \rangle$ i contingut a $V_j = \langle \vec{v}_1, \dots, \vec{v}_j \rangle$. Considerem V_{j-1}^\perp el complement ortogonal de V_{j-1} a V_j , que és de dimensió 1 ($\dim(V_{j-1}) + \dim(V_{j-1}^\perp) = \dim(V_j)$), i conté \vec{v}_j^\perp , per tant es té $V_{j-1}^\perp = \langle \vec{v}_j^\perp \rangle$. Llavors, com que ha de ser unitari, ja j -èssima columna de Q ha de ser $\vec{u}_j = \frac{\pm 1}{\|\vec{v}_j^\perp\|}\vec{v}_j^\perp$, però com que r_{jj} és positiva, $\vec{u}_j = \frac{1}{\|\vec{v}_j^\perp\|}\vec{v}_j^\perp$ i $r_{jj} = \|\vec{v}_j^\perp\|$, i tenim la unicitat de la \vec{u}_j i el coeficient r_{jj} . La resta de coeficients de la columna j d' R estan determinats per ser R triangular superior (per tant, per sota de r_{jj} són zero) i el fet de que els de sobre són les coordenades de $\vec{v}_j - \vec{v}_j^\perp$ en la base $[\vec{u}_1, \dots, \vec{u}_{j-1}]$ de V_{j-1} (i les coordenades en una base són úniques pel Teorema 3.36).

□

Exemple 5.14. Podem aprofitar els càlculs fets a l'Exemple 5.12 per trobar la factorització QR següent:

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

5.4 Aplicacions i matrius ortogonals

Definició 5.15. Diem que una aplicació lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és *ortogonal* si conserva la longitud dels vectors, o sigui, si

$$\|f(\vec{v})\| = \|\vec{v}\| \text{ per a tot } \vec{v} \in \mathbb{R}^n.$$

Si $f = f_A$ amb $A \in M_{m \times n}(\mathbb{R})$, direm que A és una *matriu ortogonal*.

Exemple 5.16. Les rotacions definides a la Secció 3.2.4 són ortogonals: es pot interpretar geomètricament o bé fent els càlculs: un vector $\begin{pmatrix} x \\ y \end{pmatrix}$ té norma

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2$$

i va a parar a:

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

que té norma

$$\begin{aligned} \|f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)\|^2 &= \left(\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right)^T \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \\ &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \\ &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 \end{aligned}$$

i per tant la rotació és ortogonal.

Exercici 5.1. Demostreu que reflexions definies a la Secció 3.2.3 són ortogonals.

Exemple 5.17. Les projeccions a un subespai $V \subset \mathbb{R}^n$ ($V \neq \mathbb{R}^n$) no són ortogonals: considerem $\vec{0} \neq \vec{u} \in V^\perp$, llavors $\text{Proj}_V(\vec{u}) = \vec{0}$ i per tant no es conserva la longitud de \vec{u} .

El lema següent ens permet calcular el producte escalar en funció dels mòduls:

Lema 5.18. Si \vec{u}, \vec{v} són vectors d' \mathbb{R}^n , llavors:

$$\vec{u} \cdot \vec{v} = \frac{1}{2} (\|\vec{u} + \vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2).$$

Demostració. Considerem $\vec{u}, \vec{v} \in \mathbb{R}^n$. Tenim que:

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

I per tant:

$$\vec{u} \cdot \vec{v} = \frac{1}{2} (\|\vec{u} + \vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2).$$

□

I d'aquí deduïm:

Teorema 5.19. Una aplicació lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és ortogonal si i només si $f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v}$ per a tot \vec{u} i \vec{v} de \mathbb{R}^n .

Demostració. Si l'aplicació és ortogonal, es compleix que, per a tot $\vec{u}, \vec{v} \in \mathbb{R}^n$, $\|f(\vec{u})\| = \|\vec{u}\|$ i $\|f(\vec{v})\| = \|\vec{v}\|$. Llavors:

$$\begin{aligned} f(\vec{u}) \cdot f(\vec{v}) &= \frac{1}{2} (\|f(\vec{u}) + f(\vec{v})\|^2 - \|f(\vec{u})\|^2 - \|f(\vec{v})\|^2) = \\ &= \frac{1}{2} (\|f(\vec{u} + \vec{v})\|^2 - \|f(\vec{u})\|^2 - \|f(\vec{v})\|^2) = \\ &= \frac{1}{2} (\|\vec{u} + \vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2) = \vec{u} \cdot \vec{v}, \end{aligned}$$

on hem utilitzat el Lema 5.18 dues vegades i que f és lineal.

Si $f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v}$ per a tot \vec{u} i \vec{v} de \mathbb{R}^n , com que $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$, tenim que, agafant $\vec{u} = \vec{v}$, $\|f(\vec{u})\|^2 = \|\vec{u}\|^2$ i per tant f és ortogonal. \square

Corol·lari 5.20. Si $f = f_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és una aplicació lineal i $\vec{e}_1, \dots, \vec{e}_n$ és la base estàndard de \mathbb{R}^n , llavors:

- (a) f és ortogonal si i només si $f(\vec{e}_1), \dots, f(\vec{e}_n)$ són vectors ortonormals. En particular, si f és ortogonal, $f(\vec{e}_1), \dots, f(\vec{e}_n)$ són linealment independents i si $n = m$, són una base ortonormal.
- (b) f_A és una aplicació lineal ortogonal si i només si A és una matriu ortogonal.

Demostració. Vegem primer (a). Si f és ortogonal, del fet que $f(\vec{e}_i) \cdot f(\vec{e}_j) = \vec{e}_i \cdot \vec{e}_j$, es dedueix que $f(\vec{e}_i)$ són unitaris i ortogonals dos a dos, per tant, ortonormals.

Suposem ara que $f(\vec{e}_i)$ són unitaris i ortogonals dos a dos i unitaris i $\vec{u} \in \mathbb{R}^n$. Tenim que existeixen nombres reals λ_i tals que $\vec{u} = \lambda_1 \vec{e}_1 + \dots + \lambda_n \vec{e}_n$. Llavors $\|\vec{u}\|^2 = \lambda_1^2 + \dots + \lambda_n^2$, i $\|f(\vec{u})\|^2 = \|\lambda_1 f(\vec{e}_1) + \dots + \lambda_n f(\vec{e}_n)\|^2 = \lambda_1^2 + \dots + \lambda_n^2$ (pel Teorema de Pitàgores, utilitzant que són ortogonals i unitaris).

Vegem ara que els $f(\vec{e}_i)$ són linealment independents: més en general, tenim que si $\vec{u}_1, \dots, \vec{u}_k$ són vectors ortonormals, llavors són linealment independents: del teorema de Pitàgores es dedueix que:

$$\|\lambda_1 \vec{u}_1 + \dots + \lambda_k \vec{u}_k\|^2 = \lambda_1^2 + \dots + \lambda_k^2$$

I per tant si tenim una combinació $\vec{0} = \lambda_1 \vec{u}_1 + \dots + \lambda_k \vec{u}_k$, tots els λ han de ser zero. Llavors tenim que, en el nostre cas, $f(\vec{e}_1), \dots, f(\vec{e}_n)$ són linealment independents.

Si $n = m$, $f(\vec{e}_1), \dots, f(\vec{e}_m)$ són m vectors de \mathbb{R}^m linealment independents, i per tant, base.

Per a veure (b), si f_A és ortogonal i fem el producte $A^T A$, tenim que el coeficient (i, j) és $f(\vec{e}_i) \cdot f(\vec{e}_j)$, per tant, com que són ortonormals, tenim $A^T A = \mathbf{1}_n$.

Les columnes d' A són els vectors $f(\vec{e}_j)$, per tant, $A^T A = \mathbf{1}_n$ si i només si $f(\vec{e}_1), \dots, f(\vec{e}_n)$ són ortogonals i unitaris (per tant f és ortogonal). \square

Teorema 5.21. Si $A, B \in M_n(\mathbb{R})$ són matrius ortogonals, llavors:

- (a) AB també és ortogonal i
- (b) A^{-1} és ortogonal (i $A^{-1} = A^T$).

Demostració. AB serà ortogonal si f_{AB} conserva la longitud dels vectors, però $f_{AB} = f_A \circ f_B$ i, per hipòtesi, tant f_B com f_A conserven la longitud, per tant, f_{AB} també.

Per a veure (b): si $\vec{v} = f_{A^{-1}}(\vec{w})$, llavors $\vec{w} = f_A(\vec{v})$ i, per hipòtesi $\|\vec{w}\| = \|f_A(\vec{v})\| = \|\vec{v}\|$, per tant A^{-1} és ortogonal. \square

5.5 Matriu d'una projecció ortogonal en una base ortonormal

Proposició 5.22. Si $V \subset \mathbb{R}^n$ un subespai vectorial, $[\vec{v}_1, \dots, \vec{v}_k]$ és una base ortonormal de V i Q és la matriu que té per columnes els vectors \vec{v}_j , llavors, la matriu de la projecció ortogonal a V és:

$$[\text{Proj}_V] = QQ^T.$$

Demostració. Com que $[\vec{v}_1, \dots, \vec{v}_k]$ és una base ortonormal de V , la projecció ortogonal a V del vector \vec{u} és:

$$\text{Proj}_V(\vec{u}) = (\vec{u} \cdot \vec{v}_1)\vec{v}_1 + \dots + (\vec{u} \cdot \vec{v}_k)\vec{v}_k = Q \begin{pmatrix} \vec{u} \cdot \vec{v}_1 \\ \vdots \\ \vec{u} \cdot \vec{v}_k \end{pmatrix} = Q \begin{pmatrix} \vec{v}_1^T \vec{u} \\ \vdots \\ \vec{v}_k^T \vec{u} \end{pmatrix} = Q \begin{pmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_k^T \end{pmatrix} \vec{u} = (QQ^T)\vec{u}$$

I per tant la matriu associada a la projecció és QQ^T . \square

Exemple 5.23. Considerem el vector $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ de \mathbb{R}^2 i calculem la matriu de la projecció a $V = \langle \begin{pmatrix} x \\ y \end{pmatrix} \rangle$. Una base ortonormal és $\vec{v}_1 = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} x \\ y \end{pmatrix}$ i per tant la matriu és (compareu-ho amb la Proposició 3.7):

$$[\text{Proj}_V] = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} x \\ y \end{pmatrix} \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} x & y \end{pmatrix} = \frac{1}{x^2+y^2} \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}.$$

Exemple 5.24. Volem calcular la matriu de la projecció ortogonal de \mathbb{R}^4 a $V = \langle \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \rangle \subset \mathbb{R}^4$. A l'Exemple 5.12 ja hem calculat una base ortonormal, i era $\mathcal{C} = [\begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}]$, per tant la matriu corresponent a la projecció ortogonal és:

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

5.6 Mínims quadrats

Aquesta secció suposarem que tenim un conjunt E i un subconjunt $V \subset E$ i que, a cada element d' E , volem assignar-li l'element de V que millor l'aproxima. Per això, suposem que E és un \mathbb{R} -espai vectorial i que V és un subespai. Llavors:

Proposició 5.25. Si E és un \mathbb{R} -espai vectorial i $V \subset E$ és un subespai, llavors, per a tot $\vec{u} \in E$ i $\vec{v} \in V$ es compleix:

$$\|\vec{u} - \text{Proj}_V(\vec{u})\| \leq \|\vec{u} - \vec{v}\|.$$

Demostració. Tenim que $\vec{u} - \vec{v} = (\vec{u} - \text{Proj}_V(\vec{u})) + (\text{Proj}_V(\vec{u}) - \vec{v})$ amb el primer vector de V^\perp i el segon de V . Aplicant el Teorema de Pitàgores tenim:

$$\|\vec{u} - \vec{v}\|^2 = \|\vec{u} - \text{Proj}_V(\vec{u})\|^2 + \|\text{Proj}_V(\vec{u}) - \vec{v}\|^2 \geq \|\vec{u} - \text{Proj}_V(\vec{u})\|^2,$$

i tenim la desigualtat que volem. \square

Per tant, considerant la mètrica de la distància, la millor aproximació en aquest cas és la projecció ortogonal.

5.6.1 Recta de regressió

Suposem que tenim un núvol de punts $(x_1, y_1), \dots, (x_k, y_k)$ de \mathbb{R}^2 que, en principi, no estan alineats. Suposem, a més, que hi ha una dependència del tipus $y = f(x)$ (per tant, al núvol de punts seria un problema que hi hagués dos punts amb la mateixa x i diferents y). Suposem que els punts estan *aproximadament* alineats i que el que volem trobar és la recta $y = a_0 + a_1x$ que aproxima millor aquests punts.

Una altra manera de pensar-ho és que volem resoldre el sistema d'equacions amb incògnites a_0 i a_1 :

$$\begin{aligned} a_0 + x_1 a_1 &= y_1 \\ &\vdots \\ a_0 + x_k a_1 &= y_k \end{aligned}$$

O, en forma matricial:

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$$

Com que molt probablement els punts no estan alineats, com a sistema d'equacions, és un sistema incompatible. Dit d'una altra manera, el vector $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$ no pertany al subespai

$V = \langle \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \rangle$. Si apliquem la Proposició 5.25, el vector de V que serà més proper a \vec{y} serà $\text{Proj}_V(\vec{y})$:

$$\text{Proj}_V \left(\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} \right) = a_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + a_1 \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}.$$

Podríem calcular a_0 i a_1 fent primer Gram-Schmidt per a trobar una base ortonormal de V i fent els productes escalars corresponents, o bé fent les consideracions següents: el vector

$$\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} - \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$$

és perpendicular a V , per tant ha de complir que:

$$\begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_k \end{pmatrix} \left(\begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} - \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

I per tant:

$$\begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_k \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_k \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_k \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix}$$

I queda el sistema compatible determinat:

$$\begin{pmatrix} k & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

Que es pot escriure: si $\bar{x} = \frac{1}{k} \sum x_i$, $\bar{y} = \frac{1}{k} \sum y_i$, $\overline{x^2} = \frac{1}{k} \sum x_i^2$ i $\overline{xy} = \frac{1}{k} \sum x_i y_i$:

$$\begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \overline{xy} \end{pmatrix}$$

I la solució es pot escriure com:

$$a_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - (\bar{x})^2} \text{ i } a_0 = \bar{y} - a_1\bar{x}. \quad (12)$$

Exemple 5.26. Considerem les notes següents d'un grup de 21 alumnes corresponent a una avaluació parcial i la nota final que van treure:

Parcial	Final	Parcial	Final	Parcial	Final
8.15	8.995	6.25	7.845	2.2	4.58
0.3	1.98	5.85	7.075	4.6	4.95
4.3	5.01	3.55	5.715	4.7	7.08
1.8	2.88	1.7	2.82	2.3	3.86
5.4	5.08	4.6	7.13	2.7	5.33
0.2	2.78	7.7	8.8	7.95	8.595
5.4	7.77	3.15	1.245	3	4.95

I volem aproximar la nota final a partir de la del parcial amb una recta:

$$\text{final} = a_0 + a_1 \text{parcial}$$

Per tant, podem aplicar la Fórmula (12) als valors (aproximem a 3 decimals):

$$\bar{x} = \overline{\text{parcial}} = 4.085, \quad \bar{y} = \overline{\text{final}} = 5.451, \quad \overline{x^2} = 21,839 \text{ i } \overline{xy} = 26,816.$$

I queda (amb 3 decimals):

$$\text{final} = 1.843 + 0.883 \text{parcial}.$$

La Figura 5 té una representació gràfica d'aquestes dades i de la recta de regressió. A més, l'equació de la recta es pot utilitzar per aproximar una nota final a partir d'una parcial: per exemple, un alumne que hagi tret un 4 al parcial, la predicció diria que trauria un 5,376 a la nota final.

5.6.2 Cas general

En el cas general, suposem que tenim un sistema d'equacions lineals:

$$Ax = b$$

amb $A \in M_{m \times n}(\mathbb{R})$ amb $m \geq n$ (més equacions que incògnites) i tal que $\text{Rang}(A) = n$. És molt probable que aquest sistema no tingui solució, i per aquests casos ens interessa tenir la millor aproximació:

Lema 5.27. *Amb les hipòtesis anteriors, el sistema*

$$A^T A x = A^T b$$

té solució única. Aquesta solució s'anomena solució per mínim quadrats del sistema $Ax = b$.

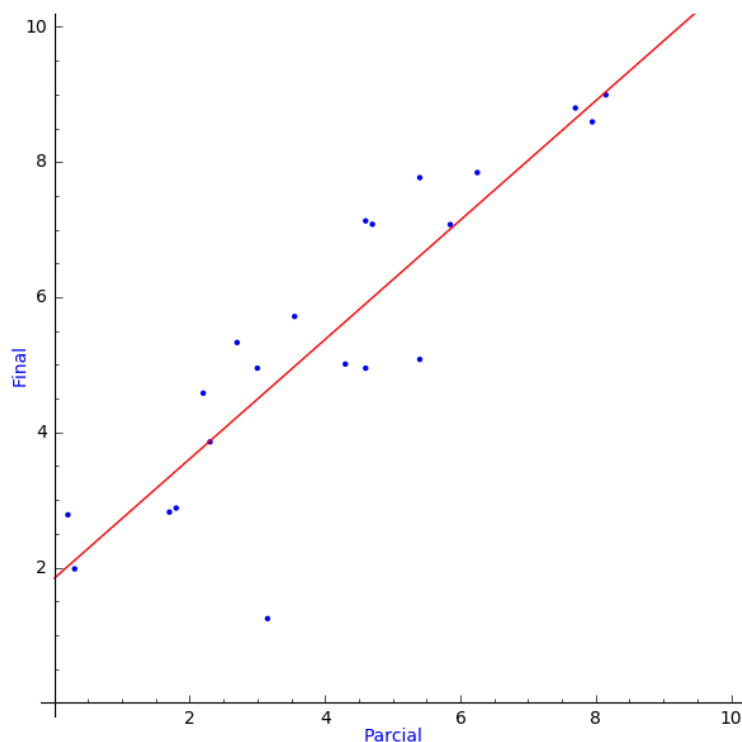


Figura 5: Recta de regressió

Demostració. Observem que $A^T A \in M_{n \times n}(\mathbb{R})$, pel que tant sols cal veure que $\text{Rang}(A^T A) = n$ (i llavors serà un sistema compatible determinat), o el que és el mateix, que l'aplicació lineal $f_{A^T A}$ és injectiva (llavors el sistema homogeni que té per matriu associada $A^T A$ serà compatible determinat i per tant $\text{Rang}(A^T A) = n$). Però si $\vec{u} \in \mathbb{R}^n$ compleix $A^T A \vec{u} = \vec{0}$, tenim $A^T(A\vec{u}) = \vec{0}$, per tant $A\vec{u}$ és solució del sistema homogeni que té per matriu associada A^T , que és de rang n , per tant $A\vec{u} = \vec{0}$. Aplicant ara que $\text{Rang}(A) = n$, tenim que $\vec{u} = \vec{0}$ i per tant $\text{Rang}(A^T A) = n$. \square

Lema 5.28. *Continuem amb les hipòtesis anteriors. De tots els vectors $\vec{v} \in \mathbb{R}^n$, la solució per mínims quadrats del sistema $Ax = b$ és la que és més propera a v en el sentit següent: si x és la solució del sistema $A^T Ax = A^T b$ i $v \in \mathbb{R}^n$, llavors*

$$\|Ax - b\|^2 \leq \|Av - b\|^2.$$

Demostració. Es dedueix del fet que la solució per mínims quadrats sigui la projecció ortogonal de b a l'espai generat per les columnes d' A , que, com que el rang d' A és n , queda caracteritzada per complir:

$$A^T(Ax - b) = \vec{0}.$$

 \square

D'aquí també es dedueix que la projecció del vector b a V , el subespai generat per les columnes d' A és, en la base de V formada per les columnes d' A és:

$$x = (A^T A)^{-1} A^T b$$

Per tant, el vector d' \mathbb{R}^n és:

$$Ax = A(A^T A)^{-1} A^T b$$

I per tant:

Proposició 5.29. Si $V \subset \mathbb{R}^n$ és un subespai vectorial i $[\vec{v}_1, \dots, \vec{v}_k]$ és una base de V , la projecció de \mathbb{R}^n a V ve donada per la matriu:

$$[\text{Proj}_V] = A(A^T A)^{-1} A^T$$

on A és la matriu que té per columnes els vectors \vec{v}_j .

Exemple 5.30. Suposem que tenim un grup de 24 estudiants als que s'ha fet una avaluació parcial i un lliurament a mig semestre amb les notes següents. La tercera columna diu quina ha sigut la nota final de l'assignatura (amb més avaluacions entremig):

Lliurament	Parcial	Final	Lliurament	Parcial	Final
9.25	9.75	8.25	9.5	5.2	7.27
9.5	2.4	3.66	9	7.1	8.32
8	1.6	5.65	8	4	7.61
8	7.1	6.71	8.5	3.4	6.77
8.5	5.7	7.77	8	6	7.74
6	2.9	6.5	10	5.9	8.37
9	9	8.13	10	2.6	8.13
8	7.6	7.15	7.5	1.6	5.21
9	6.2	6.8	10	9	8.52
10	10	9.07	9	8	7.76
9	6.5	7.43	9.75	10	8.71
8.5	4	6.39	9.5	7.4	8.57

Volem aproximar la nota final a partir de les dues que se saben a mig semestre mitjançant una fórmula:

$$\text{final} = a * \text{Lliurament} + b * \text{Parcial} + c$$

I per tant, fem mínim quadrats per trobar a, b i c . En aquest cas, el sistema que té moltes més equacions que incògnites seria:

$$\begin{pmatrix} 9.25 & 9.75 & 1 \\ 9.5 & 2.4 & 1 \\ 8 & 1.6 & 1 \\ \vdots & \vdots & \vdots \\ 9.5 & 7.4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 8.25 \\ 3.66 \\ 5.65 \\ \vdots \\ 8.57 \end{pmatrix}$$

I per tant hem de resoldre el sistema compatible determinat:

$$\begin{pmatrix} 9.25 & 9.5 & 8 & \dots & 9.5 \\ 9.75 & 2.4 & 1.6 & \dots & 7.4 \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 9.25 & 9.75 & 1 \\ 9.5 & 2.4 & 1 \\ 8 & 1.6 & 1 \\ \vdots & \vdots & \vdots \\ 9.5 & 7.4 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9.25 & 9.5 & 8 & \dots & 9.5 \\ 9.75 & 2.4 & 1.6 & \dots & 7.4 \\ 1 & 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} 8.25 \\ 3.66 \\ 5.65 \\ \vdots \\ 8.57 \end{pmatrix}$$

Quedant:

$$\begin{pmatrix} 1885.375 & 1287.5375 & 211.5 \\ 1287.5375 & 1015.0425 & 142.95 \\ 211.5 & 142.95 & 24 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1568.205 \\ 1108.1755 \\ 176.49 \end{pmatrix}$$

Que té per solució (amb 3 decimals):

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0.191 \\ 0.316 \\ 3.790 \end{pmatrix}$$

Per ant l'aproximació és:

$$\text{final} = 0.191 * \text{Lliurament} + 0.316 * \text{Parcial} + 3.79.$$

Per exemple, d'un alumne que tregui un 4 al lliurament i un altre 4 al parcial, aquest model diu que de l'assignatura acabés amb un 5.817

5.7 Formes bilineals i productes escalars

Definició 5.31. Donat un espai vectorial E sobre un cos \mathbb{K} , una *forma bilineal* és una aplicació

$$\phi: E \times E \rightarrow \mathbb{K}$$

tal que:

- $\phi(\vec{u}_1 + \vec{u}_2, \vec{v}) = \phi(\vec{u}_1, \vec{v}) + \phi(\vec{u}_2, \vec{v})$ per a tots \vec{u}_1, \vec{u}_2 i \vec{v} d' E ,
- $\phi(\lambda \vec{u}, \vec{v}) = \lambda \phi(\vec{u}, \vec{v})$ per a tots \vec{u} i \vec{v} d' E i $\lambda \in \mathbb{K}$,
- $\phi(\vec{u}, \vec{v}_1 + \vec{v}_2) = \phi(\vec{u}, \vec{v}_1) + \phi(\vec{u}, \vec{v}_2)$ per a tots \vec{u}, \vec{v}_1 i \vec{v}_2 d' E i
- $\phi(\vec{u}, \lambda \vec{v}) = \lambda \phi(\vec{u}, \vec{v})$ per a tots \vec{u} i \vec{v} d' E i $\lambda \in \mathbb{K}$.

A més diem que una aplicació bilineal ϕ és:

- *simètrica* si $\phi(\vec{u}, \vec{v}) = \phi(\vec{v}, \vec{u})$ per a tots \vec{u} i \vec{v} d' E ,
- *degenerada* si existeix $\vec{u} \neq \vec{0}$ tal que $\phi(\vec{u}, \vec{v}) = 0$ per a tot \vec{v} d' E , i
- *definida positiva* si $\phi(\vec{u}, \vec{u}) > 0$ per a tot $\vec{u} \neq \vec{0}$.

Definició 5.32. Si E és un \mathbb{K} -espai vectorial de dimensió finita i $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ és una base d' E , i ϕ és una aplicació bilineal, la *matriu de l'aplicació bilineal en la base \mathcal{B}* és:

$$[\phi]_{\mathcal{B}} = \begin{pmatrix} \phi(\vec{v}_1, \vec{v}_1) & \phi(\vec{v}_1, \vec{v}_2) & \cdots & \phi(\vec{v}_1, \vec{v}_n) \\ \phi(\vec{v}_2, \vec{v}_1) & \phi(\vec{v}_2, \vec{v}_2) & \cdots & \phi(\vec{v}_2, \vec{v}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\vec{v}_n, \vec{v}_1) & \phi(\vec{v}_n, \vec{v}_2) & \cdots & \phi(\vec{v}_n, \vec{v}_n) \end{pmatrix}.$$

Lema 5.33. Si ϕ és una aplicació bilineal definit sobre un \mathbb{K} -espai vectorial E amb base $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ i $[\phi]_{\mathcal{B}}$ és la matriu de ϕ en la base \mathcal{B} , llavors: si \vec{u} i \vec{w} són vectors d' E amb coordenades en la base \mathcal{B} :

$$[\vec{u}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad i \quad [\vec{w}]_{\mathcal{B}} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

llavors:

$$\phi(\vec{u}, \vec{w}) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \phi(\vec{v}_i, \vec{v}_j) = [\vec{u}]_{\mathcal{B}}^T [\phi]_{\mathcal{B}} [\vec{w}]_{\mathcal{B}}.$$

Demostració. La primera igualtat és per bilinealitat, i és igual a l'última expressió, que està escrita en forma matricial. \square

Lema 5.34. Si tenim $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ i $\mathcal{C} = [\vec{u}_1, \dots, \vec{u}_n]$ bases de \mathbb{K}^n , $[\text{Id}_n]_{\mathcal{C}, \mathcal{B}}$ la matriu del canvi de base (la que té per columnes les coordenades dels vectors \vec{u}_j en la base \mathcal{B}), i ϕ una aplicació bilineal, hi ha la relació següent entre les matrius de l'aplicació bilineal en cada base:

$$[\phi]_{\mathcal{C}} = [\text{Id}_n]_{\mathcal{C}, \mathcal{B}}^T [\phi]_{\mathcal{B}} [\text{Id}_n]_{\mathcal{C}, \mathcal{B}}$$

Demostració. Les matrius de $[\phi]_{\mathcal{B}}$ i $[\phi]_{\mathcal{C}}$ han de complir que, per a tot \vec{u} i \vec{v} de \mathbb{K}^n es compleixi:

$$[\vec{u}]_{\mathcal{C}}^T [\phi]_{\mathcal{C}} [\vec{v}]_{\mathcal{C}} = \phi(\vec{u}, \vec{v}) = [\vec{u}]_{\mathcal{B}}^T [\phi]_{\mathcal{B}} [\vec{v}]_{\mathcal{B}}. \quad (13)$$

Però, per les propietats de la matriu $[\text{Id}]_{\mathcal{C}, \mathcal{B}}$ tenim:

$$[\vec{u}]_{\mathcal{B}} = [\text{Id}]_{\mathcal{C}, \mathcal{B}} [\vec{u}]_{\mathcal{C}},$$

i si ho substituïm a l'Equació (13), tenim, que per a tot \vec{u} i \vec{v} de \mathbb{K}^n :

$$[\vec{u}]_{\mathcal{C}}^T [\phi]_{\mathcal{C}} [\vec{v}]_{\mathcal{C}} = ([\text{Id}]_{\mathcal{C}, \mathcal{B}} [\vec{u}]_{\mathcal{C}})^T [\phi]_{\mathcal{B}} ([\text{Id}]_{\mathcal{C}, \mathcal{B}} [\vec{v}]_{\mathcal{C}}) = [\vec{u}]_{\mathcal{C}}^T ([\text{Id}]_{\mathcal{C}, \mathcal{B}}^T [\phi]_{\mathcal{B}} [\text{Id}]_{\mathcal{C}, \mathcal{B}}) [\vec{v}]_{\mathcal{C}},$$

pel que tenim la igualtat de matrius que volem. \square

Exemple 5.35. Considerem la forma bilineal simètrica $\phi: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ donada per la matriu (en la base estàndard $\mathcal{B} = [\vec{e}_1, \vec{e}_2]$)

$$[\phi] = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$$

I volem calcular la matriu de ϕ en la base $\mathcal{C} = [(\frac{1}{1}), (\frac{1}{-1})]$. Llavors hem de calcular:

$$[\text{Id}]_{\mathcal{C}, \mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

I per tant la matriu

$$[\phi]_{\mathcal{C}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^T \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Per tant, les matrius $\begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$ i $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ representen la mateixa forma bilineal en bases diferents.

Mirem ara un cas particular de forma bilineal:

Definició 5.36. Un *producte escalar* sobre un \mathbb{R} -espai vectorial E és una forma bilineal simètrica definida positiva. En general, escriurem el resultat amb un punt: $\vec{u} \cdot \vec{v}$.

Si \mathcal{B} és una base finita d'un espai vectorial E i \cdot és un producte escalar, la matriu $[\cdot]_{\mathcal{B}}$ serà simètrica.

Exemple 5.37. Si considerem \mathbb{R}^n i $\mathcal{B} = [\vec{e}_1, \dots, \vec{e}_n]$ la base estàndard, el producte escalar definit com $\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v}$ és un producte escalar amb matriu la matriu identitat $\mathbf{1}_n$.

Exemple 5.38. Si $E = C^\infty(\mathbb{R})$ definim el producte escalar:

$$f \cdot g = \int_{-1}^1 f(x)g(x)dx.$$

Com que a $C^\infty(\mathbb{R})$ no hi ha una base finita, no té sentit parlar de la matriu del producte escalar.

Exemple 5.39. Considerem els nombres complexos \mathbb{C} . Definits com \mathbb{R} espai vectorial, tenen dimensió 2 i una base és $[1, i]$, i d'aquí deduïm l'estructura de la suma. A més, però, es poden multiplicar, i la multiplicació ve caracteritzada pel fet $i^2 = -1$ i que sigui commutativa i distributiva amb la suma: o sigui,

$$\mathbb{C} = \{a + bi \mid a \in \mathbb{R}, b \in \mathbb{R}\}$$

amb les operacions:

$$(a + bi) + (c + di) = (a + c) + (b + d)i \text{ i } (a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Veiem que $\mathbb{R} \subset \mathbb{C}$, considerant els nombres de la forma $a + 0i$.

Si $z = a + bi$, amb a i b reals, és un nombre complex, definim el *conjugat de z* i escrivim \bar{z} com:

$$\bar{z} = a - bi.$$

Veiem que:

- $z\bar{z} = a^2 + b^2 \in \mathbb{R}$,
- $z = \bar{z} \iff z \in \mathbb{R}$,
- $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ i
- $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

Considerem ara \mathbb{C}^n com a \mathbb{C} -espai vectorial de dimensió n (suma coordenada a coordenada i multiplicació per un escalar a totes les coordenades). Definim la forma bilineal següent:

$$\text{si } \vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \text{ i } \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \text{ llavors } \vec{v} \cdot \vec{w} = \sum_{i=1}^n v_i \bar{w}_i = \vec{v}^T \vec{\bar{w}} \text{ on } \vec{\bar{w}} = \begin{pmatrix} \bar{w}_1 \\ \vdots \\ \bar{w}_n \end{pmatrix}.$$

Que compleix, a més de les propietats de ser forma bilineal:

- $\vec{v} \cdot \vec{\bar{w}} = \overline{\vec{w} \cdot \vec{v}}$ per a tot $\vec{v}, \vec{w} \in \mathbb{C}^n$ (diem que és hermitica) i
- $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2 \in \mathbb{R}$ per a tot $\vec{v} \in \mathbb{C}^n$.

5.8 Tota matriu simètrica sobre \mathbb{R} diagonalitza

El resultat que volem demostrar a aquest apartat és:

Teorema 5.40 (Teorema espectral). *Una matriu $A \in M_n(\mathbb{R})$ diagonalitza en una base ortonormal si, i només si, A és simètrica.*

Per demostrar el teorema espectral necessitem utilitzar el *Teorema fonamental de l'àlgebra*, que no demostrarem aquí perquè la seva demostració surt dels objectius del curs:

Teorema 5.41 (Teorema fonamental de l'àlgebra). *Si $p(x) \in \mathbb{C}[x]$ ($p(x)$ és un polinomi amb coeficients a \mathbb{C}) de grau ≥ 1 , existeix $z \in \mathbb{C}$ tal que $p(z) = 0$.*

Demostració del teorema espectral. Una implicació és directa: si A diagonalitza en una base ortonormal vol dir que podem escriure:

$$A = QDQ^T$$

amb D una matriu diagonal i Q una matriu ortogonal ($Q^{-1} = Q^T$). Per tant:

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QD^T Q^T = A$$

i A és simètrica.

Per demostrar l'altra implicació, ho farem per inducció sobre n :

- Per a $n = 1$, tota matriu és simètrica i diagonal.
- Suposem cert per a $n - 1$, i volem demostrar-ho per n : considerem $A \in M_n(\mathbb{R})$ i $p_A(x)$ el polinomi característic. Pensem $p_A(x) \in \mathbb{C}[x]$ i, pel teorema fonamental de l'àlgebra, existeix $\lambda \in \mathbb{C}$ tal que $p_A(\lambda) = 0$. També tindrem un vector propi $\vec{v} \in \mathbb{C}^n$ tal que $A\vec{v} = \lambda\vec{v}$.

Considerem ara la forma bilineal hermítica definida a l'Exemple 5.39:

$$A\vec{v} \cdot \vec{v} = \lambda\vec{v} \cdot \vec{v} = \lambda\|\vec{v}\|^2 \text{ amb } 0 \neq \|\vec{v}\|^2 \in \mathbb{R}.$$

Però també:

$$A\vec{v} \cdot \vec{v} = (A\vec{v})^T \vec{v} = \vec{v}^T A^T \vec{v} = \vec{v}^T (\overline{A^T \vec{v}}) = \vec{v}^T A \vec{v} = \vec{v}^T (\overline{\lambda \vec{v}}) = \overline{\lambda} \|\vec{v}\|^2,$$

on hem utilitzat que $A = A^T$ (A és simètrica) i $\overline{A} = A$ (A té coeficients reals).

Per tant, tenim que $\lambda = \overline{\lambda}$ i obtenim $\lambda \in \mathbb{R}$.

Com que $p_A(\lambda) = 0$, tenim que existeix un vector propi real $\vec{v} \neq \vec{0}$ tal que $A\vec{v} = \lambda\vec{v}$. Podem considerar que $\|\vec{v}\| = 1$ (si cal, substituïm \vec{v} per $\frac{1}{\|\vec{v}\|}\vec{v}$).

Ampliem \vec{v} amb $\vec{v}_2, \dots, \vec{v}_n$ de tal manera que $[\vec{v}, \vec{v}_2, \dots, \vec{v}_n]$ sigui una base ortonormal (primer ampliem fins a base, i després apliquem Gram-Schmidt). Si considerem Q la matriu que té per columnes els vectors $\vec{v}, \vec{v}_2, \dots, \vec{v}_n$, tenim:

$$QAQ^T = \begin{pmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

on hem utilitzat que $Q^{-1} = Q^T$. Però com que $A = A^T$, tenim $(QAQ^T)^T = QAQ^T$, i per tant $b_{12} = b_{13} = \cdots = b_{1n} = 0$.

Ara restringim l'aplicació f_A a l'espai $\langle v_2, \dots, v_n \rangle$, que en la base $[v_2, \dots, v_n]$ té per matriu

$$B = \begin{pmatrix} b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

que és simètrica i $B \in M_{n-1}(\mathbb{R})$. Per hipòtesis d'inducció, B diagonalitza en una base ortonormal, pel que existeix $P \in M_{n-1}(\mathbb{R})$ matriu ortogonal i D matriu diagonal tals que $B = P^T D P$. Amb això tenim:

$$QAQ^T = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & P^T D P \end{array} \right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P^T \end{array} \right) \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & D \end{array} \right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array} \right)$$

Per tant, considerem

$$Q' = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P \end{array} \right) Q \text{ i } D' = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & D \end{array} \right)$$

i tenim que Q' és una matriu ortogonal i $A = Q'^{-1} D' Q'$, amb D' una matriu diagonal, pel que A diagonalitza a una base ortonormal.

□

5.9 Descomposició en valors singulars

A aquest apartat resoldrem la pregunta següent amb les eines que hem vist: si $f = f_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és una aplicació lineal, existeix una base ortonormal $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ d' \mathbb{R}^n tal que $f(\vec{v}_1), \dots, f(\vec{v}_n)$ siguin ortogonals?

Veurem que sí:

Teorema 5.42. $f = f_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és una aplicació lineal, llavors existeix una base ortonormal $\mathcal{B} = [\vec{v}_1, \dots, \vec{v}_n]$ d' \mathbb{R}^n tal que $f(\vec{v}_1), \dots, f(\vec{v}_n)$ tal que són ortogonals.

Demostració. Considerem $f = f_A$, i per tant $A \in M_{m \times n}(\mathbb{R})$ la matriu tal que $f(\vec{v}) = A\vec{v}$. Llavors, la matriu $A^T A \in M_n(\mathbb{R})$ és simètrica, i per tant diagonalitza en una base ortonormal, per tant, existeix una base ortonormal $\mathcal{B}[\vec{v}_1, \dots, \vec{v}_n]$ de \mathbb{R}^n i escalars $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tals que

$$A^T A \vec{v}_i = \lambda_i \vec{v}_i \text{ per a tot } i = 1, \dots, n.$$

Volem veure que $f(\vec{v}_1), \dots, f(\vec{v}_n)$ són ortogonals, o sigui, $f(\vec{v}_i) \cdot f(\vec{v}_j) = 0$ si $i \neq j$:

$$\begin{aligned} f(\vec{v}_i) \cdot f(\vec{v}_j) &= (A\vec{v}_i) \cdot (A\vec{v}_j) = (A\vec{v}_i)^T (A\vec{v}_j) = (\vec{v}_i^T A^T) (A\vec{v}_j) = \\ &= \vec{v}_i^T (A^T A \vec{v}_j) = \vec{v}_i^T (\lambda_j \vec{v}_j) = \lambda_j (\vec{v}_i \cdot \vec{v}_j) = 0. \end{aligned}$$

□

Observació 5.43. El Teorema 5.42 no diu que \mathcal{B} sigui única, i de fet, en general, no ho és: considereu l'aplicació identitat de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, que envia qualsevol base ortonormal a vectors ortogonals.

La interpretació geomètrica a \mathbb{R}^2 és la següent: la imatge d'una circumferència de radi 1 centrada a l'origen és una el·lipse, els vectors propis d' $A^T A$ corresponen a les direccions principals de l'el·lipse, i les longituds dels semieixos principals són les arrels quadrades dels valors propis d' $A^T A$. Posem nom a aquests valors:

Definició 5.44. Donada una matriu $A \in M_{m \times n}(\mathbb{R})$, definim els *valors singulars* d' A com les arrels quadrades dels valors propis de la matriu $A^T A$: $\sigma_1, \dots, \sigma_m$ positius tals que σ_i^2 és valor propi d' $A^T A$.

Observació 5.45. σ_i és el mòdul $\|f(\vec{v}_i)\|$ a l'enunciat del Teorema 5.42.

Proposició 5.46. Si $f = f_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ és una aplicació lineal i $\text{Rang}(A) = r$, llavors hi ha r valors singulars diferent de zero i els $n - r$ restants valen zero.

Demostració. Tenim que en la base \mathcal{B} del Teorema 5.42, f_A té per matriu $[f_A]_{\mathcal{B}}$, una matriu amb σ_i a la diagonal i zeros a la resta. Llavors, com que el rang és la dimensió de la imatge, el rang de f_A és igual al de $[f_A]_{\mathcal{B}}$, que és el nombre de σ_i no nuls, i la resta han de ser zero. □

Podem aprofitar aquests raonaments per demostrar:

Teorema 5.47. Si $A \in M_{m \times n}(\mathbb{R})$, llavors existeixen $P \in M_m(\mathbb{R})$ ortogonal, $D \in M_{m \times n}(\mathbb{R})$ amb els únics elements no nuls a la diagonal i positius, i $Q \in M_n(\mathbb{R})$ ortogonal tals que:

$$A = PDQ.$$

A més, podem ordenar les entrades diagonals de D de més gran a més petita. Anomenem a aquesta descomposició com una descomposició en valors singulars d' A .

Demostració. Considerem l'aplicació lineal f_A i veiem que necessitem bases ortonormals \mathcal{B} d' \mathbb{R}^n i \mathcal{C} d' \mathbb{R}^m tals que $D = [f_A]_{\mathcal{B}, \mathcal{C}}$ sigui diagonal.

El Teorema 5.42 ens dóna una base \mathcal{B} d' \mathbb{R}^n que és ortonormal. Ordenem aquesta base de tal manera que $\|f(\vec{v}_i)\| \geq \|f(\vec{v}_{i+1})\|$. Sigui $Q^T = Q^{-1}$ la matriu que té per columnes aquesta base.

Els vectors $f(\vec{v}_1), \dots, f(\vec{v}_n)$ són ortogonals. Suposem que els k primers són no nuls, llavors, considerant: $\vec{w}_i = \frac{1}{\|f(\vec{v}_i)\|}$ per a $1 \leq i \leq k$, tindrem vectors ortonormals d' \mathbb{R}^m , per tant linealment independents, i que podem ampliar amb $\vec{w}_{k+1}, \dots, \vec{w}_m$ fins a tenir una base ortonormal \mathcal{C} d' \mathbb{R}^m . Sigui P la matriu que té per columnes aquests vectors.

Com que $f(\vec{v}_i) = \|f(\vec{v}_i)\|\vec{w}_i$ si $i \leq k$ i $f(\vec{v}_i) = \vec{0}$ si $i > k$, tenim que $[f]_{\mathcal{B}, \mathcal{C}}$ és diagonal amb els valors singulars d' A a la diagonal ordenats de més gran a més petit. \square

Exemple 5.48. Considerem l'aplicació lineal $f_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ amb:

$$A = \begin{pmatrix} 3 & -9/5 \\ -1 & 13/5 \end{pmatrix}$$

i volem fer-ne una descomposició en valors singulars.

Calculem primer $A^T A$:

$$A^T A = \begin{pmatrix} 10 & -8 \\ -8 & 10 \end{pmatrix}$$

I la diagonalitzem. El polinomi característic és:

$$p_{A^T A}(x) = x^2 - 20x + 36 = (x - 18)(x - 2)$$

Per tant, els valors propis d' $A^T A$ són $\lambda_1 = 18$ i $\lambda_2 = 2$ i els corresponents vectors propis:

$$\text{Ker}(A^T A - 18\mathbf{1}_2) = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle \text{ i } \text{Ker}(A^T A - 2\mathbf{1}_2) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle$$

I com que els hem de considerar unitaris, obtenim:

$$Q^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \text{ i per tant } Q = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

La matriu diagonal són els valors singulars, que són les arrels quadrades dels valors propis d' $A^T A$, per tant són $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ i $\sigma_2 = \sqrt{2}$ i tenim:

$$D = \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}.$$

I la matriu P té per columnes les imatges de $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ i $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ per f_A dividides per les seves normes, per tant:

$$\begin{pmatrix} 3 & -9/5 \\ -1 & 13/5 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 24/5 \\ -18/5 \end{pmatrix} = 6 \begin{pmatrix} 4/5 \\ -3/5 \end{pmatrix}$$

i

$$\begin{pmatrix} 3 & -9/5 \\ -1 & 13/5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 8/5 \end{pmatrix} = 2 \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}$$

Per tant:

$$P = \begin{pmatrix} 4/5 & 3/5 \\ -3/5 & 4/5 \end{pmatrix}$$

I una descomposició d' A en valors singulars és:

$$\begin{pmatrix} 3 & -9/5 \\ -1 & 13/5 \end{pmatrix} = \begin{pmatrix} 4/5 & 3/5 \\ -3/5 & 4/5 \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

A la Figura 6 podem veure com es modifica la circumferència unitat quan apliquem la matriu A . Veiem que a la direcció $\begin{pmatrix} 4 \\ -3 \end{pmatrix}$ hi ha el primer eix principal amb semilongitud $3\sqrt{2}$ i a la direcció $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$ l'altre eix, amb semilongitud $\sqrt{2}$.

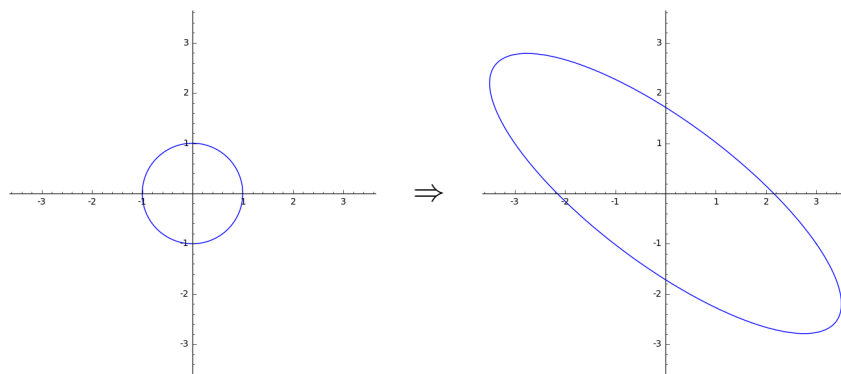


Figura 6: Deformació de la circumferència unitat per l'aplicació lineal f_A .

5.10 Classificació de formes bilineals simètriques sobre \mathbb{R}^n

A aquesta secció l'objectiu és classificar les formes bilineals. Definim abans què vol dir que siguin equivalents:

Definició 5.49. Diem que dues formes bilineals $\phi_1, \phi_2: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$ amb matrius corresponents $[\phi_1]$ i $[\phi_2]$ són *equivalents* si existeix una base \mathcal{C} de \mathbb{K}^n tal que

$$[\phi_1] = [\phi_2]_{\mathcal{C}}.$$

Dit d'una altra manera, si existeix una matriu invertible $\mathcal{S} \in M_n(\mathbb{K})$ (la matriu del canvi de base, que té per columnes els vectors de \mathcal{C}) tal que

$$[\phi_1] = \mathcal{S}^T [\phi_2] \mathcal{S}.$$

Escriurem $\phi_1 \sim \phi_2$.

Lema 5.50. Ser equivalents com a formes bilineals de $\mathbb{K}^n \times \mathbb{K}^n$ a \mathbb{K} és una relació d'equivalència. O sigui:

- $\phi \sim \phi$ per a tot ϕ forma bilineal (reflexiva),
- $\phi_1 \sim \phi_2 \Leftrightarrow \phi_2 \sim \phi_1$ per a totes ϕ_1, ϕ_2 formes bilineals (simètrica) i
- $\phi_1 \sim \phi_2$ i $\phi_2 \sim \phi_3$ implica $\phi_1 \sim \phi_3$ per a totes ϕ_1, ϕ_2, ϕ_3 formes quadràtiques (transitiva).

Demostració. Cal veure en cada cas quina és la matriu \mathcal{S} del canvi de base:

- Per veure que $\phi \sim \phi$, considerem $\mathcal{S} = \mathbf{1}_n$,
- Si $\phi_1 \sim \phi_2$ tenim que existeix \mathcal{S} una matriu de canvi de base (i per tant invertible) tal que:

$$[\phi_1] = \mathcal{S}^T [\phi_2] \mathcal{S}.$$

Però aquesta igualtat també es pot escriure com:

$$[\phi_2] = (\mathcal{S}^{-1})^T [\phi_1] \mathcal{S}^{-1}$$

i per tant $\phi_2 \sim \phi_1$.

- Per hipòtesis tenim que existeixen \mathcal{S}_1 i \mathcal{S}_2 tals que:

$$[\phi_1] = \mathcal{S}_1^T [\phi_2] \mathcal{S}_1 \text{ i } [\phi_2] = \mathcal{S}_2^T [\phi_3] \mathcal{S}_2.$$

Llavors:

$$[\phi_1] = \mathcal{S}_1^T \mathcal{S}_2^T [\phi_3] \mathcal{S}_2 \mathcal{S}_1 = (\mathcal{S}_2 \mathcal{S}_1)^T [\phi_3] (\mathcal{S}_2 \mathcal{S}_1)$$

i tenim $\phi_1 \sim \phi_3$.

□

A partir d'ara considerem que $\mathbb{K} = \mathbb{R}$ i l'objectiu és classificar les formes bilineals $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$:

Teorema 5.51. *Tota forma bilineal simètrica $\phi: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ és equivalent a una forma bilineal que té per matriu una matriu diagonal amb coeficients (a la diagonal): 1 a les r primeres files, -1 a les s següents i 0 a les t últimes ($n = r + s + t$). O sigui una matriu de la forma:*

$$\begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & \ddots & & & \cdots & & & \cdots & \\ 0 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ & & & & \ddots & & & \cdots & \\ 0 & \cdots & 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & & \cdots & & & \ddots & \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

A més, si definim la parella (r, s) com la signatura de ϕ ($\text{Sign}(\phi)$), dues formes bilineals bilineals $\phi_1, \phi_2: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ són equivalents si i només si $\text{Sign}(\phi_1) = \text{Sign}(\phi_2)$.

Demostració. Considerem $[\phi]$ la matriu simètrica $n \times n$ sobre \mathbb{R} de la forma bilineal ϕ . Pel Teorema espectral, existeix una matriu ortogonal Q i una matriu diagonal D amb valors $\lambda_1, \dots, \lambda_n$ tal que:

$$D = Q^T[\phi]Q.$$

Ara fem els canvis següents:

- Podem reordenar els elements de la diagonal de D reordenant les columnes de Q : si Q' és la matriu que resulta d'intercanviar les columnes j i k de Q , llavors, Q' continua sent ortogonal i el producte:

$$(Q')^T[\phi]Q'$$

també és diagonal, però intercanviant les posicions i i j , o sigui, els valors λ_i i λ_j .

Aprofitem aquesta propietat per reordenar la diagonal de D i posar primer els $\lambda_i > 0$, llavors els $\lambda_i < 0$ i finalment els $\lambda_i = 0$. Per tant, podem suposar que tenim Q una matriu ortogonal tal que

$$D = Q^T[\phi]Q.$$

amb D una matriu diagonal tal que a la diagonal té els r primers coeficients positius, els s següents negatius i els t últims zero.

- Considerem A la matriu diagonal següent, definida a partir dels coeficients de D : el coeficient a_{ii} es defineix com:

$$a_{ii} = \begin{cases} 1/\sqrt{\lambda_i} & \text{si } 1 \leq i \leq r, \\ 1/\sqrt{-\lambda_i} & \text{si } r < i \leq r + s, \\ 1 & \text{si } r + s < i \leq n. \end{cases}$$

Tenim que $A^T D A$ és una matriu diagonal amb r uns a les primeres files, s menys uns a les següents i zero a les últimes.

Per tant:

$$[\phi] \sim A^T Q^T[\phi]Q A = (Q A)^T[\phi]Q A$$

i $(Q A)^T[\phi]Q A$ és diagonal i com diu l'enunciat del teorema.

De moment hem vist que, com que ser equivalent a té la propietat transitiva, si ϕ_1 i ϕ_2 tenen la mateixa signatura, les dues són equivalents a una mateixa forma bilineal i per tant $\phi_1 \sim \phi_2$. Cal veure el recíproc, o el que és equivalent, que dues matrius diagonals equivalents D i D' amb (r, s, t) i (r', s', t') coeficients 1, -1 i 0 respectivament, llavors $(r, s, t) = (r', s', t')$.

Com que D i D' són equivalents, existeix A una matriu invertible que té per columnes v_i tal que:

$$D' = A^T D A.$$

Si fem el producte de la dreta, les r primeres files són $\|v_1\|^2$ (D coincideix amb la identitat a les r primeres files i columnes), \dots , $\|v_r\|^2$, per tant són números positius, per tant $r' \geq r$ i argumentant canviant els papers de D i D' (i A per A^{-1}), tenim $r \geq r'$, pel que $r = r'$.

$r + s$ és el rang de D . Com que D' és el producte de D per matrius invertibles, $r + s$ també és el rang de D' , que és $r' + s'$, per tant, $s = s'$.

Finalment, $t = n - (r + s) = n - (r' + s') = t'$. □

Exemple 5.52. Considerem la forma bilineal $\phi: \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{R}$ donada per la matriu:

$$[\phi] = \begin{pmatrix} 1 & 5 & -1 & 3 \\ 5 & 1 & 3 & -1 \\ -1 & 3 & 1 & 5 \\ 3 & -1 & 5 & 1 \end{pmatrix}.$$

Una manera de classificar-la és de calcular el polinomi característic i estudiar el signe dels valors on s'anul·la:

$$p_{[\phi]}(x) = \det(A - x\mathbf{1}_4) = x^4 - 4x^3 - 64x^2 + 256x.$$

I $p_{[\phi]}(x)$ s'anul·la als valors $\{-8, 0, 4, 8\}$, per tant té signatura $(2, 1)$ i és equivalent a la forma bilineal que té per matriu:

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

Acabem aquesta secció aprofitant aquesta classificació per fer la corresponent de les formes quadràtiques.

Definició 5.53. Una *forma quadràtica* sobre un cos \mathbb{K} és una aplicació $q: \mathbb{K}^n \rightarrow \mathbb{K}$ que es pot expressar com:

$$q(x_1, \dots, x_n) = \sum_{1 \leq i, j \leq n} \lambda_{ij} x_i x_j$$

amb $\lambda_{ij} \in \mathbb{K}$.

També es pot escriure com:

$$q(x_1, \dots, x_n) = \vec{x}^T A \vec{x},$$

on $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ i A és la matriu (simètrica) que té per coeficients

$$a_{ij} = \begin{cases} \lambda_{ii} & \text{si } i = j, \\ \frac{\lambda_{ij} + \lambda_{ji}}{2} & \text{si } i \neq j. \end{cases}$$

Observació 5.54. Amb el que hem vist, tenim una bijecció entre les matrius simètriques (que es poden pensar com formes bilineals en una base donada) i les formes quadràtiques. Escriurem $[q]$ per denotar la matriu simètrica corresponent a la forma quadràtica q .

Exemple 5.55. La forma quadràtica $q: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida per $q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$ correspon a la matriu simètrica

$$\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

Podem recuperar la forma quadràtica fent:

$$q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1^2 + bx_1x_2 + cx_2^2.$$

Exemple 5.56. Considerem les formes quadràtiques q_1 i q_2 de \mathbb{R}^2 a \mathbb{R} definides com $q_1(x_1, x_2) = x_1x_2$ i $q_2(y_1, y_2) = y_1^2 - y_2^2$. Observem que si fem el canvi de base:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

es poden escriure:

$$y_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 \text{ i } y_2 = \frac{1}{2}x_1 - \frac{1}{2}x_2$$

i tenim

$$q_2(y_1, y_2) = \left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right)^2 - \left(\frac{1}{2}x_1 - \frac{1}{2}x_2\right)^2 = x_1x_2.$$

Per tant, tenen la mateixa forma. Aquesta igualtat també es pot veure en forma matricial observant que si \mathcal{S} és la matriu del canvi de base, tenim $[q_1] = \mathcal{S}^T[q_2]\mathcal{S}$, on $[q_1]$ i $[q_2]$ són les matrius simètriques corresponents a les formes bilineals q_1 i q_2 respectivament:

$$\begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

Definició 5.57. Diem que dues formes quadràtiques $q_1, q_2: \mathbb{K}^n \rightarrow \mathbb{K}$ són *equivalents*, amb matrius si existeix un canvi de base \mathcal{S} a \mathbb{K}^n tal que

$$[q_1] = \mathcal{S}^T[q_2]\mathcal{S}.$$

Escriurem $q_1 \sim q_2$.

Com que els canvis de base afecten de la mateixa manera que afectaven a les formes bilineals (veure Lema 5.34), obtenim que “*ser equivalent a*” també es una relació d’equivalència a les formes quadràtiques i l’anàleg al Teorema 5.51:

Teorema 5.58. Tota forma quadràtica $q: \mathbb{R}^n \rightarrow \mathbb{R}$ és equivalent a una forma bilineal del tipus:

$$q_1(x_1, \dots, x_n) = x_1^2 + \dots + x_r^2 - x_{r+1}^2 - \dots - x_{r+s}^2.$$

A més, si definim la parella (r, s) com la signatura de q ($\text{Sign}(q)$), dues formes quadràtiques $q_1, q_2: \mathbb{R}^n \rightarrow \mathbb{R}$ són equivalents si i només si $\text{Sign}(q_1) = \text{Sign}(q_2)$.

Exemple 5.59. Classifiquem la forma quadràtica $q: \mathbb{R}^4 \rightarrow \mathbb{R}$:

$$q(x_1, x_2, x_3, x_4) = x_1^2 + 10x_1x_2 - 2x_1x_3 + 6x_1x_4 + x_2^2 - 6x_2x_3 - 2x_2x_4 + x_3^3 + 10x_3x_4 - x_4^2.$$

Que té per matriu:

$$\begin{pmatrix} 1 & 5 & -1 & 3 \\ 5 & 1 & 3 & -1 \\ -1 & 3 & 1 & 5 \\ 3 & -1 & 5 & 1 \end{pmatrix}$$

Per tant, podem aprofitar els càlculs de l’Exemple 5.52, i tenim que és equivalent a:

$$q(y_1, y_2, y_3, y_4) = y_1^2 + y_2^2 - y_3^2.$$

Exercicis recomanats

Els exercicis que segueixen són útils per practicar el material presentat. La numeració és la de [1].

Secció 5.1: 12, 16, 18.

Secció 5.2: 14, 34, 38.

Secció 5.3: 5-11, 13-20, 32.

Secció 5.4: 2, 8, 10.

Secció 5.5: 4, 10, 16, 22.

Secció 8.1: 6, 12, 16, 22.

Secció 8.2: 4, 10, 18, 22.

Secció 8.3: 6, 16, 18, 20.

A Nombres complexos

Podem definir els nombres complexos com

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}.$$

Com a espai vectorial sobre \mathbb{R} , podem identificar \mathbb{C} amb \mathbb{R}^2 , i per tant tenim definida una suma i una producte per reals:

$$(a + bi) + (c + di) = (a + c) + (b + d)i, \quad \lambda(a + bi) = \lambda a + \lambda bi.$$

Tenim però una operació addicional, el producte de nombres complexos:

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i,$$

on la fórmula es recorda fàcilment si tenim en compte que $i^2 = -1$, i fent servir la propietat distributiva.

Definim la *part real* i la *part imaginària* d'un nombre complex com:

$$\Re(a + bi) = a, \quad \Im(a + bi) = b.$$

També el *conjugat* d'un nombre complex es defineix com:

$$\overline{(a + bi)} = a - bi,$$

i per tant tenim les fórmules

$$\Re(z) = \frac{z + \bar{z}}{2}, \quad \Im(z) = \frac{z - \bar{z}}{2i}, \quad \forall z \in \mathbb{C}.$$

La *norma* d'un element $z = a + bi$ és

$$N(z) = z\bar{z} = (a + bi)(a - bi) = a^2 + b^2 \geq 0.$$

Observem que $z\bar{z} \geq 0$, i que $z\bar{z} = 0$ si i només si $z = 0$. Això ens permet calcular fàcilment una fórmula per l'invers d' $a + bi$:

$$(a + bi)^{-1} = \frac{1}{a + bi} = \frac{a - bi}{(a + bi)(a - bi)} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}i.$$

Pensat com un element de \mathbb{R}^2 , $N(z) = |z|^2$, on $|z| = \sqrt{a^2 + b^2}$ és el mòdul de z pensat com un element de \mathbb{R}^2 . Les coordenades polars del punt $(a, b) \in \mathbb{R}^2$ es calculen fent servir que:

$$a = r \cos(\theta), \quad b = r \sin(\theta),$$

i per tant:

$$r = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}, \quad \theta = \arg(z) = \arctan(b/a)(+\pi),$$

on haurem de sumar π a $\arctan(b/a) \in [-\pi/2, \pi/2]$ si el punt (a, b) pertany al II o III quadrant.

Podem escriure (aquí ho podem pensar com una notació, encara que té una justificació algebraica) que

$$a + bi = re^{i\theta},$$

i així podem recordar les fórmules

$$|zw| = |z||w|, \quad \arg(zw) = \arg(z) + \arg(w) \pmod{2\pi}.$$

El següent resultat és el motiu pel què ens interessa treballar a \mathbb{C} :

Teorema A.1 (Teorema fonamental de l'àlgebra). *Tot polinomi $f(x) \in \mathbb{C}[x]$ de grau $n \geq 0$ es pot escriure com*

$$f(x) = c(x - \lambda_1) \cdots (x - \lambda_n),$$

on c i $\lambda_1, \dots, \lambda_n$ són nombres complexos (possiblement repetits).

Referències

- [1] Otto Bretscher, *Linear Algebra with Applications*, 1997, Prentice-Hall International Inc.
- [2] Enric Nart, Xavier Xarles, *Apunts d'àlgebra lineal*, 2016, Col·lecció Materials, 237, UAB.