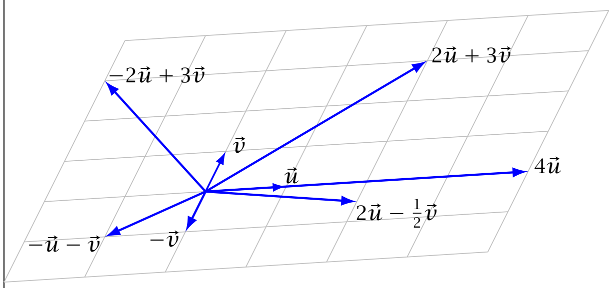
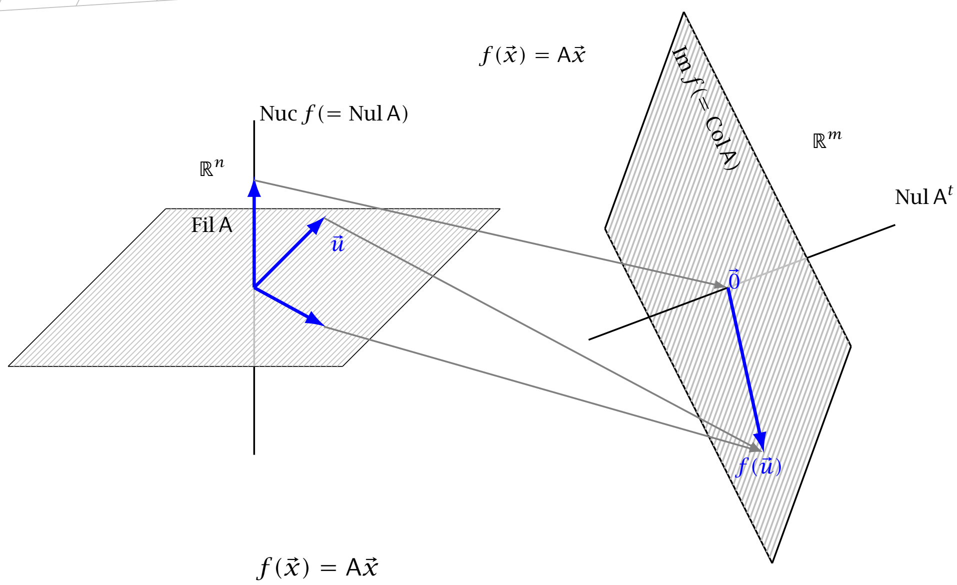
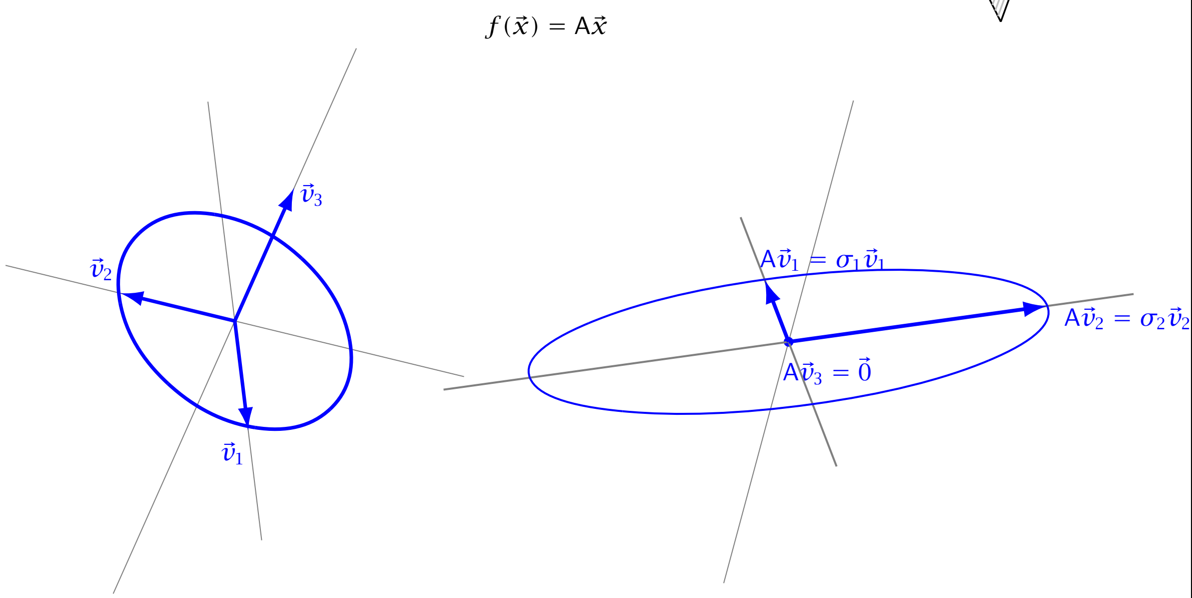
Universitat Politècnica de València

Grau en enginyeria informàtica

ÀLGEBRA







Robert Fuster

Departament de Matemàtica Aplicada

Curs 2017–2018

Darrera actualització: 23 d'abril de 2018

Continguts

INTRODUCCIÓ

Continguts

1. PROBLEMES *lineals*

En aquesta primera sessió del curs, mirarem de definir què és l'Àlgebra Lineal a través de diversos problemes que es resolen amb tècniques lineals i descriurem la relació que tenen aquests problemes amb la nostra assignatura.

L'adjectiu *lineal* apareixerà diverses vegades en aquest curs: hi ha sistemes d'equacions *lineals*, combinacions *lineals*, transformacions (o aplicacions) *lineals*, vectors *lineal*ment dependents o independents, equacions en diferències *lineals*, equacions diferencials *lineals*... En anglès, dels espais vectorials en diuen *vector spaces*, però també *linear spaces*.

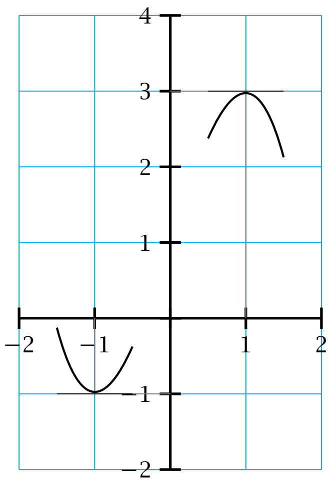
L'objecte d'estudi de l'Àlgebra Lineal són les operacions i les transformacions lineals. Com que les operacions lineals es fan amb matrius (si més no, quan la dimensió és finita), podem concloure que l'Àlgebra Lineal s'ocupa del càlcul amb matrius i de les propietats de les matrius.
Però, com com que les transformacions lineals tenen lloc en els espais vectorials, també resulta que l'àlgebra lineal s'ocupa dels espais vectorials, és a dir, dels espais de vectors, i de les transformacions entre espais vectorials, és a dir, de les aplicacions lineals.

Podem posar l'èmfasi en les matrius o en els espais vectorials i les aplicacions lineals. En qualsevol cas, però, l'instrument bàsic de l'Àlgebra Lineal són les combinacions lineals. Una operació entre vectors és lineal quan l'únic que fem amb els vectors és sumar-los i/o multiplicar-los per nombres. Combinant aquestes dues operacions fem una *combinació lineal*: si $\vec{u}_1$ i $\vec{u}_2$ són dos vectors i α_1, α_2 dos nombres, aleshores $\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2$ és un nou vector, obtingut com a combinació lineal de $\vec{u}_1$ i $\vec{u}_2$.

El que farem ara serà mostrar alguns problemes de matèries diverses on s'hi apliquen les tècniques de l'Àlgebra Lineal.

1.1. SISTEMES D'EQUACIONS LINEALS. APROXIMACIÓ POLINÒMICA D'UNA FUNCIÓ

La gràfica d'una funció polinòmica de tercer grau passa pels punts $(-1,-1)$ i $(1,3)$ i té la tangent horitzontal en aquests dos punts. Quina és aquesta funció?



Encara que el problema es planteja en termes d'anàlisi matemàtica, la solució consisteix a resoldre un sistema lineal de quatre equacions amb quatre incògnites:

- La funció és $f(x) = ax^3 + bx^2 + cx + d$ i la seua derivada, $f'(x) = 3ax^2 + 2bx + c$.
- Que la corba passe pel punt $(-1,-1)$ amb tangent horitzontal en aquest punt vol dir que $f(-1) = -1$ i $f'(-1) = 0$, és a dir,

$$\begin{aligned} -a + b - c + d &= -1 \\ 3a - 2b + c \quad &= 0 \end{aligned}$$

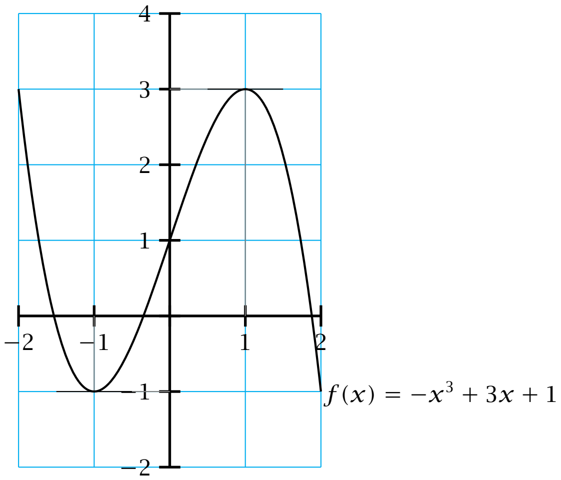
- I que passe també per $(1,3)$ amb tangent horitzontal significa que

$$\begin{aligned} a + b + c + d &= 3 \\ 3a + 2b + c \quad &= 0 \end{aligned}$$

Per tant, els coeficients del polinomi s'obtenen resolent el sistema d'equacions lineals

$$\begin{aligned} -a + b - c + d &= -1 \\ 3a - 2b + c \quad &= 0 \\ a + b + c + d &= 3 \\ 3a + 2b + c \quad &= 0 \end{aligned}$$

La solució d'aquest sistema és $a = -1$, $b = 0$, $c = 3$, $d = 1$. En conseqüència, la funció que busquem és $f(x) = -x^3 + 3x + 1$.



Com que aquest sistema d'equacions lineals només té una solució, podem assegurar que el polinomi que hem trobat és únic: només hi ha una funció polinòmica de grau menor o igual a tres que passe pels dos punts donats i hi tinga una tangent horitzontal.

☞ En el tema 1 estudiarem la resolució dels sistemes d'equacions lineals. Tot i que aquest és un tema aparentment trivial i ben conegut, les tècniques que s'hi fan servir i les propietats que s'hi troben són la base de gran part de l'Àlgebra Lineal.

1.2. Cadenes de Màrkov i matrius estocàstiques

Suposem que en un país l'evolució de les migracions internes al llarg d'un cert període de temps ha estat aquesta:

- Cada any, el 10 % de la població de la capital es trasllada a viure a una altra ciutat i el 5 % d'aquesta població se'n va al medi rural.
- Dels habitants de les altres ciutats, el 15 % emigra a la capital i un altre 10 % al medi rural.
- El 12 % dels habitants del medi rural emigren tots els anys a la capital i el 15 % a una altra ciutat.

En aquest context, ens plantegem les qüestions següents:

- Si en un moment determinat la població de la capital és el 20 % del total, la de les altres ciutats el 45 % i la resta dels habitants del país, és a dir, el 35 %, viuen al medi rural, quina serà la distribució al cap d'un any?
- I després de 5, 10, 20 anys?
- Hi ha alguna tendència a l'estabilitat, de manera que a llarg termini la distribució de la població no patisca canvis significatius?
- Si la distribució de la població inicial fos una altra, per exemple, 30 %, 20 %, 50 %, quins serien els resultats a llarg termini?

Introduïm les dades migratòries en un quadre de doble entrada:

	capital	altres ciutats	medi rural
capital	0,85	0,15	0,12
altres ciutats	0,10	0,75	0,15
medi rural	0,05	0,10	0,73

(1.1)

Cada columna d'aquest quadre és un vector de probabilitats: la primera columna conté les probabilitats que una persona que avui viu a la capital estiga d'ací un any (a) a la capital (0,85 = 85 %), (b) a una altra ciutat (0,10 = 10 %) o (c) al medi rural (0,05 = 5 %). Aquests tres nombres han de sumar necessàriament 1, atés que estem suposant que tota la població del país habita un d'aquests tres medis, i són no negatius, perquè un valor negatiu no té cap sentit en el problema que ens ocupa.

De manera semblant s'interpreten les altres dues columnes. Anomenem $\vec{p}_1$, $\vec{p}_2$ i $\vec{p}_3$ aquests tres vectors i representem-los en forma de columna:

$$\vec{p}_1 = \begin{bmatrix} 0{,}85 \\ 0{,}10 \\ 0{,}05 \end{bmatrix} \qquad \vec{p}_2 = \begin{bmatrix} 0{,}15 \\ 0{,}75 \\ 0{,}10 \end{bmatrix} \qquad \vec{p}_3 = \begin{bmatrix} 0{,}12 \\ 0{,}15 \\ 0{,}73 \end{bmatrix}$$

D'altra banda, el vector

$$\vec{u}_0 = \begin{bmatrix} 0{,}20 \\ 0{,}45 \\ 0{,}35 \end{bmatrix}$$

representa la distribució inicial de la població. Per saber quina serà la distribució de la població d'ara en un any, hem de

(a) multiplicar el vector $\vec{p}_1$ per $0,20$, per a esbrinar on seran l'any que ve les persones que avui viuen a la capital,
(b) multiplicar el vector $\vec{p}_2$ per $0,45$, per fer el mateix amb els habitants actuals de les altres ciutats,
(c) multiplicar el vector $\vec{p}_3$ per $0,35$ i
(d) sumar aquests tres vectors.

Llavors obtindrem:

$$\vec{u}_1 = 0,20\vec{p}_1 + 0,45\vec{p}_2 + 0,35\vec{p}_3 = 0,20 \begin{bmatrix} 0{,}85 \\ 0{,}10 \\ 0{,}05 \end{bmatrix} + 0,45 \begin{bmatrix} 0{,}15 \\ 0{,}75 \\ 0{,}10 \end{bmatrix} + 0,35 \begin{bmatrix} 0{,}12 \\ 0{,}15 \\ 0{,}73 \end{bmatrix} = \begin{bmatrix} 0{,}2795 \\ 0{,}41 \\ 0{,}3105 \end{bmatrix}$$

així que, d'ací a un any la població de la capital serà del 27,95 %, la de les altres ciutats del 41 % i la del medi rural, del 31,05 %.

Hem resolt la primera qüestió fent una combinació lineal amb els tres vectors $\vec{p}_1$, $\vec{p}_2$ i $\vec{p}_3$. Aquesta combinació lineal també la podem interpretar com un producte matriu-vector: si amb els tres vectors de probabilitats construïm la matriu

$$\mathsf{M} = \begin{bmatrix} 0{,}85 & 0{,}15 & 0{,}12 \\ 0{,}10 & 0{,}75 & 0{,}15 \\ 0{,}05 & 0{,}10 & 0{,}73 \end{bmatrix}$$

(la matriu M no és altra cosa que el quadre (1.1)) llavors

$$\vec{u}_1 = 0,20\vec{p}_1 + 0,45\vec{p}_2 + 0,35\vec{p}_3 = \begin{bmatrix} 0{,}85 & 0{,}15 & 0{,}12 \\ 0{,}10 & 0{,}75 & 0{,}15 \\ 0{,}05 & 0{,}10 & 0{,}73 \end{bmatrix} \begin{bmatrix} 0{,}20 \\ 0{,}45 \\ 0{,}35 \end{bmatrix} = \mathsf{M}\vec{u}_0$$

Per a decidir com es distribueix la població al cap de 2, 3, 4, 5..., 10..., 20 anys hem de anar calculant els productes

$$\vec{u}_1 = \mathsf{M}\vec{u}_0,\ \vec{u}_2 = \mathsf{M}\vec{u}_1,\ \vec{u}_3 = \mathsf{M}\vec{u}_2,\ \dots$$

El quadre següent mostra els resultats que obtindríem (arrodonits a quatre xifres decimals):

	anys									
	0	1	2	3	4	5	···	10	···	20
capital	0,20	0,2795	0,3363	0,3770	0,4061	0,4269		0,4695		0,4791
altres ciutats	0,45	0,41	0,3820	0,3623	0,3486	0,3388		0,3194		0,3152
medi rural	0,35	0,3105	0,2816	0,2606	0,2453	0,2343		0,2111		0,2057

És clar que la població de la capital creix i les de les altres ciutats i del camp decreixen. Però aquestes variacions van amortint-se, de manera que el primer any el creixement de població de la capital és aproximadament del 8 %, el segon any, del 5 %, en el pas del quart al cinqué any només del 2 % i en els deu anys que van del desé al vinté, només de l'1 %.[1] De fet, el vector de població a l'any 100 serà aquest: a la capital 0,4795, a les altres ciutats 0,3151 i al medi rural 0,2055, amb una diferència màxima entre l'any 20 i l'any 100 de només el 4 per 10000.

En conseqüència, podem concloure que *a la llarga* la població tendirà a estabilitzar-se al voltant dels percentatges següents: 48 % 31,5 % i 20,5 %.

[1] No és un 1 % annual, sinó un 1 % en deu anys!

Ara bé, quin significat té el fet que *la població s'estabilitze*? Notem que les persones continuen migrant (sempre en els mateixos percentatges: tots els anys, el 15 % dels habitants de la capital emigren a un altre medi), però els moviments es compensen, de manera que, a la llarga, el nombre de persones que deixen la capital és aproximadament igual al de les que hi van a viure.

En termes de l'anàlisi matemàtica, la successió de vectors $\vec{u}_0$, $\vec{u}_1$, $\vec{u}_2$... convergeix a un vector $\vec{u}$ i aquest vector, que representa la distribució estable de la població al cap d'uns quants anys, té la propietat següent:

$$\mathsf{M}\vec{u} = \vec{u}$$

Quan un vector té aquesta propietat diem que és *un vector estacionari de la matriu* M. De fet, es tracta del primer vector propi que veiem en aquest curs. Podem calcular *exactament* aquest vector?

Suposem finalment que el vector de la població inicial és $\begin{bmatrix} 0{,}30 \\ 0{,}20 \\ 0{,}50 \end{bmatrix}$. En els anys subsegüents tindrem aquestes proporcions:

	anys									
	0	1	2	3	4	5	···	10	···	20
capital	0,30	0,345	0,3795	0,4056	0,4251	0,4397		0,4742		0,4791
altres ciutats	0,20	0,255	0,2857	0,3025	0,3112	0,3155		0,3171		0,3152
medi rural	0,50	0,4	0,3347	0,2919	0,2636	0,2448		0,2115		0,2057

així que al cap de vint anys la distribució serà la mateixa que amb les dades inicials.

El que hem descrit és un exemple típic de *cadena de Màrkov*, és a dir, una successió de vectors *d'estat* $\vec{u}_0$, $\vec{u}_1$, $\vec{u}_2$..., que es defineixen per recurrència com

$$\vec{u}_{n+1} = \mathsf{M}\vec{u}_n = \mathsf{M}^n\vec{u}_0$$

on

- Els components dels vectors $\vec{u}_n$ estan tots entre 0 i 1 (perquè representen probabilitats o, si es vol, percentatges).
- La suma de tots els components d'aquests vectors és 1 (perquè representa el conjunt unió de tots els estats).

  Un vector que compleix aquestes dues propietats és un vector *estocàstic*.
- Les columnes de la *matriu de transició* M són també vectors estocàstics.

Una cadena de Màrkov és *regular* si la matriu de transició M o una potència d'aquesta matriu té totes les entrades no nul·les. Doncs bé, es pot provar que si la cadena és regular, aleshores la successió $\vec{u}_n$ convergeix a un vector *estacionari* $\vec{u}$, de manera que $\mathsf{M}\vec{u} = \vec{u}$. Aquest vector es pot calcular per mètodes estrictament lineals: l'única cosa que hem de fer és resoldre el sistema d'equacions lineals

$$\mathsf{M}\vec{x} = \vec{x}$$

i, entre totes les solucions, elegir-ne una que siga un vector estocàstic.

☞ En l'estudi de les matrius estocàstiques i les cadenes de Markov es fan servir diverses eines de l'Àlgebra Lineal: multiplicació de matrius, sistemes lineals, valors i vectors propis... Ens n'ocuparem en alguns exercicis pràctics del curs.

1.3. Valors i vectors propis. Equacions en diferències

Una equació en diferències (o recurrència) lineal i homogènia és una expressió com aquesta:

$$\begin{aligned} a_1 &= 0 \\ a_2 &= 1 \\ a_n - 3a_{n-1} + 2a_{n-2} &= 0 \end{aligned}$$

Aquesta equació és d'ordre 2, perquè hi apareixen els termes a_n, a_{n-1} i a_{n-2}, però podem transformar-la en una equació (matricial) d'ordre 1: Definim el vector $\vec{a}_n = (a_n, a_{n+1})$ i llavors les equacions anteriors són equivalents a

$$\vec{a}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\vec{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} a_n \\ -2a_{n-1} + 3a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \vec{a}_{n-1}$$

D'ací es dedueix que

$$\vec{a}_n = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}^{n-1} \vec{a}_1 = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}^{n-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

I el problema és ara el càlcul de les potències de la matriu

$$\mathsf{A} = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}$$

Quan estudiarem el problema de la diagonalització podrem provar que

$$\underbrace{\begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}}_{\mathsf{P}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}}_{\mathsf{D}} \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}}_{\mathsf{P}^{-1}}$$

i, a partir d'ací, es prova fàcilment que

$$\vec{a}_n = \mathsf{A}^{n-1}\vec{a}_1 = \begin{bmatrix} -1 + 2^{n-1} \\ -1 + 2^n \end{bmatrix}$$

de manera que la solució de l'equació recurrent original és

$$\boxed{a_1 = 0, \quad a_2 = 1, \quad a_n = -1 + 2^{n-1}, \ (n = 2, 3, \dots)}$$

Les columnes de la matriu P (és a dir, $(2,1)$ i $(1,1)$) són vectors propis de la matriu A, i les entrades diagonals de D (2 i 1) en són els valors propis.

1.3.1. Els nombres de Fibonacci

La famosa successió de Fibonacci, $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \dots$, es defineix per recurrència d'aquesta manera:

$$f_1 = 1$$
$$f_2 = 1$$
$$f_n = f_{n-2} + f_{n-1}$$

Aplicant la tècnica que acabem de presentar es pot provar que

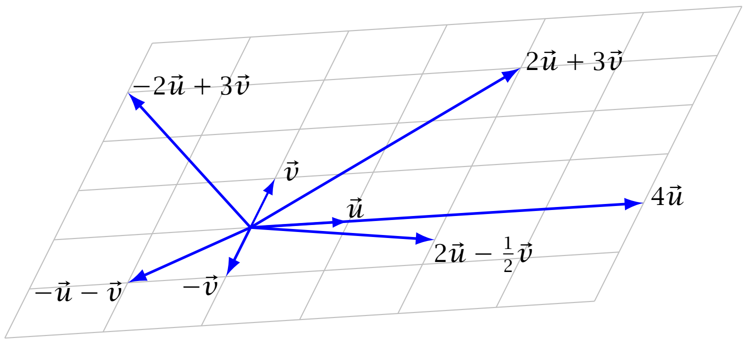
$$f_n = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n$$

Per sorprenent que puga semblar, aquests nombres són enters! De fet es pot simplificar el càlcul si es té en compte que els nombres $\frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^n$ són tan petits que f_n és l'enter més pròxim a $\frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^n$. Per exemple, fent servir una calculadora obtindrem que

$$\frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{12} = 144{,}001388875\dots$$

de manera que al cap de dotze mesos tindrem $f_{12} = 144$ parelles de conills.

☞ En els temes 1 i 2 estudiarem les operacions amb matrius (que hem fet servir en aquest exemple per a calcular potències, productes i inverses) i, com ja hem dit, en el tema 6 tractarem els problemes relacionats amb els valors propis. Per tal de calcular els valors propis farem servir els determinants, que estudiem al tema 5.



Primera part
Matrius, vectors i sistemes d'equacions lineals

Capítol 1

Sistemes d'equacions lineals

Continguts

2. MATRIUS I VECTORS

"You can't add apples and oranges."
In a strange way, this is the reason for vectors!
Gilbert Strang

En aquesta lliçó repassarem les operacions amb vectors i amb matrius, fixant-nos especialment en el producte escalar de dos vectors i les aplicacions geomètriques d'aquest producte. També ens interessa remarcar la interpretació de les matrius com a conjunts de vectors.
L'operació amb matrius més important, la multiplicació, l'estudiarem en la propera sessió.

2.1. MATRIUS

Una *matriu* és un conjunt de nombres reals o complexos, ordenats en una taula de doble entrada. Per exemple,

- $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & 3 \end{bmatrix}$ és una matriu 2×3 (llegiu *dos per tres*).
- $\begin{bmatrix} 0 & -2 \\ 1+\mathrm{i} & 0 \end{bmatrix}$ és una matriu *quadrada* 2×2 (o una matriu quadrada *d'ordre* 2),
- $\begin{bmatrix} 1 & 2 & 3\mathrm{i} \end{bmatrix}$ és una matriu *fila* 1×3
- $\begin{bmatrix} 1 \\ 2 \\ 3\mathrm{i} \end{bmatrix}$ és una matriu *columna* (o *vector*) 3×1.
- $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ és la matriu *nul·la* 3×2.
- $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ és la matriu *identitat* 3×3.
- $\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & & & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}$ és una matriu $m \times n$ (m files per n columnes).

  ☞ El primer índex es refereix a la fila i, el segon, a la columna. Per exemple, a_{23} és el nombre que hi ha a la segona fila i a la tercera columna.

El conjunt de totes les matrius $m \times n$ el representarem com $\mathcal{M}_{m\times n}$; en el cas que necessitem distingir si treballem amb nombres reals o complexos escriurem $\mathcal{M}_{m\times n}(\mathbb{R})$ o bé $\mathcal{M}_{m\times n}(\mathbb{C})$. Els nombres que hi ha a la matriu són els *elements* o les *entrades* de la matriu.

☞ Fixeu-vos que quan parlem d'una matriu $m \times n$, m sempre és el nombre de files i n el de columnes.

2.1.1. OPERACIONS AMB MATRIUS: SUMA, PRODUCTE ESCALAR-MATRIU I COMBINACIONS LINEALS

Amb les matrius es poden fer tres operacions bàsiques: la suma, la multiplicació per un escalar i la multiplicació de dues matrius.

Per poder sumar dues matrius cal que tinguen les mateixes dimensions (el mateix nombre de files i el mateix nombre de columnes). Llavors, l'únic que hem de fer és sumar-les *element a element*. Per exemple,

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

Per a restar dues matrius també ho farem element a element:

$$\begin{bmatrix}1 & -1 & 0\\ 0 & 2 & -1\end{bmatrix} - \begin{bmatrix}2 & 3 & -2\\ 0 & -1 & 0\end{bmatrix} = \begin{bmatrix}1-2 & -1-3 & 0+2\\ 0-0 & 2+1 & -1-0\end{bmatrix} = \begin{bmatrix}-1 & -4 & 2\\ 0 & 3 & -1\end{bmatrix}$$

(No es pot sumar ni restar dues matrius que no tinguen les mateixes dimensions).

Representarem com $-\mathsf{A}$ la matriu que obtenim si canviem el signe a totes les entrades de A; per exemple, $-\begin{bmatrix}1 & 2 & -1\\ 0 & -3 & 0\end{bmatrix} = \begin{bmatrix}-1 & -2 & 1\\ 0 & 3 & 0\end{bmatrix}$. Aleshores, la diferència $\mathsf{A}-\mathsf{B}$ és el mateix que la suma de les matrius A i $-\mathsf{B}$, és a dir, $\mathsf{A}-\mathsf{B} = \mathsf{A}+(-\mathsf{B})$.

Propietats de la suma i la diferència

1. $\mathsf{A}+\mathsf{B} = \mathsf{B}+\mathsf{A}$ (commutativa)
2. $(\mathsf{A}+\mathsf{B})+\mathsf{C} = \mathsf{A}+(\mathsf{B}+\mathsf{C})$ (associativa)
3. $\mathsf{A}+\mathsf{O} = \mathsf{A}$ (O és el neutre de la suma)
4. $\mathsf{A}-\mathsf{B} = \mathsf{A}+(-\mathsf{B})$
5. $\mathsf{A}-\mathsf{A} = \mathsf{A}+(-\mathsf{A}) = \mathsf{O}$ ($-\mathsf{A}$ és el simètric de A)

En Àlgebra Lineal se sol anomenar *escalars* als nombres, i sovint es representen amb lletres gregues (α, β,...), per tal de distingir-los dels vectors i de les matrius. Per a multiplicar un escalar (és a dir, un nombre) per una matriu, hi multipliquem cada entrada de la matriu. Per exemple,

$$3\begin{bmatrix}1 & -1 & 0\\ 0 & 2 & -1\end{bmatrix} = \begin{bmatrix}3\cdot 1 & 3(-1) & 3\cdot 0\\ 3\cdot 0 & 3\cdot 2 & 3(-1)\end{bmatrix} = \begin{bmatrix}3 & -3 & 0\\ 0 & 6 & -3\end{bmatrix}$$

Propietats del producte escalar-matriu

1. $\alpha_1(\alpha_2\mathsf{A}) = (\alpha_1\alpha_2)\mathsf{A}$ (associativa)
2. $(\alpha_1+\alpha_2)\mathsf{A} = \alpha_1\mathsf{A}+\alpha_2\mathsf{A}$ (distributiva respecte a la suma d'escalars)
3. $\alpha(\mathsf{A}+\mathsf{B}) = \alpha\mathsf{A}+\alpha\mathsf{B}$ (distributiva respecte a la suma de vectors)
4. $1\mathsf{A} = \mathsf{A}$ (el nombre 1 és l'element neutre)
5. $(-\alpha)\mathsf{A} = -(\alpha\mathsf{A}) = \alpha(-\mathsf{A})$ (elements oposats)

Una combinació qualsevol d'aquestes operacions és una *combinació lineal.* Per exemple,

$$3\begin{bmatrix}1 & -1 & 0\\ 0 & 2 & -1\end{bmatrix} + 2\begin{bmatrix}2 & 3 & -2\\ 0 & -1 & 0\end{bmatrix} = \begin{bmatrix}3\cdot 1+2\cdot 2 & 3(-1)+2\cdot 3 & 3\cdot 0+2(-2)\\ 3\cdot 0+2\cdot 0 & 3\cdot 2+2(-1) & 3(-1)+2\cdot 0\end{bmatrix} = \begin{bmatrix}7 & 3 & -4\\ 0 & 4 & -3\end{bmatrix}$$

és una combinació lineal de les matrius $\begin{bmatrix}1 & -1 & 0\\ 0 & 2 & -1\end{bmatrix}$ i $\begin{bmatrix}2 & 3 & -2\\ 0 & -1 & 0\end{bmatrix}$. Els nombres 3 i 2, que multipliquen les matrius, són els *pesos* de la combinació lineal.

En general, la matriu A és la combinació lineal de les matrius $\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_p$, amb pesos $\alpha_1, \alpha_2, \ldots, \alpha_p$ si

$$\mathsf{A} = \alpha_1\mathsf{A}_1 + \alpha_1\mathsf{A}_2 + \cdots + \alpha_p\mathsf{A}_p$$

☞ Les combinacions lineals són les eines més importants del càlcul amb vectors i matrius.

2.2. Vectors

Un vector és una matriu columna. Per exemple,

$$\vec{u} = \begin{bmatrix}1\\ 1\\ 0\\ -2\end{bmatrix} \quad \vec{v} = \begin{bmatrix}1+2\mathrm{i}\\ 1-\mathrm{i}\end{bmatrix} \quad \vec{w} = \begin{bmatrix}\sqrt{2}\\ 3\\ -\pi\end{bmatrix} \quad \vec{x} = \begin{bmatrix}-\frac{1}{2}\end{bmatrix} \quad \vec{0} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix} \quad \vec{a} = \begin{bmatrix}a_1\\ a_2\\ \cdots\\ a_n\end{bmatrix}$$

☞ En aquest curs representarem els vectors amb una lletra amb una fletxeta a sobre (d'acord amb la visió geomètrica que identifica els vectors amb fletxes), com ara, $\vec{u}$, $\vec{v}$ o $\vec{x}$, o bé amb lletres majúscules, com ara U, V o X. Les matrius en general les representarem sempre amb lletres majúscules.

Cadascun dels nombres que hi apareixen és un *component* del vector. Segons el nombre de components que tinguen, els vectors s'anomenen bidimensionals, tridimensionals, etc. (o, alternativament, de dimensió dos, de dimensió tres...). El conjunt de tots els vectors n-dimensionals es representa com $\mathbb{R}^n$ o $\mathbb{C}^n$, segons que treballem amb nombres reals o complexos. Com que moltes vegades serà indiferent que els nombres siguen reals o complexos, usarem la lletra $\mathbb{K}$ per representar indistintament el conjunt dels nombres reals o el dels complexos, de manera que quan escrivim $\mathbb{K}^n$ ens estem referint tant als vectors reals com als complexos.

Anomenem *zero* al vector de dimensió n que té tots els components iguals a 0 (aquest vector el representarem com $\vec{0}$ o O).

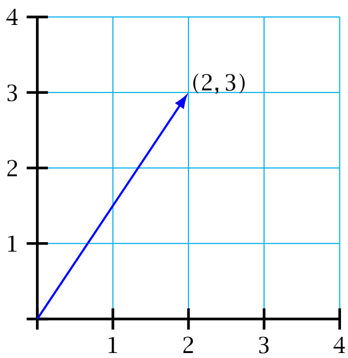
Com que la notació en columnes provoca un espaiat excessiu del text en el sentit vertical (especialment quan s'inclou un vector a dins d'un paràgraf), sovint farem servir la notació tradicional $(1, 1, 0, -2)$ per a representar el vector $\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}$:

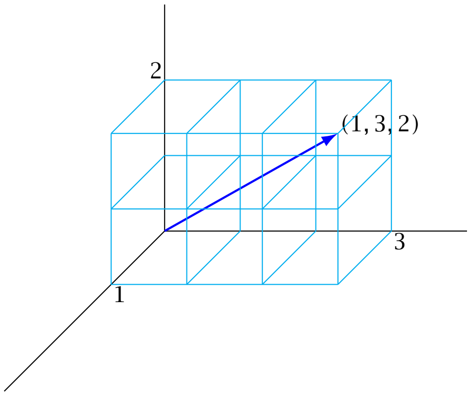
> Un vector és una expressió del tipus $\begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}$. Però, per comoditat, el representarem com $(1, 1, 0, -2)$.
>
> És a dir, que $(1, 1, 0, -2) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -2 \end{bmatrix}$.

☞ Insistim encara més, sobre aquest punt: $(1, 1, 0, -2)$, escrit entre parèntesis (rodons) i amb els components separats per comes, és un vector; en canvi, $\begin{bmatrix} 1 & 1 & 0 & -2 \end{bmatrix}$, entre claudàtors (parèntesis rectangulars) i amb les entrades separades per espais en blanc, és una matriu fila: $(1, 1, 0, -2) \neq \begin{bmatrix} 1 & 1 & 0 & -2 \end{bmatrix}$.

2.2.1. Interpretació geomètrica dels vectors i de les operacions amb vectors

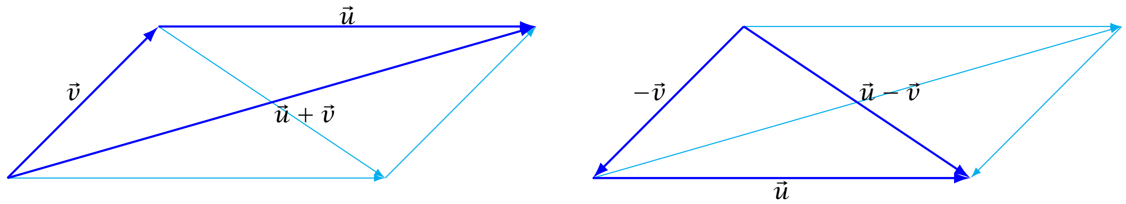
Els vectors de $\mathbb{R}^2$ i $\mathbb{R}^3$ es visualitzen molt bé com a punts o com a fletxes en un pla (en el cas de $\mathbb{R}^2$) o en un espai tridimensional (quan es tracta de vectors de $\mathbb{R}^3$).





Per als vectors amb més de tres components, la imatge dels vectors com a fletxes segueix sent una representació intuïtiva molt adient. D'altra banda, les operacions i les propietats que coneixem bé en els casos bidimensional i tridimensional s'hi poden traslladar sense gaire dificultat. D'això ens n'encarregarem tot seguit.

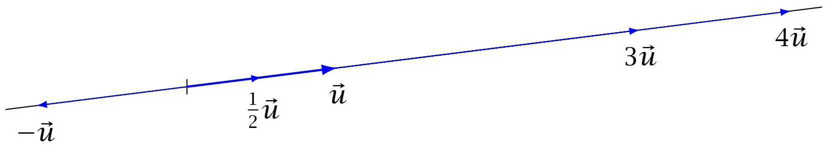
D'acord amb aquesta visió, les operacions entre vectors també es poden interpretar geomètricament: dos vectors, $\vec{u}$, i $\vec{v}$, defineixen un paral·lelogram. Llavors, la suma dels dos vectors és la diagonal d'aquest paral·lelogram, orientada des de l'origen fins a l'extrem oposat; i la diferència, $\vec{u} - \vec{v}$, és l'altra diagonal, orientada des de l'extrem de $\vec{v}$ cap al de $\vec{u}$.



Suma i diferència de dos vectors

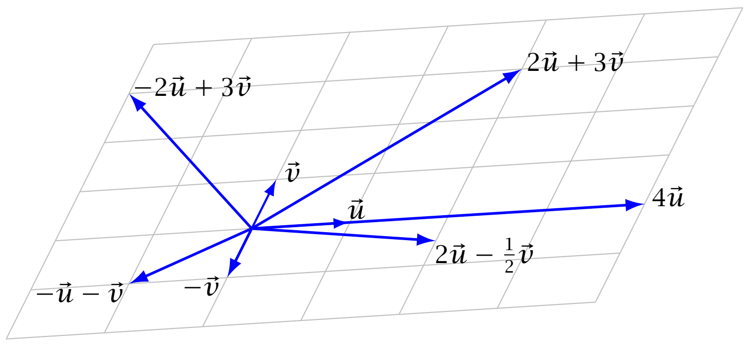
Multiplicar un vector per un escalar produeix una dilatació (o una contracció) del vector: multiplicar-lo per 3, triplica la longitud. Per $1/2$, el divideix a la meitat. I, quan l'escalar és negatiu, llavors l'orientació del vector s'inverteix.

Notem que tots els productes es troben sobre la mateixa recta.



Producte de diversos escalars per un vector

Aleshores, per representar gràficament la combinació lineal $\alpha_1\vec{u} + \alpha_2\vec{v}$, haurem de dilatar els vectors, segons els pesos, α_1 i α_2 i fer-ne la suma.



Diverses combinacions lineals dels dos vectors $\vec{u}$ i $\vec{v}$

2.3. El producte escalar

La tercera operació important amb vectors és el producte escalar $\vec{u} \cdot \vec{v}$. Per a calcular el producte escalar de dos vectors $\vec{u}$ i $\vec{v}$ també cal que tinguen el mateix nombre de components; però, a més a més, convé que distingim si els vectors són reals o complexos.

2.3.1. El producte escalar (cas real)

En aquest apartat estudiem el producte escalar de dos vectors *reals*.

El producte escalar de $\vec{u}$ i $\vec{v}$ es calcula multiplicant-los component a component i sumant tots aquests productes.

$$\vec{u} \cdot \vec{v} = (u_1, u_2, \dots, u_n) \cdot (v_1, v_2, \dots, v_n) = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Per exemple,

$$(2, -1, 3, 0) \cdot (-1, 1, 4, 1) = 2 \cdot (-1) + (-1) \cdot 1 + 3 \cdot 4 + 0 \cdot 1 = -2 - 1 + 12 + 0 = 9$$

☞ Fixeu-vos bé que el producte *escalar* de dos vectors és un escalar (no un vector).

Propietats del producte escalar real

1. $\vec{u}_1 \cdot \vec{u}_2 = \vec{u}_2 \cdot \vec{u}_1$ (commutativa)
2. $(\vec{u}_1 + \vec{u}_2) \cdot \vec{v} = \vec{u}_1 \cdot \vec{v} + \vec{u}_2 \cdot \vec{v}$ (distributiva)
3. $\alpha(\vec{u}_1 \cdot \vec{u}_2) = (\alpha\vec{u}_1) \cdot \vec{u}_2 = \vec{u}_1 \cdot (\alpha\vec{u}_2)$ (associativitat amb el producte escalar-vector)
4. $\vec{0} \cdot \vec{u} = 0$ ($\vec{0}$ és ortogonal a tots els vectors)
5. Si $\vec{u} \neq \vec{0}$, llavors $\vec{u} \cdot \vec{u} > 0$ (definit positiu)

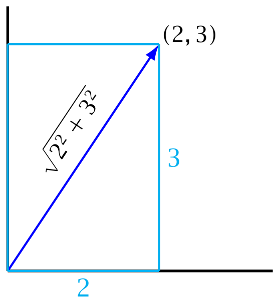
☞ El producte escalar és un instrument molt potent per a l'estudi de les propietats geomètriques dels vectors, així que al llarg d'aquest curs l'estudiarem amb certa profunditat. En principi, el podem fer servir per a mesurar la longitud dels vectors i l'angle entre dos vectors.

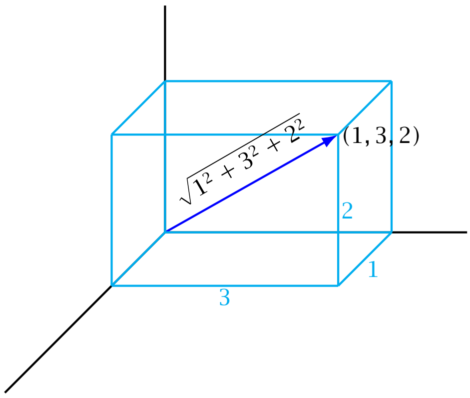
2.3.2. Aplicacions geomètriques del producte escalar

La *norma* (o *longitud*) del vector $\vec{u} \in \mathbb{R}^n$ és el nombre

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} \tag{2.1}$$

Els gràfics següents justifiquen aquesta definició en el cas dels vectors de dos o tres components.





Els vectors de norma igual a la unitat s'anomenen *vectors unitaris*.

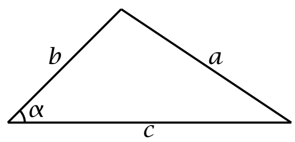
La *distància* entre dos vectors $\vec{u}$ i $\vec{v}$ és la norma de la diferència:

$$\mathrm{d}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

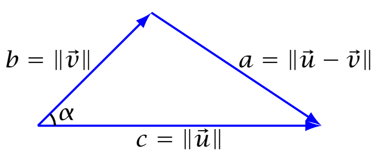
Finalment, per a calcular l'angle entre dos vectors $\vec{u}$ i $\vec{v}$, aplicarem el teorema del cosinus:

En qualsevol triangle de costats a, b i c, si α és l'angle oposat al costat a, llavors

$$a^2 = b^2 + c^2 - 2bc\cos\alpha$$



Suposem que α és l'angle entre els dos vectors $\vec{u}$ i $\vec{v}$. Llavors, les longituds d'aquests vectors són, respectivament, $\|\vec{u}\|$ i $\|\vec{v}\|$ i el vector $\vec{u} - \vec{v}$ completa el triangle



Aleshores, segons el teorema del cosinus,

$$\begin{aligned}
a^2 &= b^2 + c^2 - 2bc\cos\alpha \\
\|\vec{u} - \vec{v}\|^2 &= \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\,\|\vec{u}\|\,\|\vec{v}\|\cos\alpha \\
(\vec{u} - \vec{v})\cdot(\vec{u} - \vec{v}) &= \vec{u}\cdot\vec{u} + \vec{v}\cdot\vec{v} - 2\,\|\vec{u}\|\,\|\vec{v}\|\cos\alpha \\
\vec{u}\cdot\vec{u} + \vec{v}\cdot\vec{v} - 2\vec{u}\cdot\vec{v} &= \vec{u}\cdot\vec{u} + \vec{v}\cdot\vec{v} - 2\,\|\vec{u}\|\,\|\vec{v}\|\cos\alpha \\
-2\vec{u}\cdot\vec{v} &= \qquad\qquad\qquad - 2\,\|\vec{u}\|\,\|\vec{v}\|\cos\alpha
\end{aligned}$$

Per tant,

$$\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos\alpha \tag{2.2}$$

De manera que, si $\vec{u} \neq \vec{0}$ i $\vec{v} \neq \vec{0}$,

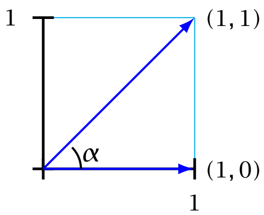
$$\boxed{\cos\alpha = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}} \tag{2.3}$$

Aquesta expressió ens permet calcular l'angle entre els dos vectors. Noteu que si un dels vectors és nul, aleshores no hi ha cap angle (un angle és l'obertura determinada per dos segments de recta; com que el vector zero es limita a un sol punt no s'hi pot definir cap angle).

Per exemple, si α és l'angle determinat pels vectors $(1,1)$ i $(1,0)$,

$$\cos\alpha = \frac{(1,1)\cdot(1,0)}{\|(1,1)\|\,\|1,0\|} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$$

Així que $\alpha = \pi/4$ (o 45^{o}).



D'altra banda, de la fórmula (2.2) es dedueix una propietat molt important, coneguda com *desigualtat de Cauchy-Schwarz*:

$$\boxed{|\vec{u}\cdot\vec{v}| \leq \|\vec{u}\|\,\|\vec{v}\|}, \qquad \forall\,\vec{u},\vec{v}$$

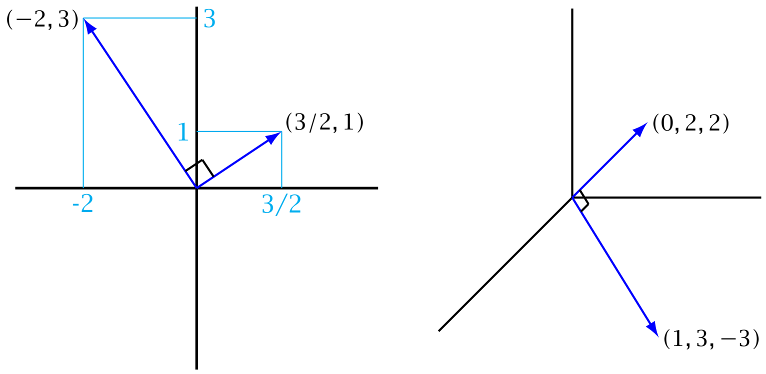
Dos vectors $\vec{u}$ i $\vec{v}$ de $\mathbb{R}^n$ són *ortogonals* si el producte escalar $\vec{u}\cdot\vec{v}$ és igual a zero. Per exemple, els vectors $(1,3,-3)$ i $(0,2,2)$ són ortogonals, perquè

$$(1,3,-3)\cdot(0,2,2) = 1\cdot 0 + 3\cdot 2 - 3\cdot 2 = 0 + 6 - 6 = 0$$

Si tenim dos vectors $\vec{u}$ i $\vec{v}$ diferents de zero i són ortogonals, llavors aquests dos vectors formen un angle recte, perquè el cosinus d'aquest angle és

$$\cos\alpha = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|} = 0$$

així que el mot *ortogonal* és un sinònim de *perpendicular*.



Vectors ortogonals en dues i tres dimensions

Observem, però, que el vector $\vec{0}$ és ortogonal a qualsevol altre vector (per exemple, $(0,0,0)\cdot(1,2,3) = 0+0+0 = 0$), encara que amb el vector zero no podem parlar d'angles.

Un conjunt de vectors és *ortogonal* si cada vector del conjunt és ortogonal a tots els altres. Si, a més a més, tots els vectors d'un conjunt ortogonal són unitaris, llavors direm que el conjunt és *ortonormal*.

2.3.3. El producte escalar (cas complex)

Quan els vectors són complexos, el producte escalar es defineix com[1]

$$\vec{u}\cdot\vec{v} = (u_1,u_2,\dots,u_n)\cdot(v_1,v_2,\dots,v_n) = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n$$

Per exemple, si $\vec{u} = (1+2\mathrm{i}, -1)$ i $\vec{v} = (1-\mathrm{i}, 2\mathrm{i})$, llavors,

$$\begin{aligned}\vec{u}\cdot\vec{v} &= (\overline{1+2\mathrm{i}})(1-\mathrm{i}) + (\overline{-1})(2\mathrm{i})\\ &= (1-2\mathrm{i})(1-\mathrm{i}) + (-1)(2\mathrm{i}) = (-1-3\mathrm{i}) + (-2\mathrm{i})\\ &= -1-5\mathrm{i}\end{aligned}$$

Amb aquesta definició, el producte escalar no compleix exactament les mateixes propietats que hem enumerat per al cas real, perquè no és una operació commutativa.

Propietats del producte escalar complex

1. $\vec{u}_1\cdot\vec{u}_2 = \overline{\vec{u}_2\cdot\vec{u}_1}$
2. $(\vec{u}_1+\vec{u}_2)\cdot\vec{v} = \vec{u}_1\cdot\vec{v} + \vec{u}_2\cdot\vec{v}$ (distributives)
   $\vec{u}\cdot(\vec{v}_1+\vec{v}_2) = \vec{u}\cdot\vec{v}_1 + \vec{u}\cdot\vec{v}_2$
3. $\alpha(\vec{u}_1\cdot\vec{u}_2) = (\overline{\alpha}\vec{u}_1)\cdot\vec{u}_2 = \vec{u}_1\cdot(\alpha\vec{u}_2)$
4. $\vec{0}\cdot\vec{u} = 0$ ($\vec{0}$ és ortogonal a tots els vectors)
5. Si $\vec{u}\neq\vec{0}$, llavors $\vec{u}\cdot\vec{u} > 0$ (definit positiu)

Provarem la més interessant: si $\vec{u}\neq\vec{0}$, llavors $\vec{u}\cdot\vec{u} > 0$.
Com que

$$\begin{aligned}\vec{u}\cdot\vec{u} &= (u_1,u_2,\dots u_n)\cdot(u_1,u_2,\dots u_n)\\ &= \overline{u_1}u_1 + \overline{u_2}u_2 + \cdots + \overline{u_n}u_n\\ &= |u_1|^2 + |u_2|^2 + \cdots + |u_n|^2 > 0 \quad \square\end{aligned}$$

La longitud i els conceptes relatius a l'ortogonalitat que hem estudiat adés són també vàlids en el cas complex: anomenem *norma* (o longitud) del vector complex $\vec{u}$ al nombre no negatiu $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$, i diem que els vectors $\vec{u}$ i $\vec{v}$ són *ortogonals* si $\vec{u}\cdot\vec{v} = 0$.

2.4. Exercicis

Exercici 2.1. (Operacions amb matrius) Donades les matrius $\mathsf{A} = \begin{bmatrix} 4 & 0\\ -1 & 1\end{bmatrix}$ i $\mathsf{B} = \begin{bmatrix} -1 & 2\\ -2 & 0\end{bmatrix}$ calculeu $\mathsf{A}+\mathsf{B}$, $3\mathsf{A}$ i $\mathsf{A}-2\mathsf{B}$.

Exercici 2.2. (Operacions amb vectors) Considerem els vectors de $\mathbb{R}^3$

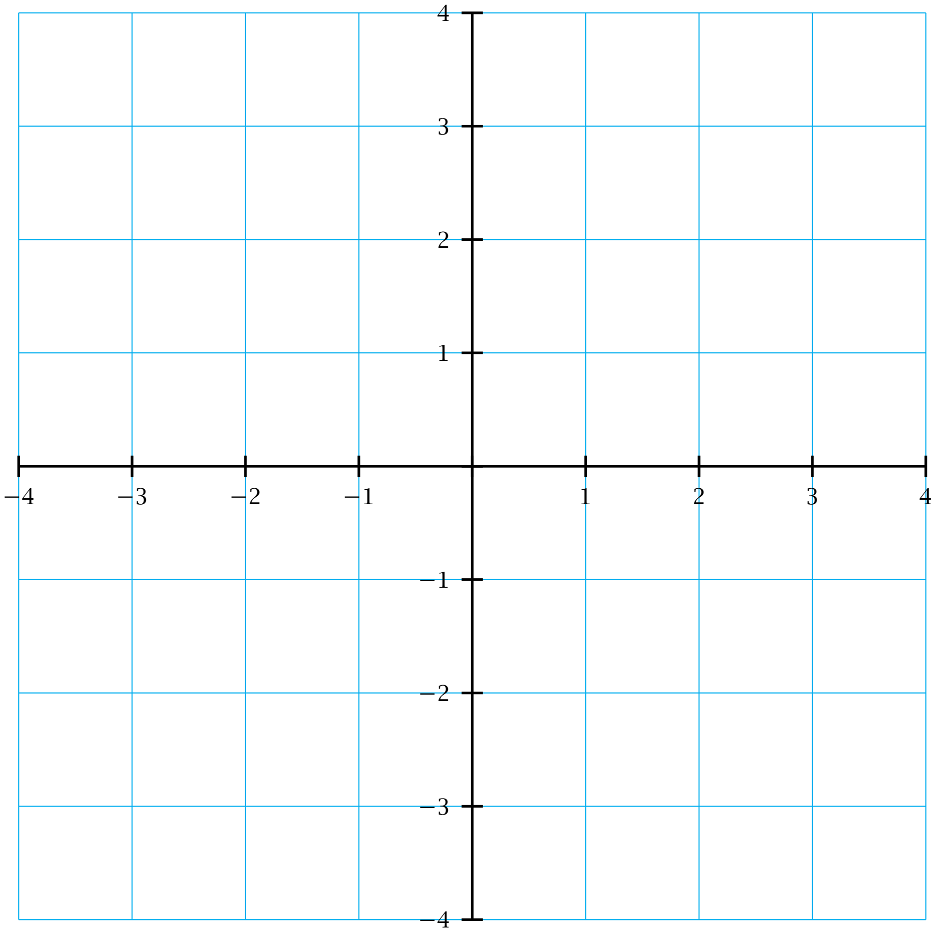
$$\vec{u}_1 = (1,-1,2), \vec{u}_2 = (0,1,-3), \vec{u}_3 = (-1,3,-8)$$

Calculeu (a) $\vec{u}_1+\vec{u}_2$, (b) $3\vec{u}_3$ i (c) $\vec{u}_1 - 2\vec{u}_2 + \vec{u}_3$.

[1] Recordem que $\overline{a+b\mathrm{i}} = a - b\mathrm{i}$ és el *conjugat* del nombre complex $a+b\mathrm{i}$.
En alguns textos es defineix el producte escalar conjugant el segon vector, en comptes del primer: $\vec{u}\cdot\vec{v} = u_1\overline{v_1} + u_2\overline{v_2} + \cdots + u_n\overline{v_n}$.

EXERCICI 2.3. (Combinacions lineals) Expresseu, si és possible, les matrius $A = \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix}$ i $B = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}$ com a combinació lineal de les matrius $A_1 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ i $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

EXERCICI 2.4. (Representació gràfica dels vectors en $\mathbb{R}^2$)
Siguen $\vec{u}_1 = (1, -2)$ i $\vec{u}_2 = (-1, 1)$. Representeu gràficament els vectors $\vec{u}_1$, $\vec{u}_2$, $2\vec{u}_1$, $-\vec{u}_2$, $3\vec{u}_1 + 3\vec{u}_2$, $\vec{u}_1 - 2\vec{u}_2$, $5\vec{u}_1 + 9\vec{u}_2$ i $-4\vec{u}_1 - 6\vec{u}_2$.



EXERCICI 2.5. (Combinacions lineals)

(a) Proveu que qualsevol vector de $\mathbb{R}^2$, $\vec{u} = (a, b)$, és combinació lineal de $\vec{u}_1 = (1, 1)$ i $\vec{u}_2 = (1, -1)$.

(b) És cert que qualsevol vector de $\mathbb{R}^3$ és combinació lineal de $(2, -1, -1)$, $(-1, 2, -1)$ i $(-1, -1, 2)$?

EXERCICI 2.6. (Norma d'un vector) Calculeu les longituds dels vectors $\vec{e}_1 = (1, 0)$, $\vec{e}_2 = (0, 1)$, $\vec{u} = (3, 4)$ i $\vec{v} = (-1, 2)$.

EXERCICI 2.7. (Angle entre dos vectors) Calculeu l'angle entre les següents parelles de vectors:

1. $\vec{u} = (\sqrt{3}, 1)$ i $\vec{v} = (0, 1)$
2. $\vec{u} = (\sqrt{3}, 1)$ i $\vec{v} = (2, 2)$
3. $\vec{u} = (1, 2, 3)$ i $\vec{v} = (1, 2, 6)$ (ací podeu fer servir la calculadora)
4. $\vec{u} = (1, 2, 1, 2)$ i $\vec{v} = (2, -1, -2, 1)$

EXERCICI 2.8. (Conjunts ortogonals) Digueu si els conjunts següents són ortogonals, ortonormals o cap de les dues coses.

1. $A = \{(1,1,1),(1,-1,0),(-1,-1,2)\}$
2. $B = \left\{\frac{1}{2}(1,-1,1,-1), \frac{1}{2}(-1,-1,1,1), \frac{1}{2}(-1,1,1,-1)\right\}$
3. $C = \{(-1,1,1),(1,-1,2),(1,-1,-1)\}$

EXERCICI 2.9. Trobeu tots els vectors de $\mathbb{R}^2$ que són ortogonals a $(1,2)$ i interpreteu-los geomètricament.

EXERCICI 2.10. (Producte escalar complex)

1. Calculeu el producte escalar $(1+\mathrm{i}, 1-\mathrm{i}) \cdot (2, \mathrm{i})$.
2. Comproveu que els vectors $\vec{u} = (-2+3\mathrm{i}, 1+5\mathrm{i})$ i $\vec{v} = (1-\mathrm{i}, \mathrm{i})$ són ortogonals.

EXERCICI 2.11. (Ortogonalitat entre vectors complexos) (a) Siguen $\vec{u}$ i $\vec{v}$ dos vectors de $\mathbb{C}^n$. Proveu que si $\vec{u}$ és ortogonal a $\vec{v}$, llavors $\vec{v}$ és ortogonal a $\vec{u}$. (b) Trobeu tots els vectors de $\mathbb{C}^2$ que són ortogonals al vector $(1, \mathrm{i})$.

EXERCICI 2.12. (Vectors unitaris) Si $\vec{u}$ i $\vec{v}$ són vectors unitaris (de longitud 1), calculeu els productes escalars

$$\vec{u} \cdot (-\vec{u}) \qquad (\vec{u}+\vec{v}) \cdot (\vec{u}-\vec{v}) \qquad (\vec{u}+2\vec{v}) \cdot (\vec{u}-2\vec{v})$$

EXERCICI 2.13. (Projecció ortogonal) Donats els vectors $\vec{u} = (1,1)$ i $\vec{v} = (1,5)$, trobeu el valor de α perquè el vector $\vec{v} - \alpha\vec{u}$ siga ortogonal a $\vec{u}$. Representeu sobre un diagrama cartesià els vectors $\vec{u}$, $\vec{v}$, $\vec{v} - \alpha\vec{u}$ i $\alpha\vec{u}$.

3. Multiplicació de matrius

Multiplícate por cero
Bart Simpson

Ací estudiem la multiplicació de matrius.
Tot i que els estudiants ja hi esteu familiaritzats, el que més ens interessa és interpretar el producte de matrius des de diversos punts de vista, sobre tot com a operacions amb les columnes o les files de les matrius.

3.1. El producte matriu-vector

El producte d'una matriu A per un vector $\vec{b}$ és la combinació lineal de les columnes de la matriu que es compon fent servir les components del vector com a pesos. Per exemple, si

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 3 & -1 & 5 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}$$

llavors

$$\mathsf{A}\vec{b} = \begin{bmatrix} 1 & 1 & 0 \\ 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 1\begin{bmatrix} 1 \\ -1 \end{bmatrix} - 4\begin{bmatrix} 0 \\ 5 \end{bmatrix} \tag{3.1}$$

$$= \begin{bmatrix} 2 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ -20 \end{bmatrix} \tag{3.2}$$

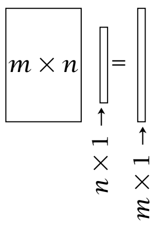
$$= \begin{bmatrix} 3 \\ -15 \end{bmatrix}$$

En general,

Si $\vec{a}_1$, $\vec{a}_2$, $\dots\vec{a}_n$ són les columnes de A i $\vec{b} = (b_1, b_2, \dots, b_n)$, llavors

$$\mathsf{A}\vec{b} = b_1\vec{a}_1 + b_2\vec{a}_2 + \cdots + b_n\vec{a}_n$$

- ☞ El producte d'una matriu per un vector és un vector.
- ☞ Per a poder fer el producte $\mathsf{A}\vec{b}$, el nombre de columnes de A ha de coincidir amb el nombre de components de $\vec{b}$ (perquè hem de multiplicar cada columna per un element del vector). Per exemple, si A és una matriu $m \times n$, aleshores $\vec{b}$ ha de ser un vector amb n components. Aleshores, el producte $\mathsf{A}\vec{b}$ és un vector amb m components.



- ☞ És molt important que recordem que el producte d'una matriu per un vector és una combinació lineal de les columnes de la matriu.

3.2. El producte fila-matriu

El producte d'una matriu fila U per una matriu A és el vector fila AB que s'obté com a combinació lineal de les files de la matriu A posant-hi, com a pesos, les entrades de U. Per exemple, si

$$\mathsf{A} = \begin{bmatrix} 2 & 1 \end{bmatrix} \qquad \mathsf{B} = \begin{bmatrix} 1 & 1 & 0 \\ 3 & -1 & 5 \end{bmatrix}$$

llavors

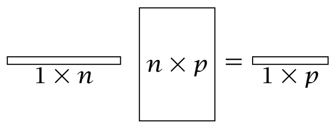
$$\mathsf{AB} = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 3 & -1 & 5 \end{bmatrix} = 2\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} + 1\begin{bmatrix} 3 & -1 & 5 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 3 & -1 & 5 \end{bmatrix}$$
$$= \begin{bmatrix} 5 & 1 & 5 \end{bmatrix}$$

En general,

Si $\mathsf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ i $\mathsf{B}_1, \mathsf{B}_2, \ldots \mathsf{B}_n$ són les files de B, llavors

$$\mathsf{AB} = a_1\mathsf{B}_1 + a_2\mathsf{B}_2 + \cdots + a_n\mathsf{B}_n$$

- ☞ El producte d'una matriu fila per una matriu és una matriu fila.
- ☞ Per a poder fer el producte AB, el nombre d'entrades de A ha de coincidir amb el nombre de files de B. Per exemple, si A és una matriu $1 \times n$, B ha de ser una matriu $n \times p$. Aleshores, el producte AB és una matriu fila $1 \times p$.



- ☞ El vector AB és una combinació lineal de les files de A.

3.3. El producte fila-columna

El cas més simple de producte de dues matrius és el producte d'una fila per una columna. Observem que aquest producte es pot veure com un producte matriu-vector o, també, com un producte fila-matriu. En qualsevol cas, és clar que el que hem de fer, per calcular-lo, és component a component i sumar els resultats. Per exemple,

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = 1 \cdot 1 + (-1) \cdot (-2) = 3$$

Si $\mathsf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$ i $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix}$, llavors $\mathsf{A}\vec{b} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

- ☞ El producte fila-columna és una matriu 1×1, però això ho identificarem amb un escalar.

3.4. El producte matriu-matriu

La regla pràctica que sol donar-se del producte de dues matrius (i la que segurament coneixeu) és la següent:

Per a multiplicar les matrius reals A i B es fa el producte de cada fila de A per cada columna de B, de manera que l'element situat en la fila i i en la columna j de la matriu AB és el producte de la fila i de A per la columna j de B.

Exemple 3.1.

Producte de les matrius

$$\mathsf{A} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \qquad \mathsf{B} = \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix}$$

Per a multiplicar aquestes dues matrius seguirem els passos següents:

- Multipliquem la primera fila de A per la primera columna de B: $\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = 3$ i posem aquest nombre en la posició $(1,1)$ de la matriu AB:

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & \\ & \\ & \end{bmatrix}$$

- Fem el mateix amb la primera fila de A i la segona columna de B i posem el resultat en la posició $(1,2)$:

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ & \\ & \end{bmatrix}$$

  i això completa la primera fila de AB.

- Multipliquem la segona fila de A per la primera columna de B:

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & \\ & \end{bmatrix}$$

- Segona fila per segona columna,

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ & \end{bmatrix}$$

- Tercera fila per primera columna,

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ 0 & \end{bmatrix}$$

- I, finalment, tercera fila per segona columna,

$$AB = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ 0 & 4 \end{bmatrix}$$

Com que hem de fer els productes de les files de A per les columnes de B cal que el nombre de components dels vectors fila de la primera matriu coincidisca amb el nombre de components dels vectors columna de la segona, és a dir, que *el nombre de columnes de* A *ha de ser el mateix que el nombre de files de* B.

☞ El producte d'una matriu $m \times n$ per una matriu $n \times p$ és una matriu $m \times p$.

$$AB = \begin{bmatrix} A_1 \\ A_2 \\ \cdots \\ A_m \end{bmatrix} \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} A_1\vec{b}_1 & A_1\vec{b}_2 & \cdots & A_1\vec{b}_p \\ A_2\vec{b}_1 & A_2\vec{b}_2 & \cdots & A_2\vec{b}_p \\ \cdots & & & \\ A_m\vec{b}_1 & A_m\vec{b}_2 & \cdots & A_m\vec{b}_p \end{bmatrix}$$

Així és com es defineix normalment el producte de dues matrius. Però a nosaltres ens interessa entendre el producte *columna a columna*, fent combinacions lineals de les columnes de B, o bé *fila a fila*, fent combinacions lineals de les files de A.

Tornem a l'exemple 3.1.: multiplicar les matrius $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix}$ i $B = \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix}$.

Producte per columnes Aquest producte es pot fer multiplicant A per cada columna de B:

$$A\vec{b}_1 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} \qquad A\vec{b}_2 = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$$

$$AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ 0 & 4 \end{bmatrix}$$

En general, si B té p columnes,

$$AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix}$$

Producte per files Alternativament, també podem multiplicar cada fila de A per B:

$$A_1B = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \end{bmatrix}$$

$$A_2B = \begin{bmatrix} 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} -4 & 0 \end{bmatrix}$$

$$A_3B = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \end{bmatrix}$$

$$AB = \begin{bmatrix} A_1B \\ A_2B \\ A_3B \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ 0 & 4 \end{bmatrix}$$

En general, si A té m files,

$$AB = \begin{bmatrix} A_1B \\ A_2B \\ \cdots \\ A_mB \end{bmatrix}$$

Columnes per files I encara hi ha una altra opció: si multipliquem cada columna de A per la fila de B corresponent i les sumem, també obtindrem el producte AB:

$$\vec{a}_1B_1 + \vec{a}_2B_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 2 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ -4 & 0 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -4 & 0 \\ 0 & 4 \end{bmatrix}$$

En general,

$$AB = \vec{a}_1B_1 + \vec{a}_2B_2 + \cdots + \vec{a}_nB_n$$

El quadre 3.1 resumeix tot el que hem vist a prop del producte de matrius.

3.4.1. Propietats del producte

El quadre següent enumera algunes propietats del producte de dues matrius.

1. $(AB)C = A(BC)$ (associativa)
2. $(A + B)C = AC + BC$ (distributiva per l'esquerra)
3. $A(B + C) = AB + AC$ (distributiva per la dreta)
4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$ (associativitat amb el producte escalar-matriu)

☞ Cal tenir molt en compte que la multiplicació de matrius no té la propietat commutativa. Per exemple, si $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ i $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$ aleshores

$$AB = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \qquad BA = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 1 & 3 \end{bmatrix}$$

☞ Per a poder fer el producte AB el nombre de columnes de A ha de coincidir amb el de files de B, per exemple, si A és una matriu $m \times n$, aleshores B ha de ser una matriu $n \times p$. Aleshores, el producte AB és una matriu $m \times p$.

$$\boxed{\quad m \times n \quad}\ \boxed{\begin{matrix} \\ \\ n \times p \\ \\ \\ \end{matrix}} = \boxed{\ m \times p\ }$$

1. **Files per columnes:**

   Cada element de la matriu AB és el producte d'una fila de A per una columna de B (per exemple, l'element $(2,3)$ de AB és el producte de la fila 2 de A per la columna 3 de B).

$$\begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \\ \cdots \\ \mathsf{A}_m \end{bmatrix} \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} \mathsf{A}_1\vec{b}_1 & \mathsf{A}_1\vec{b}_2 & \cdots & \mathsf{A}_1\vec{b}_p \\ \mathsf{A}_2\vec{b}_1 & \mathsf{A}_2\vec{b}_2 & \cdots & \mathsf{A}_2\vec{b}_p \\ \cdots & & & \\ \mathsf{A}_m\vec{b}_1 & \mathsf{A}_m\vec{b}_2 & \cdots & \mathsf{A}_m\vec{b}_p \end{bmatrix}$$

☞ Cada entrada de AB és un producte fila-columna.

2. A **per columnes de** B**:**

   Cada columna de AB és el producte de la matriu A per una columna de B:

$$\mathsf{AB} = \mathsf{A}\begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} \mathsf{A}\vec{b}_1 & \mathsf{A}\vec{b}_2 & \cdots & \mathsf{A}\vec{b}_p \end{bmatrix}$$

☞ Cada columna de AB és una combinació lineal de les columnes de A.

3. **Files de** A **per** B**:**

   Cada fila de AB és el producte d'una fila de A per la matriu B:

   Si A_1, A_2, ..., A_m són les files de A, llavors

$$\mathsf{AB} = \begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \\ \cdots \\ \mathsf{A}_m \end{bmatrix} \mathsf{B} = \begin{bmatrix} \mathsf{A}_1\mathsf{B} \\ \mathsf{A}_2\mathsf{B} \\ \cdots \\ \mathsf{A}_m\mathsf{B} \end{bmatrix}$$

☞ Cada fila de AB és una combinació lineal de les files de B.

4. **Columnes per files:**

$$\mathsf{AB} = \vec{a}_1\mathsf{B}_1 + \vec{a}_2\mathsf{B}_2 + \cdots + \vec{a}_n\mathsf{B}_n$$

☞ AB és una suma de productes columna-fila.

Quadre 3.1: Diverses interpretacions del producte de dues matrius així que $\mathsf{AB} \neq \mathsf{BA}$. En realitat, aquest resultat no ens ha de sorprendre, si recordem que les columnes de AB són combinacions lineals de les columnes de A, mentre que les columnes de BA ho són de les columnes de B.

El fet que el producte no siga commutatiu té com a conseqüència que algunes simplificacions que estem acostumats a fer mecànicament quan treballem amb nombres no es puguen traslladar al càlcul amb matrius. Vegem-ne algunes:

1. La igualtat $\mathsf{ABA} = \mathsf{A}^2\mathsf{B}$ no té perquè ser correcta, perquè es basa en el següent *raonament*: $\mathsf{ABA} = \mathsf{AAB}$; és a dir, estem suposant A commuta amb B.
2. Tampoc no és cert sempre que $(\mathsf{A}+\mathsf{B})^2 = \mathsf{A}^2 + 2\mathsf{AB} + \mathsf{B}^2$; en realitat, $(\mathsf{A}+\mathsf{B})^2 = \mathsf{A}^2 + \mathsf{AB} + \mathsf{BA} + \mathsf{B}^2$.
3. Ni tampoc podem afirmar que $(\mathsf{AB})^2 = \mathsf{A}^2\mathsf{B}^2$.

3.5. Exercicis

Exercici 3.1. (Un producte matriu-vector) Calculeu el producte $\mathsf{A}\vec{b}$, essent $\mathsf{A} = \begin{bmatrix} 4 & 1 \\ 2 & -5 \\ 0 & 6 \end{bmatrix}$ i $\vec{b} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, (a) element a element i (b) fent combinacions lineals de les columnes de A.

Exercici 3.2. Siga $\mathsf{A} = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & 8 & -1 & 7 \end{bmatrix}$ i $\vec{x} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}$. Escriviu el vector $\vec{b} = \mathsf{A}\vec{x}$ com a combinació lineal de les columnes de A.

Exercici 3.3. Donades les matrius $\mathsf{A} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \\ -1 & 1 \end{bmatrix}$ i $\mathsf{B} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$, calculeu el producte AB de quatre maneres diferents: (a) fent productes de les files de A per les columnes de B, (b) fent combinacions lineals de les columnes de A, (c) fent combinacions lineals de les files de B i (d) fent productes de les columnes de A per les files de B.

Exercici 3.4. En cadascun dels casos següents, què podem dir de la matriu AB? Justifiqueu les respostes.

(a) si la primera fila de A és nul·la

(b) si la primera fila de A és $\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}$

(c) si la primera fila de A és $\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \end{bmatrix}$

(d) si la primera columna de B és $(0, 1, 0, \dots, 0)$

(e) Si $\mathsf{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$

(f) Si $\mathsf{B} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$

Exercici 3.5. Donades les matrius

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 3 & -1 & 1 \end{bmatrix} \quad \mathsf{B} = \begin{bmatrix} -3 & 2 \\ 1 & -1 \end{bmatrix} \quad \mathsf{C} = \begin{bmatrix} 2 & 1 & -3 \end{bmatrix} \quad \mathsf{D} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

calculeu tots els productes que siguen possibles.

Exercici 3.6. Calculeu totes les potències A^n, $n \geq 1$ de la matriu $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Exercici 3.7. Trobeu totes les matrius B que commuten amb la matriu $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Exercici 3.8. **(Producte de matrius i operacions elementals)** Calculeu els productes següents fent combinacions lineals de les files de la matriu $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$.

(a) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

Exercici 3.9. **(Matrius inverses i determinants)** Calculeu el producte AB essent

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Exercici 3.10. **(Matrius transposades, matrius simètriques i matrius ortogonals)** La transposada de la matriu A és la matriu A^t les files de la qual són les columnes de la matriu A (la primera fila de A^t és igual a la primera columna de A, la segona fila de A^t és igual a la segona columna de A, etc.).

(a) Quina són les matrius transposades de les matrius següents?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

Es diu que la matriu A és simètrica si és igual a la seua matriu transposada.

(b) Quines de les matrius de l'apartat anterior són simètriques?

Es diu que la matriu real quadrada A és ortogonal si les seues columnes formen un conjunt de vectors ortonormal.

(c) Proveu que la matriu $A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ és ortogonal.

(d) Si la matriu A és ortogonal, què podem dir del producte A^tA?

4. Equacions i sistemes lineals

Els continguts d'aquesta lliçó i les següents són molt simples i, probablement, ja els coneixeu dels vostres estudis anteriors; segurament esteu familiaritzats amb el mètode d'eliminació per a discutir i resoldre sistemes d'equacions lineals i amb la classificació dels sistemes segons el nombre de solucions que tenen. El que trobareu de nou és el concepte de matrius esglaonades i esglaonades reduïdes, que tampoc no són gaire complicats.
Tot i això, és molt important que estudieu la lliçó amb molt de compte, perquè en aquests continguts es troben les claus principals de l'Àlgebra Lineal.

4.1. Equacions lineals

Començarem amb una sola equació lineal amb dues incògnites:
Exemple 4.1.

Estudi de l'equació
$$x_1 + 2x_2 = 3 \tag{4.1}$$

Aquesta equació és *lineal* perquè les úniques operacions que fem amb les incògnites consisteixen a multiplicar-les per una constant i sumar els resultats. Qualsevol equació que requerisca un altre tipus d'operació no és lineal.[2] En gran mesura, l'Àlgebra Lineal és l'estudi de les equacions lineals.

És clar que l'equació (4.1) té moltes solucions, com ara, aquesta: $x_1 = 3$, $x_2 = 0$; o aquesta altra: $x_1 = 0$, $x_2 = 3/2$ (simplement substituint els valors de x_1 i x_2 donats es comprova de seguida que efectivament són solucions de l'equació).

De fet, aquesta equació té infinites solucions, que podem trobar fàcilment: aïllant en (4.1) la primera incògnita tindrem
$$x_1 = 3 - 2x_2 \tag{4.2}$$
de manera que donant a x_2 qualsevol valor podem *calcular* el valor adequat de x_1 per a obtenir la solució. Per exemple, per a $x_2 = 0$ obtindrem $x_1 = 3 - 2 \cdot 0 = 3$, així que retrobem la nostra primera solució: $x_1 = 3$, $x_2 = 0$; en canvi, si partim de $x_2 = 1$ obtindrem $x_1 = 3 - 2 \cdot 1 = 1$ i podem assegurar que $x_1 = 1$, $x_2 = 1$ és una altra solució.

Atés que cada solució de l'equació (4.1) és un parell de nombres, x_1, x_2, resulta molt convenient veure aquesta solució com el vector (x_1, x_2).

D'acord amb aquesta notació, les tres solucions *particulars* de l'equació (4.1) que ja coneixem són aquestes:
$$\begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 3/2 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Aquestes solucions s'anomenen *particulars*, pel fet que n'hi ha moltes. Anomenarem solució *general* a una expressió que represente totes les solucions particulars; per a obtenir-la, tornem a la fórmula (4.2). Si α representa qualsevol nombre, llavors podem reescriure aquesta fórmula de la manera següent,
$$\begin{aligned} x_1 &= 3 - 2\alpha \\ x_2 &= \alpha \end{aligned}$$
de manera que el conjunt de totes les solucions de l'equació (4.1) és
$$\mathscr{S} = \{(3 - 2\alpha, \alpha) \,:\, \alpha \in \mathbb{R}\}$$
així que anomenarem solució general de l'equació al vector genèric $(3 - 2\alpha, \alpha)$. Com que ací α representa un nombre qualsevol, direm que α és un *paràmetre* i que hem expressat la solució general de l'equació (4.1) en *forma paramètrica*.

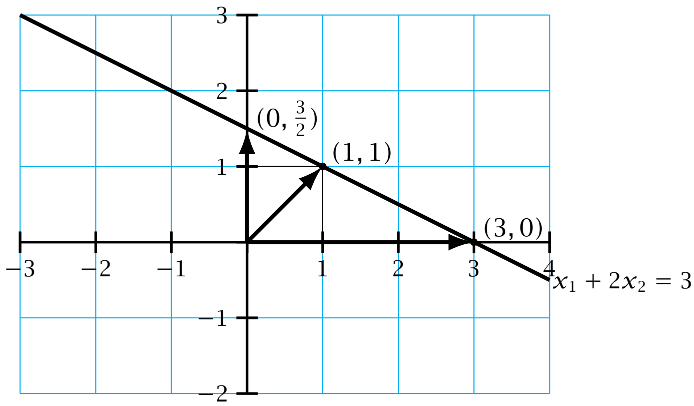
Per simple que parega, el fet d'interpretar les solucions de les equacions lineals com a vectors és un dels pilars de l'Àlgebra Lineal. La primera conseqüència d'aquesta interpretació és que podem fer servir el càlcul vectorial per a treballar amb les solucions; per exemple, la solució general de la nostra equació es pot expressar com
$$(x_1, x_2) = (3, 0) + \alpha(-2, 1)$$

[2] Per exemple, $x_1x_2 = 1$ o $x_1 + \cos x_2 = 0$ no són equacions lineals.

o, millor, el conjunt de totes les solucions d'aquesta equació és

$$\mathscr{S} = \{(3,0) + \alpha(-2,1) \colon \ \alpha \in \mathbb{R}\} \tag{4.3}$$

Per a visualitzar gràficament les solucions de l'equació (4.1), com que $x_1 + 2x_2 = 3$ és l'equació d'una recta en el pla i ja en coneixem diverses solucions, podem marcar-ne un parell i dibuixar la recta que les conté.

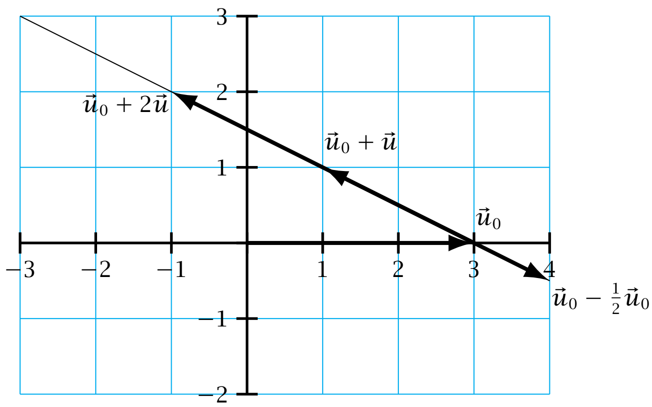


Les solucions de l'equació es corresponen amb tots els punts d'aquesta recta.

Que la solució general siga

$$\vec{x} = (3,0) + \alpha(-2,1)$$

significa que qualsevol solució de l'equació és la suma del vector $\vec{u}_0 = (3,0)$ amb un múltiple de $\vec{u} = (-2,1)$. El significat geomètric d'aquest fet es veu en el gràfic següent: les solucions són tots els punts de la recta que passa per $(3,0)$ i té la direcció del vector $\vec{u}$.



Així conclou l'estudi de l'equació (4.1). Ara tractarem una equació amb tres incògnites.

Exemple 4.2.

Solucions i interpretació gràfica de l'equació

$$2x_1 + 2x_2 - x_3 = 1 \tag{4.4}$$

Aquesta equació també té infinites solucions. Si aïllem la primera incògnita tindrem

$$x_1 = 1/2 - x_2 + (1/2)x_3 \tag{4.5}$$

així que ara encara sembla que tenim més llibertat per a trobar solucions: podem donar qualsevol valor a x_2 i a x_3 i, llavors, calcular el valor adequat de x_1 per a obtenir una solució. Per exemple, si posem

$x_2 = x_3 = 0$ haurà de ser $x_1 = 1/2$, de manera que $(1/2, 0, 0)$ és una solució de l'equació (4.4). Altres eleccions de x_2 i x_3 ens proporcionen aquestes altres solucions particulars:

$$\begin{bmatrix} 0 \\ 1/2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Per a determinar el conjunt de totes les solucions, assignarem un parell de paràmetres a les variables x_2 i x_3, com ara $x_2 = \alpha_1, x_3 = \alpha_2$, i obtindrem $x_1 = 1/2 - x_2 + (1/2)x_3 = 1/2 - \alpha_1 + (1/2)\alpha_2$, de manera que el conjunt de les solucions és aquest:

$$\mathscr{S} = \{(1/2 - \alpha_1 + (1/2)\alpha_2, \alpha_1, \alpha_2) : \alpha_1, \alpha_2 \in \mathbb{R}\}$$

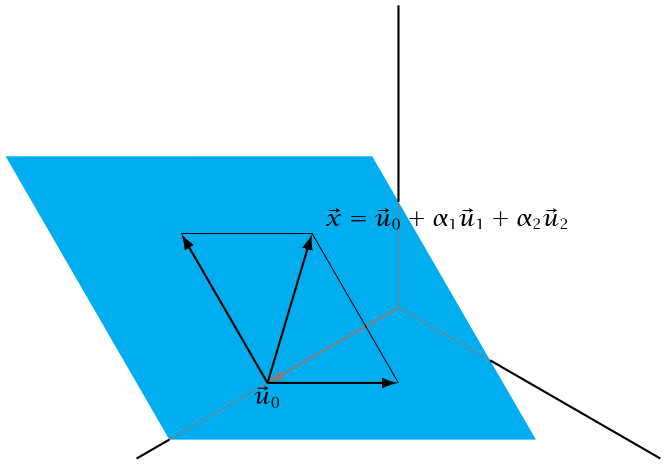
o bé,

$$\mathscr{S} = \{(1/2, 0, 0) + \alpha_1(-1, 1, 0) + \alpha_2(1/2, 0, 1) : \alpha_1, \alpha_2 \in \mathbb{R}\} \tag{4.6}$$

i direm que la solució general d'aquesta equació és

$$\vec{x} = (1/2, 0, 0) + \alpha_1(-1, 1, 0) + \alpha_2(1/2, 0, 1)$$

Així com en l'exemple anterior el conjunt de totes les solucions es podia interpretar com una recta en el pla, ara l'equació (4.4) representa un pla en l'espai tridimensional. Tots els vectors de posició dels punts d'aquest pla es poden obtenir sumant el vector $\vec{u}_0 = (1/2, 0, 0)$ amb una combinació lineal dels vectors $\vec{u}_1 = (-1, 1, 0)$ i $\vec{u}_2 = (1/2, 0, 1)$.



Sembla clar que qualsevol equació lineal es pot resoldre de la mateixa manera, per moltes incògnites que tinga. Únicament pot haver-hi algun problema quan els coeficients (els nombres que multipliquen les incògnites) són zero. Per exemple, en l'equació

$$0x_1 + 2x_2 + 3x_3 = 5$$

no podem aïllar la primera incògnita (de fet, aquesta incògnita pot ser invisible, si escrivim l'equació com $2x_2 + 3x_3 = 5$), però si la segona,

$$x_2 = \frac{5}{2} - \frac{3}{2}x_3$$

de manera que ara són x_1 i x_3 les incògnites a les quals donarem valors arbitraris i les solucions es poden expressar com $x_1 = \alpha_1$, $x_2 = 5/2 - (3/2)\alpha_2$, $x_3 = \alpha_2$. Així que el conjunt de totes les solucions és

$$\mathscr{S} = \{(0, 5/2, 0) + \alpha_1(1, 0, 0) + \alpha_2(0, -3/2, 1) : \alpha_1, \alpha_2 \in \mathbb{R}\}$$

Així, doncs, per a resoldre una equació lineal l'únic que hem de fer és aïllar una incògnita el coeficient de la qual no siga zero i substituir totes les altres incògnites per paràmetres. Tot i això, hi ha un parell de casos *patològics*:

1. En l'equació

$$0x_1 + 0x_2 = 0$$

no es pot aïllar cap incògnita! Però tot i això, és evident que qualsevol vector bidimensional n'és solució, és a dir, que la solució és $x_1 = \alpha_1$, $x_2 = \alpha_2$.

2. I tampoc podem aïllar cap incògnita en l'equació

$$0x_1 + 0x_2 = a, \quad (a \neq 0)$$

Però el que passa ara és que aquesta equació no té cap solució.

El quadre següent resumeix totes les possibilitats que pot haver-hi:

Discussió i resolució de l'equació lineal

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

1. Si algun coeficient a_i és no nul, llavors l'equació és *compatible* (o *consistent*), és a dir, té alguna solució.

   Les solucions s'obtenen aïllant x_i i substituint totes les altres incògnites per paràmetres.

2. Si tots els coeficients a_1, a_2,..., a_n són zero, llavors

   2.1. Si $b \neq 0$ llavors, l'equació és *incompatible* (o *inconsistent*), és a dir, no té cap solució.

   2.2. Si $b = 0$ llavors, l'equació és *compatible* i qualsevol conjunt de nombres $x_1, x_2, \ldots, x_n$ és solució.

4.2. Sistemes d'equacions lineals

Ara ens ocuparem dels sistemes d'equacions lineals. Es tracta de trobar les solucions comunes a dues o més equacions.

Exemple 4.3.

Estudi del sistema d'equacions lineals $\begin{cases} x_1 + 2x_2 = 3 \\ x_1 - x_2 = -3 \end{cases}$

El sistema es pot resoldre molt fàcilment sense més que fer una senzilla operació d'eliminació: si a la segona equació li restem la primera obtindrem un nou sistema que és equivalent al sistema inicial però en el qual la segona equació només té una incògnita:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ -3x_2 &= -6 \end{aligned}$$

així que dividint la segona equació entre -3 obtenim el valor de la segona incògnita, $x_2 = 2$, i el sistema es converteix en

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ x_2 &= 2 \end{aligned}$$

Finalment, podem restar a la primera equació el doble de la segona,

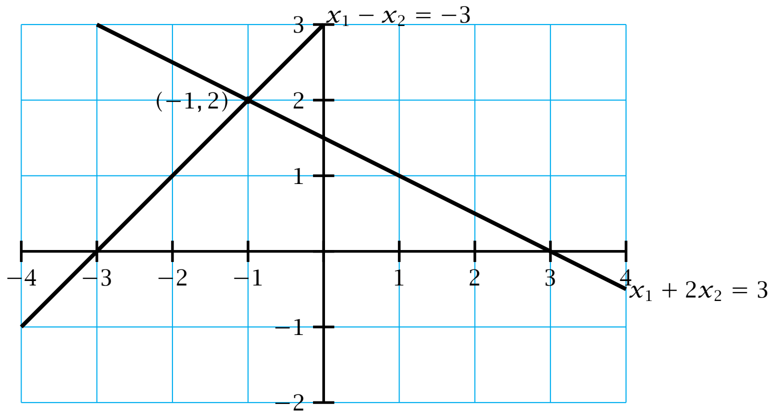
$$\begin{aligned} x_1 &= -1 \\ x_2 &= 2 \end{aligned}$$

En definitiva, el nostre sistema té una solució única: $(x_1, x_2) = (-1, 2)$. Com que sempre que es resol un problema cal comprovar els resultats, convé que ens assegurem que la solució és correcta substituint $x_1 = -1$, $x_2 = 2$ en l'equació original:

$$\begin{aligned} x_1 + 2x_2 &= -1 + 2 \cdot 2 = 3 \\ x_1 - x_2 &= -1 - 2 = -3 \end{aligned}$$

☞ En aquestes lliçons no sempre comprovarem els resultats, perquè això només serviria per a fer-los més tediosos, però sempre que resolgueu un problema heu de comprovar que els resultats són correctes.

En farem dues interpretacions gràfiques. En primer lloc, com que la solució general de cadascuna d'aquestes equacions representa una recta, resulta que l'única solució comuna a les dues equacions és la intersecció d'aquestes dues rectes.



Per a visualitzar geomètricament la solució des d'un altre punt de vista, més interessant, ens convé observar que el sistema

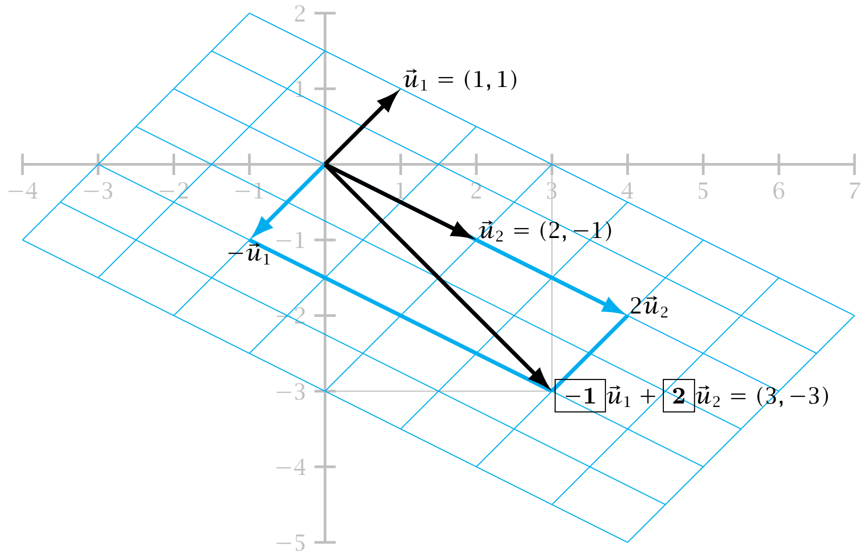
$$\begin{aligned} x_1 + 2x_2 &= 3 \\ x_1 - x_2 &= -3 \end{aligned}$$

és equivalent a l'*equació vectorial*

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

així que *trobar una solució del sistema lineal és equivalent a escriure el vector* $(3, -3)$ *com a combinació lineal dels vectors* $(1, 1)$ *i* $(2, -1)$. Per tant, la solució $x_1 = -1$, $x_2 = 2$ (que hem obtingut fa un moment) significa que podem obtenir el vector $(3, -3)$ amb la combinació lineal $-1(1, 1) + 2(2, -1)$:

$$-1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$



Si mirem de generalitzar el resultat que hem trobat en aquest exemple, sembla que un sistema de dues equacions lineals amb dues incògnites ha de tenir una solució única, atés que la solució de cada equació és una recta i llavors el punt intersecció de les dues rectes és la solució del sistema; o bé, mirant-ho des del punt de vista de les combinacions lineals dels vectors, sembla que qualsevol vector de $\mathbb{R}^2$ s'ha de poder escriure com a combinació lineal de dos vectors donats. Doncs bé, això és cert *quasi sempre*, però hi ha un parell de casos especials en els quals no és així. Els següents exemples ho posaran en clar.

Exemple 4.4.

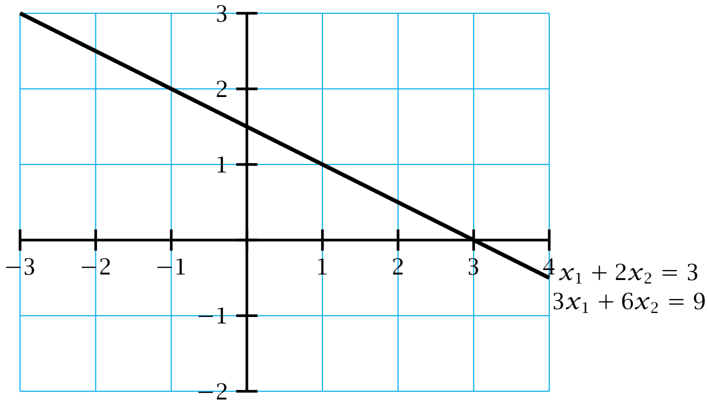
Estudi del sistema d'equacions lineals $\begin{cases} x_1 + 2x_2 = 3 \\ 3x_1 + 6x_2 = 9 \end{cases}$

Repetint l'estratègia d'eliminació, haurem de restar a la segona equació el triple de la primera, per tal d'eliminar-hi la incògnita x_1. Llavors obtenim el sistema següent:

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 0 &= 0 \end{aligned}$$

I la segona equació ha desaparegut! (perquè $3x_1 - 3x_1 = 0$, $6x_2 - 3(2x_2) = 0$ i $9 - 3 \cdot 3 = 0$), de manera que el sistema ha quedat reduït a una sola equació (la nostra ben coneguda equació de l'exemple 4.1.), la solució general de la qual és $(x_1, x_2) = (3, 0) + \alpha(-2, 1)$. En definitiva, el sistema té infinites solucions.

Evidentment, això passa perquè la segona equació és exactament el triple de la primera, de manera que en realitat no suposa cap restricció nova a les incògnites. Gràficament, si cada equació representa una recta, el que passa és que les dues equacions representen la mateixa recta.

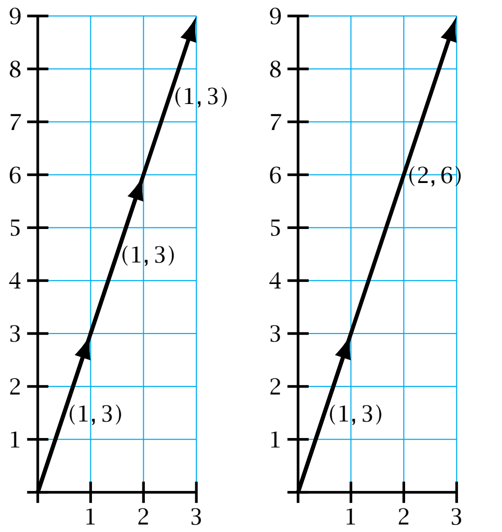


Per veure què passa amb la interpretació vectorial del problema, escrivim el sistema en forma vectorial:

$$\left.\begin{aligned} x_1 + 2x_2 &= 3 \\ 3x_1 + 6x_2 &= 9 \end{aligned}\right\} \iff x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$$

així que es tracta d'escriure el vector $(3,9)$ com a combinació lineal dels vectors $(1,3)$ i $(2,6)$, però això es pot fer de moltes maneres: per exemple, $(1,3)+(2,6)=(3,9)$; però també

$$
\begin{aligned}
3(1,3)+0(2,6)&=(3,9)\\
(1,3)+(2,6)&=(3,9)\\
0(1,3)+(3/2)(2,6)&=(3,9)\\
2(1,3)+(1/2)(2,6)&=(3,9)\\
&\dots
\end{aligned}
$$



$(3,9)=3(1,3)$ $\quad$ $(3,9)=(1,3)+(2,6)$

Exemple 4.5.

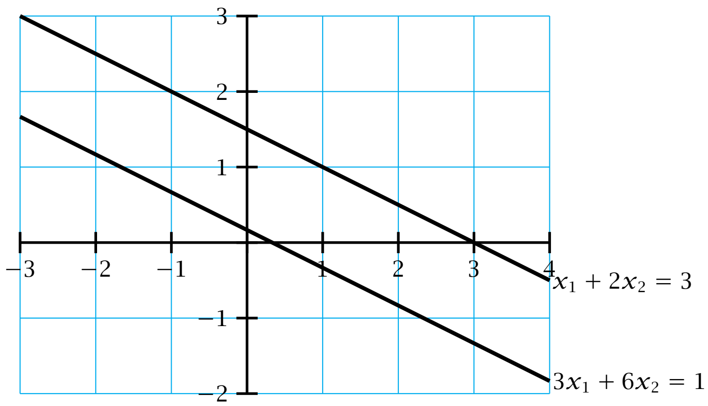
Estudi del sistema d'equacions lineals $\begin{cases} x_1+2x_2=3\\ 3x_1+6x_2=1\end{cases}$

Restem a la segona equació el triple de la primera, per a eliminar la primera incògnita,

$$
\begin{aligned}
x_1+2x_2&=3\\
0&=-8
\end{aligned}
$$

I el que passa ara és que la segona equació no pot tenir cap solució, així que el sistema tampoc no té solucions.

Si cada equació representa una recta, el problema és que les dues equacions del nostre exemple corresponen a rectes paral·leles, així que no tenen cap punt en comú:



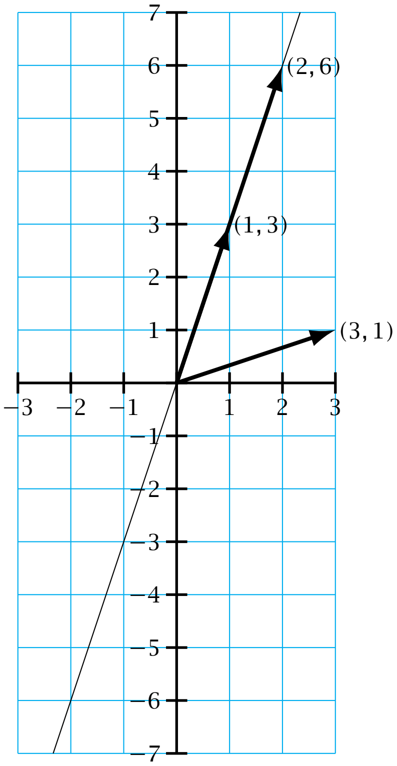
En aquest cas, la forma vectorial del sistema és aquesta:

$$x_1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

i el problema consisteix a expressar el vector $(3, 1)$ com a combinació lineal dels vectors $(1, 3)$ i $(2, 6)$. Però com que el vector $(2, 6)$ és proporcional a $(1, 3)$, resulta que tots dos determinen la mateixa recta, i només els vectors que es troben sobre aquesta recta (que són els de la forma $\alpha(1, 3)$) es poden expressar com a combinacions lineals d'aquells dos vectors. En altres paraules, les combinacions lineals del vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

són els vectors proporcionals a $\vec{u}_1$; com que el vector $\vec{u} = (3, 1)$ no és un múltiple de $\vec{u}_1$, llavors l'equació $x_1\vec{u}_1 + x_2\vec{u}_2 = \vec{u}$ no té cap solució.



El fet que $(2, 6)$ siga un múltiple de $(1, 3)$ fa que algunes equacions (com la de l'exemple 4.4.) tinguen infinites solucions i, en canvi, altres (com la de l'exemple 4.5.) no en tinguen cap!

Acabarem estudiant un sistema de dues equacions amb tres incògnites.

Exemple 4.6.

Resolució del sistema lineal $\begin{cases} 2x_1 + 2x_2 - x_3 = 1 \\ x_1 - x_2 + x_3 = 0 \end{cases}$

En primer lloc, eliminarem x_1 de la segona equació restant-hi la meitat de la primera,

$$\begin{aligned} 2x_1 + 2x_2 - \quad x_3 &= 1 \\ - 2x_2 + (3/2)x_3 &= -1/2 \end{aligned}$$

i també eliminarem x_2 en la primera equació, sumant-hi la segona equació:

$$\begin{aligned} 2x_1 \qquad + (1/2)x_3 &= 1/2 \\ - 2x_2 + (3/2)x_3 &= -1/2 \end{aligned}$$

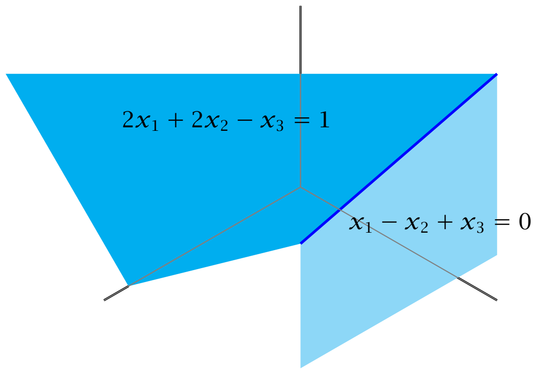
de manera que les dues equacions es poden resoldre aïllant x_1 i x_2 i assignant a x_3 un valor arbitrari. Si dividim la primera equació entre 2 i la la segona entre -2 encara serà més senzill obtenir la solució:

$$\begin{aligned} x_1 \quad + (1/4)x_3 &= 1/4 \\ x_2 - (3/4)x_3 &= 1/4 \end{aligned}$$

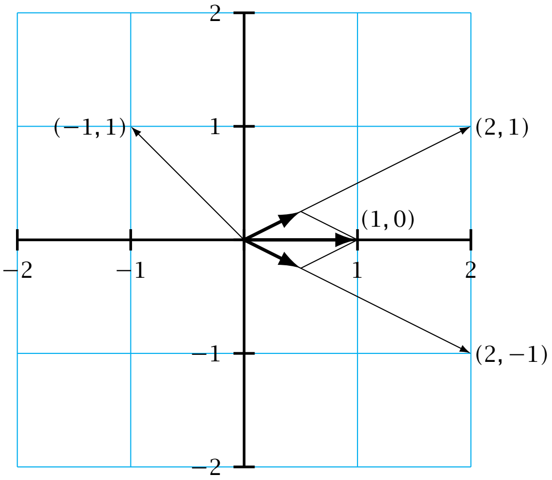
Així que, $x_1 = 1/4 - (1/4)x_3$, $x_2 = 1/4 + (3/4)x_3$, i és clar que la solució general d'aquest sistema és

$$\vec{x} = (1/4, 1/4, 0) + \alpha(-1/4, 3/4, 1)$$

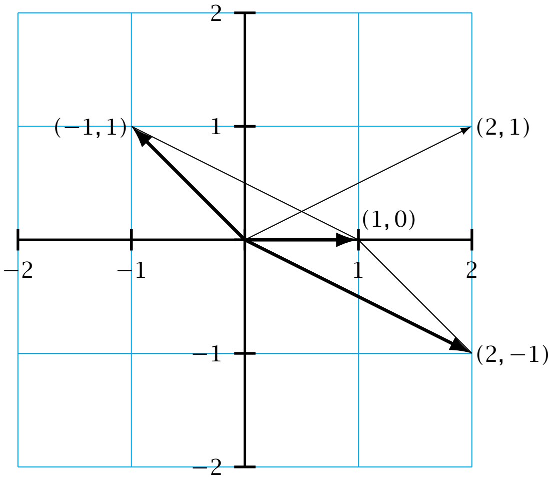
Gràficament, cada equació d'aquest sistema representa un pla, de manera que la solució del sistema és la recta intersecció dels dos plans.



Des d'un altre punt de vista, el fet que aquest sistema tinga moltes solucions vol dir que hi ha moltes maneres d'expressar el vector $(1,0)$ com a combinació lineal dels tres vectors $(2,1)$, $(2,-1)$ i $(1,-1)$. La figura següent en mostra un parell.



$$(1,0) = \frac{1}{4}(2,1) + \frac{1}{4}(2,-1)$$



$$(1,0) = (2,-1) + (-1,1)$$

4.2.1. Classificació dels sistemes lineals atenent al nombre de solucions

Els darrers exemples mostren que hi ha sistemes d'equacions lineals que tenen una solució única, d'altres que en tenen més d'una i d'altres que no en tenen cap. En conseqüència, podem classificar els sistemes en les següents categories:

Definicions 4.1.

1. *Un sistema d'equacions lineals és incompatible (o també, inconsistent) quan no té cap solució.*
2. *En cas que el sistema tinga alguna solució, aleshores l'anomenem compatible (o consistent). Els sistemes compatibles poden ser*
   (a) *Determinats, quan tenen només una solució.*
   (b) *Indeterminats, quan en tenen més d'una.*

4.2.2. Les formes vectorial i matricial d'un sistema d'equacions lineals. Matrius associades al sistema

En aquest apartat, introduirem una nova forma de representar els sistemes lineals (la més compacta i també la més interessant des del punt de vista de l'àlgebra lineal). El sistema de l'exemple 4.6.,

$$\begin{aligned} 2x_1 + 2x_2 - x_3 &= 1 \\ x_1 - x_2 + x_3 &= 0 \end{aligned}$$

és equivalent a l'equació vectorial

$$x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

o bé

$$x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 = \vec{b}$$

Aquesta és *la forma vectorial* del sistema.

Però, si recordem que el producte d'una matriu per un vector (columna) és una combinació lineal de les columnes de la matriu,

$$\begin{bmatrix} 2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

observarem que el sistema lineal també és equivalent a l'equació matricial

$$\begin{bmatrix} 2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

que podem expressar com

$$\mathsf{A}\vec{x} = \vec{b}$$

Aquesta és *la forma matricial* del sistema.

La matriu $\mathsf{A} = \begin{bmatrix} 2 & 2 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ és la *matriu de coeficients* del sistema, el *vector de termes independents* és $\vec{b} = (3, -3)$ i el *vector incògnita* és $\vec{x} = (x_1, x_2, x_3)$.

Per últim, anomenarem *matriu ampliada* a la matriu que s'obté afegint el vector columna $\vec{b}$ a la matriu A.

$$\left[\begin{array}{c|c} \mathsf{A} & \vec{b} \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 1 & -1 & 1 & -3 \end{array}\right]$$

Aquesta matriu conté tota la informació necessària a prop del sistema.[3]

4.3. Mètode de reducció de Gauss-Jordan (primera aproximació)

Fins ara, hem resolt només sistemes de dues equacions. Però l'estratègia d'eliminació (sumar o restar a una equació un múltiple adequat d'una altra) es pot fer servir per a resoldre molts altres sistemes d'equacions lineals. De fet, tots els sistemes lineals es poden discutir, i resoldre en cas que tinguen solució, mitjançant un mètode general molt simple, el mètode d'eliminació (o, millor, *de reducció*) de Gauss-Jordan.

La idea bàsica del mètode de Gauss-Jordan és eliminar x_1 de totes les equacions excepte de la primera; fer el mateix amb x_2, eliminat-lo de totes les equacions tret de la segona, i així successivament: En el millor cas, això conduirà a la separació de les incògnites, de manera que cada equació es podrà resoldre independentment de les altres.

En comptes de treballar directament sobre el sistema lineal ho farem amb la matriu ampliada. Tot i això, perquè s'entenga perfectament el procés, en el primer exemple treballarem en paral·lel amb el sistema lineal i la matriu ampliada.

Exemple 4.7.

Discussió i resolució del sistema lineal $\begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 - x_2 + x_3 = 0 \\ 2x_1 + 4x_2 + 4x_3 = 3 \end{cases}$

La matriu associada a aquest sistema és

$$\left[\begin{array}{c|c} \mathsf{A} & \vec{b} \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 0 \\ 2 & 4 & 4 & 3 \end{array}\right]$$

El primer pas consisteix a sumar (o restar) un múltiple adequat de la primera equació a la segona de tal manera que desaparega x_1 de la segona equació. Això es pot aconseguir, en aquest sistema, restant a la segona equació la primera (o restant la primera fila de la matriu ampliada a la segona fila):

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ - 2x_2 + 2x_3 &= -1 \\ 2x_1 + 4x_2 + 4x_3 &= 3 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ 2 & 4 & 4 & 3 \end{array}\right]$$

[3] La ratlla vertical entre les matrius A i $\vec{b}$ no té cap significat matemàtic; únicament ens serveix per a separar els vectors de coeficients del vector de termes independents. Com que en moltes aplicacions $\vec{b}$ serà una matriu en comptes d'un vector, aquesta ratlla ens serà molt útil.

Per a eliminar x_1 de la tercera equació (fila), hi restem el doble de la primera:

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ -2x_2 + 2x_3 &= -1 \\ 2x_2 + 6x_3 &= 1 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ 0 & 2 & 6 & 1 \end{array}\right]$$

El següent pas consistiria a eliminar x_2 de les equacions primera i tercera. Tot i això, ho farem només amb la tercera equació (la reducció de la primera la farem més endavant, perquè d'aquesta manera ens estalviarem alguns càlculs). Així que a la tercera equació hi sumem la segona

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ -2x_2 + 2x_3 &= -1 \\ 8x_3 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & -2 & 2 & -1 \\ 0 & 0 & 8 & 0 \end{array}\right]$$

Ara ja hem eliminat les incògnites *per sota* (x_1 l'hem eliminada en la segona i en la tercera equacions i x_2 en la tercera). Ara eliminarem *per damunt*. Començarem per x_3, eliminant-la en les dues primeres equacions: a la primera equació hi sumem la tercera dividida entre 8:

$$\begin{aligned} x_1 + x_2 \phantom{{}+2x_3} &= 1 \\ -2x_2 + 2x_3 &= -1 \\ 8x_3 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & -2 & 2 & -1 \\ 0 & 0 & 8 & 0 \end{array}\right]$$

i a la segona equació hi restarem la tercera dividida entre quatre:

$$\begin{aligned} x_1 + x_2 &= 1 \\ -2x_2 &= -1 \\ 8x_3 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 8 & 0 \end{array}\right]$$

Per últim, eliminem x_2 a la primera equació sumant-hi un mig de la segona:

$$\begin{aligned} x_1 &= 1/2 \\ -2x_2 &= -1 \\ 8x_3 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/2 \\ 0 & -2 & 0 & -1 \\ 0 & 0 & 8 & 0 \end{array}\right]$$

Ara cada equació conté una sola incògnita, així que la solució s'obté dividint la segona equació entre -2 i la tercera entre 8:

$$\begin{aligned} x_1 &= 1/2 \\ x_2 &= 1/2 \\ x_3 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

El sistema és determinat i la solució general és

$$\vec{x} = (1/2, 1/2, 0)$$

Les tres equacions representen tres plans, de manera que la intersecció de cada dos d'aquests plans és una recta.

El sistema és determinat perquè les tres rectes es tallen en un punt.

Exemple 4.8.

Discussió i resolució del sistema lineal $\begin{cases} \phantom{x_1 - {}} 3x_2 + 3x_3 = 1 \\ x_1 - x_2 - x_3 = 0 \\ x_1 + 2x_2 + 2x_3 = 1 \end{cases}$

Repetim la mateixa estratègia. La matriu ampliada és

$$\left[\begin{array}{ccc|c} 0 & 3 & 3 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & 2 & 2 & 1 \end{array}\right]$$

Aquest sistema presenta un problema que encara no havíem trobat: se suposa que hem d'eliminar x_1 en totes les equacions excepte la primera, però és precisament a la primera equació que el coeficient de x_1 és zero. Per resoldre aquest problema intercanviarem les dues primeres equacions (files).

$$\begin{aligned} x_1 - x_2 - x_3 &= 0 \\ 3x_2 + 3x_3 &= 1 \\ x_1 + 2x_2 + 2x_3 &= 1 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 3 & 3 & 1 \\ 1 & 2 & 2 & 1 \end{array}\right]$$

Eliminem x_1 de la tercera equació, restant-hi la primera (restem la primera fila a la tercera),

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 3 & 3 & 1 \\ 0 & 3 & 3 & 1 \end{array}\right]$$

Restem la segona fila a la tercera.

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Ara hauríem d'eliminar x_3 de les dues primeres equacions, però no ho podem fer, perquè aquesta incògnita ha desaparegut de la tercera equació. El que sí que es pot fer és eliminar x_2 a la primera equació, sumant-hi un terç de la segona.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/3 \\ 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Dividim la segona equació entre 3, perquè el coeficient de x_2 siga 1,

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Així doncs, el sistema original és equivalent a

$$\left.\begin{aligned} x_1 \phantom{{}+x_3} &= 1/3 \\ x_2 + x_3 &= 1/3 \end{aligned}\right\} \text{ o, en forma matricial, } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix}$$

Aquest sistema es resol aïllant les incògnites x_1 i x_2,

$$\begin{aligned} \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right) \vec{x} &= \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{x} &= \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \vec{x} \\ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} &= \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ x_3 \\ 0 \end{bmatrix} \end{aligned}$$

i donant a x_3 un valor arbitrari:

$$x_1 = 1/3, \quad x_2 = 1/3 - \alpha, \quad x_3 = \alpha$$

El sistema és indeterminat i la solució general, en forma vectorial, és aquesta:

$$\vec{x} = (1/3, 1/3, 0) + \alpha(0, -1, 1)$$

Exemple 4.9.

Discussió i solucions del sistema lineal $\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 0 \\ x_1 - 3x_2 - 3x_3 = 3 \end{cases}$

La matriu ampliada és aquesta:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 0 \\ 1 & -3 & -3 & 3 \end{array}\right]$$

Primer pas: a la segona i a la tercera files els restem la primera.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & -4 & -4 & 2 \end{array}\right]$$

Segon pas: restem el doble de la segona fila a la tercera.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -1 \\ 0 & 0 & 0 & 4 \end{array}\right]$$

Aquesta és la matriu ampliada del sistema

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

I com que, ací, la tercera equació és $0 = 4$, podem concloure que el sistema és incompatible.

Remarquem que en els tres exemples d'aquest apartat, per a eliminar les variables, hem fet sempre el mateix tipus de transformació del sistema: utilitzar una fila de la matriu per a modificar les altres. Aquest és el tipus d'operació elemental més important. En la fase final, si el sistema era compatible, únicament hem hagut de dividir cada equació pel primer coeficient per a trobar el valor de la incògnita corresponent. A més a més, en una ocasió hem hagut de permutar dues files de la matriu. D'això també en direm operacions elementals. En la unitat propera definirem definitivament les operacions elementals i descriurem adequadament l'algorisme de Gauss-Jordan.

4.4. Matrius esglaonades

En tots els exemples anteriors hem reduït el sistema lineal (o la matriu ampliada corresponent) a una forma especial que ens ha permés decidir si aquest sistema tenia solucions i, en aquest cas, calcular-les. Aquest tipus especial de matrius s'anomenen *esglaonades*.

Definició 4.2.

Una matriu és esglaonada si

1. *Si hi ha files nul·les i no nul·les llavors, les files nul·les estan per davall de les no nul·les, i*
2. *Si ha ha més d'una fila no nul·la, el primer element no nul (per l'esquerra) de cada fila, a partir de la segona, es troba més a la dreta que el primer element no nul de la fila anterior (en altres paraules, per davall del primer element no nul de cada fila només hi ha zeros). Aquest primer element no nul és un element principal o pivot de la matriu.*

Una matriu és esglaonada reduïda si és esglaonada i, a més,

3. *Tots els seus pivots són* 1 *(d'ací que també els anomenem uns principals).*
4. *Tots els elements situats per damunt dels uns principals són zeros.*

EXEMPLE 4.10.

Les matrius següents són esglaonades, però només la última és esglaonada reduïda:

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

En les matrius esglaonades anomenem *columnes principals* (o *columnes pivot*) a les que contenen un element principal. Si la matriu és esglaonada reduïda, totes les entrades de les columnes principals, tret de l'u principal, són zeros.

☞ L'objectiu de l'algorisme de Gauss-Jordan és transformar una matriu qualsevol en una que siga esglaonada reduïda mitjançant les transformacions que s'anomenen operacions elementals.

4.5. EXERCICIS

EXERCICI 4.1. Classifiqueu les equacions següents (és a dir, dieu si són consistents o inconsistents) i, en cas que siguen consistents, trobeu-ne totes les solucions. En tots els casos, les incògnites són x_1, x_2 i x_3.

(a) $x_1 + x_2 + x_3 = 1$ (b) $x_2 + x_3 = 0$ (c) $x_1 - x_3 = -1$

(d) $x_2 = 0$ (e) $0x_3 = 0$ (f) $0x_1 + 0x_2 + 0x_3 = 1$

EXERCICI 4.2. Discutiu i resoleu, en cas que siguen compatibles, els sistemes lineals següents. Utilitzeu l'algorisme de Gauss-Jordan tal com ho hem fet als exemples d'aquesta unitat.

$$\text{(a)}\ \begin{cases}3x_1 + 2x_2 + x_3 = 10\\ 2x_1 + x_2 \phantom{{}+x_3} = 4\\ x_1 \phantom{{}+2x_2} - x_3 = -2\end{cases}\qquad \text{(b)}\ \begin{cases}3x_1 + 2x_2 + x_3 = 10\\ 2x_1 + x_2 \phantom{{}+x_3} = 4\\ x_1 \phantom{{}+2x_2} - x_3 = 0\end{cases}\qquad \text{(c)}\ \begin{cases}3x_1 + 2x_2 + x_3 = 10\\ 2x_1 + x_2 \phantom{{}+x_3} = 4\\ x_1 + x_2 - x_3 = 0\end{cases}$$

EXERCICI 4.3. Trobeu tots els vectors que són ortogonals als dos vectors $(1, -1, 0)$ i $(1, 0, 1)$.

EXERCICI 4.4. Escriviu el sistema lineal

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1\\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2\end{aligned}$$

en forma vectorial i en forma matricial. Digueu quina és la matriu de coeficients i quina la matriu ampliada.

EXERCICI 4.5. Feu el mateix amb el sistema genèric

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

EXERCICI 4.6. Si la matriu de coeficients d'un sistema lineal és una matriu 7×5, quantes incògnites i quantes equacions té el sistema? Quants components té el vector de termes independents? I el vector incògnita?

EXERCICI 4.7. Determineu si cada una d'aquestes matrius és esglaonada, esglaonada reduïda o si no és esglaonada.

$$\text{(a)}\ \begin{bmatrix}0&1&0&3\\1&0&1&3\\0&0&0&0\end{bmatrix}\quad \text{(b)}\ \begin{bmatrix}1&0&0&0\\0&0&1&0\end{bmatrix}\quad \text{(c)}\ \begin{bmatrix}1&2&0&0\\0&2&0&0\\0&0&2&1\end{bmatrix}\quad \text{(d)}\ \begin{bmatrix}1&2&0&0\\0&1&0&1\end{bmatrix}$$

$$\text{(e)}\ \begin{bmatrix}1&2&0&0\\0&0&2&1\end{bmatrix}\quad \text{(f)}\ \begin{bmatrix}0&2&0&0\\0&1&2&1\\0&0&0&0\end{bmatrix}\quad \text{(g)}\ \begin{bmatrix}0&0&0&0\\0&0&0&0\end{bmatrix}\quad \text{(h)}\ \begin{bmatrix}1&0&2&5\\0&1&3&-3\end{bmatrix}$$

5. Matrius elementals. Algorisme de Gauss-Jordan

Divide et impera
(sentència atribuïda a Juli Cèsar)

En la lliçó anterior hem fet servir allò que anomenem operacions elementals per a transformar els sistemes lineals (o les matrius associades a aquests sistemes) a la forma esglaonada reduïda, que ens permet decidir si el sistema és compatible i,si ho és, resoldre'l. Ara estudiarem les matrius elementals, que ens permeten interpretar les operacions elementals com a productes de matrius.
Amb aquests instruments podem descriure de manera precisa l'algorisme de Gauss-Jordan.

5.1. Operacions elementals. La matriu identitat i les matrius elementals

Una *operació elemental* és una transformació matricial d'un qualsevol d'aquests tres tipus:

Operació de reducció: Sumar a una fila un múltiple adequat d'una altra fila. Més ben dit, substitució d'una fila pel resultat de sumar-hi un múltiple d'una altra fila.

Escalat d'una fila: Multiplicació d'una fila per un nombre no nul.

Permutació de dues files: Intercanviar l'ordre entre dues files.

☞ Només hi ha aquests tres tipus d'operacions elementals! Molt sovint es pot esglaonar un sistema fent servir altres transformacions, semblants a aquestes, però que no són estrictament operacions elementals; nosaltres utilitzarem només operacions elementals.

☞ En particular, cal remarcar que les operacions més importants —les de reducció— consisteixen en canviar una fila de la matriu pel resultat de sumar a aquesta fila un múltiple d'una altra: per exemple, podem canviar la primera fila, f_1, per $f_1 - (1/2)f_2$ (la primera menys la meitat de la segona), però no pel doble de la primera menys la segona: *canviar f_1 per $2f_1 - f_2$ no és una operació elemental.*

Definicions 5.1.

1. *Dues matrius* A *i* B *són equivalents per files si podem transformar* A *en* B *mitjançant un nombre finit d'operacions elementals.*
2. *Si la matriu esglaonada* S *es equivalent a* A *direm que* S *és una forma esglaonada de* A.
3. *Si la matriu esglaonada reduïda* R *és equivalent a* A *direm que* R *és la forma esglaonada reduïda de* A.

☞ Una matriu té moltes formes esglaonades, però la forma esglaonada reduïda és única, és a dir, que *per a qualsevol matriu* A *hi ha una sola matriu esglaonada reduïda que és equivalent per files a* A.

El fet més important a prop de les operacions elementals és que aquestes operacions es poden interpretar com multiplicacions matricials: *cada operació elemental sobre la matriu* A *és el producte d'una matriu especial per la matriu* A. Aquest fet serà molt i molt important i tindrà moltes aplicacions, especialment en els dos primers temes però també en la resta del curs.

Aquestes matrius especials les anomenarem *matrius elementals* i el nostre objectiu immediat és el d'investigar com han de ser les matrius elementals. Uns pocs exemples ho posaran en clar.

EXEMPLE 5.1.

Trobeu una matriu

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$$

de manera que, en multiplicar-la per

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix}$$

permuta les dues primeres files de A. És a dir, que

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -10 & 6 \\ 0 & 0 & -2 & 0 \\ 2 & 4 & -5 & 6 \end{bmatrix} \begin{matrix} \text{(fila 2a de A)} \\ \text{(fila 1a de A)} \\ \end{matrix}$$

Per a construir-la, recordem que les files del producte són combinacions lineals de les files de la segona matriu:

- La primera fila d'aquest producte és

$$e_{11}(\text{fila 1 de A}) + e_{12}(\text{fila 2 de A}) + e_{13}(\text{fila 3 de A})$$

  així que com que volem que el resultat siga la segona fila de A convé posar

$$0(\text{fila 1 de A}) + 1(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 0 & 1 & 0 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -10 & 6 \\ & & & \\ & & & \end{bmatrix}$$

- La segona fila ha de ser

$$1(\text{fila 1 de A}) + 0(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -10 & 6 \\ 0 & 0 & -2 & 0 \\ & & & \end{bmatrix}$$

- Finalment, com que la tercera fila del producte volem que siga la mateixa que la tercera fila de A, hi posem

$$0(\text{fila 1 de A}) + 0(\text{fila 2 de A}) + 1(\text{fila 3 de A})$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 & -10 & 6 \\ 0 & 0 & -2 & 0 \\ 2 & 4 & -5 & 6 \end{bmatrix}$$

La matriu que busquem és $\mathsf{E}_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

La matriu $\mathsf{E}_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ és una *matriu elemental del tipus permutació.* Els subíndexs $1, 2$ fan referència a les files que es permuten en realitzar el producte $\mathsf{E}_{1,2}\mathsf{A}$.

Exemple 5.2.

Trobeu una matriu

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$$

de manera que, en multiplicar-la per

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix}$$

divideix entre 2 la segona fila de A. És a dir, que

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 2 & -5 & 3 \\ 2 & 4 & -5 & 6 \end{bmatrix} \ \tfrac{1}{2}(\text{fila 2a de A})$$

Perquè la primera fila del producte coincidisca amb la primera fila de A hem de posar-hi

$$1(\text{fila 1 de A}) + 0(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ & & & \\ & & & \end{bmatrix}$$

La segona fila ha de ser

$$0(\text{fila 1 de A}) + (1/2)(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 2 & -5 & 3 \\ & & & \end{bmatrix}$$

Finalment, com que la tercera fila del producte volem que siga la mateixa que la tercera fila de A, hi posem

$$0(\text{fila 1 de A}) + 0(\text{fila 2 de A}) + 1(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 2 & 4 & -10 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 2 & -5 & 3 \\ 2 & 4 & -5 & 6 \end{bmatrix}$$

La matriu que cercàvem és aquesta: $\mathsf{E}_2(1/2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

La matriu $\mathsf{E}_2(1/2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ és una *matriu elemental del tipus escalat.* El subíndex 2 i el nombre $1/2$ fan referència a la fila de la matriu A que canvia d'escala (la segona) i al factor d'escala (un mig).

EXEMPLE 5.3.

Trobeu una matriu

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$$

de manera que, en multiplicar-la per

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix}$$

resta 2 vegades la fila segona a la fila tercera, és a dir, que

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 0 & 2 & 5 & -6 \end{bmatrix} \text{ (fila 3a de A)} - 2\text{(fila 2a de A)}$$

Perquè la primera fila no canvie hem de posar-hi

$$1(\text{fila 1 de A}) + 0(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ & & & \\ & & & \end{bmatrix}$$

Per a la segona fila,

$$0(\text{fila 1 de A}) + 1(\text{fila 2 de A}) + 0(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ & & \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ & & & \end{bmatrix}$$

Finalment, la tercera fila ha de ser

$$0(\text{fila 1 de A}) - 2(\text{fila 2 de A}) + 1(\text{fila 3 de A})$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 2 & 4 & -5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 1 & 1 & -5 & 6 \\ 0 & 2 & 5 & -6 \end{bmatrix}$$

Així que la matriu que busquem és $\mathsf{E}_{3,2}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$.

La matriu $\mathsf{E}_{3,2}(-2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ és una *matriu elemental del tipus reducció.* Els subíndexs $3, 2$ i el nombre -2 (entre parèntesis) fan referència al fet que sumem a la fila 3 la fila 2 multiplicada per -2.
Observem que l'entrada $(3, 2)$ d'aquesta matriu és precisament el nombre -2.

5.1.1. La matriu identitat

La *matriu identitat d'ordre n* és la matriu quadrada $n \times n$

$$\mathrm{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

(els uns ocupen tota la diagonal principal de la matriu i la resta d'entrades són zeros)

Aquesta matriu és l'element neutre de la multiplicació de matrius, perquè sempre que es multiplica a l'esquerra o a la dreta per una matriu A el resultat és la mateixa matriu A.

5.1.2. Les matrius elementals

Les matrius elementals que hem anat introduint en els exemples anteriors es poden veure com a petites modificacions de la matriu identitat; així com la identitat no modifica les matrius quan s'hi multiplica, les matrius elementals hi introdueixen petites variacions (el que anomenem operacions elementals). Hi ha tres tipus de matrius elementals:

La *matriu elemental del tipus permutació* $\mathsf{E}_{i,j}$ és la matriu que resulta d'intercanviar les files i i j de la matriu identitat.

Per exemple, si treballem amb matrius 4×4,

$$\mathsf{E}_{2,3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathsf{E}_{4,1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

☞ Multiplicar la matriu $\mathsf{E}_{i,j}$ per una matriu A és equivalent a intercanviar les files i i j de A. Per exemple,

$$E_{1,2} \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix}$$

La *matriu elemental del tipus escalat* $\mathsf{E}_i(\alpha)$ és la matriu que resulta de substituir en la matriu identitat l'element que es troba en la posició (i, i) per α.

Per exemple, si treballem amb matrius 4×4,

$$\mathsf{E}_2(3/2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathsf{E}_4(-2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

☞ Multiplicar la matriu $\mathsf{E}_i(\alpha)$ per una matriu A és equivalent a multiplicar la fila i de A per α. Per exemple,

$$E_1(-3) \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -12 & -6 \\ 2 & 3 \end{bmatrix}$$

La *matriu elemental del tipus reducció* $\mathsf{E}_{i,j}(\alpha)$ és la matriu que resulta de canviar l'entrada (i, j) de la matriu identitat per α.

Per exemple, si treballem amb matrius 4×4,

$$\mathsf{E}_{2,3}(4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \mathsf{E}_{4,1}(-2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}$$

☞ Multiplicar la matriu $\mathsf{E}_{i,j}(\alpha)$ per una matriu A és equivalent a sumar a la fila i de A α vegades la fila j. Per exemple,

$$E_{1,2}(-1) \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$$

5.2. Els algorismes de Gauss i de Gauss-Jordan

L'algorisme de Gauss és la primera fase del de Gauss-Jordan. Aquest algorisme computa una forma esglaonada de qualsevol matriu. La idea bàsica és la mateixa que hem fet servir en els exemples de la unitat anterior, però en comptes d'eliminar cada incògnita de totes les equacions excepte una, ara només l'eliminarem en les equacions que es troben *per davall* de l'element principal. Per suposat, no treballarem amb equacions sinó amb matrius.

Direm que la matriu A és esglaonada fins a la columna r si la submatriu $\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_r\end{bmatrix}$ és esglaonada. També direm que les files de A que contenen els pivots d'aquesta submatriu esglaonada són pivotades.

Per exemple, la matriu

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

és esglaonada fins a la cinquena columna i les dues primeres files són pivotades (perquè contenen els pivots 3 i 1).

Algorisme de Gauss

Aquest algorisme calcula una forma esglaonada de la matriu A.

Partim de la matriu $\mathsf{S} = \mathsf{A}$ ***i la transformarem fins que siga esglaonada.***

Repetiu...

Elecció del pivot: Entre totes les files no pivotades de S elegiu la que tinga el primer element no nul més a l'esquerra (si n'hi ha més d'una, elegiu-ne una qualsevol). Aquesta és la *fila pivot* i el seu primer element no nul és el *pivot*.

Si la fila pivot no és la primera entre les no pivotades, feu una permutació de files perquè ho siga.

Esglaonament: Feu zero tots els elements per baix del pivot, sumant a cada fila per baix de la fila pivot un múltiple adequat d'aquesta fila.

... fins que la matriu S **siga esglaonada.**

En qualsevol moment del procés podeu fer una operació elemental del tipus escalat.

☞ Notem que totes les transformacions que fem en aplicar aquest algorisme són operacions elementals, de manera que la matriu S es pot obtenir multiplicant unes quantes matrius elementals per la matriu A.

Exemple 5.4.

Apliqueu l'algorisme de Gauss per a trobar una forma esglaonada de la matriu

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 7 \\ 2 & 4 & -10 & 12 \\ 2 & 4 & -12 & 19 \\ 2 & 4 & -5 & -5 \end{bmatrix}$$

Justifiqueu cada operació elemental com un producte de matrius.

Podem elegir com a fila pivot qualsevol excepte la primera; per exemple, la segona. Així que permutem les files primera i segona:

$$\mathsf{E}_{1,2}\mathsf{A} = \begin{bmatrix} 2 & 4 & -10 & 12 \\ 0 & 0 & -2 & 7 \\ 2 & 4 & -12 & 19 \\ 2 & 4 & -5 & -5 \end{bmatrix}$$

Ara fem zeros per davall de a_{11}; restem la primera fila a la tercera i a la quarta:

$$\mathsf{E}_{4,1}(-1)\mathsf{E}_{3,1}(-1)\mathsf{E}_{1,2}\mathsf{A} = \begin{bmatrix} 2 & 4 & -10 & 12 \\ 0 & 0 & -2 & 7 \\ 0 & 0 & -2 & 7 \\ 0 & 0 & 5 & -17 \end{bmatrix}$$

La matriu ja està esglaonada fins la segona columna. Ara la fila pivot pot ser la segona, la tercera o la quarta. Ens quedem amb la segona i tornem a fer zeros per davall del pivot.

$$\mathsf{E}_{4,2}(5/2)\mathsf{E}_{3,2}(-1)\mathsf{E}_{4,1}(-1)\mathsf{E}_{3,1}(-1)\mathsf{E}_{1,2}\mathsf{A} = \begin{bmatrix} 2 & 4 & -10 & 12 \\ 0 & 0 & -2 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}$$

Ara la matriu és esglaonada fins la tercera columna; com a fila pivot hem d'escollir la quarta, així que intercanviem les files tercera i quarta.

$$\mathsf{E}_{3,4}\mathsf{E}_{4,2}(5/2)\mathsf{E}_{3,2}(-1)\mathsf{E}_{4,1}(-1)\mathsf{E}_{3,1}(-1)\mathsf{E}_{1,2}\mathsf{A} = \begin{bmatrix} 2 & 4 & -10 & 12 \\ 0 & 0 & -2 & 7 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathsf{S}$$

Aquesta matriu S és una forma esglaonada de la matriu A. □

☞ Per agilitar l'escriptura, farem servir la notació

$$\mathsf{A} \xrightarrow{\mathsf{E}} \mathsf{B}$$

per simbolitzar que la matriu B s'obté aplicant a A l'operació corresponent a la matriu E, és a dir, que EA = B (com ja hem fet a l'exercici 5.3.).

D'aquesta manera, el darrer exemple el podem esquematitzar així:

$$\mathsf{A} \xrightarrow{\mathsf{E}_{1,2}} \xrightarrow{\mathsf{E}_{3,1}(-1)} \xrightarrow{\mathsf{E}_{4,1}(-1)} \xrightarrow{\mathsf{E}_{3,2}(-1)} \xrightarrow{\mathsf{E}_{4,2}(5/2)} \xrightarrow{\mathsf{E}_{3,4}} \mathsf{S}$$

La diferència entre els algorismes de Gauss i de Gauss-Jordan és que aquest últim calcula la forma esglaonada reduïda de la matriu, en comptes d'una forma simplement esglaonada. Com veurem de seguida, per discutir si un sistema lineal és compatible és suficient conéixer-ne una forma esglaonada, de manera que bastarà que hi apliquem l'algorisme de Gauss.

Ara bé, la forma esglaonada reduïda té uns avantatges evidents sobre una forma esglaonada qualsevol: en primer lloc, com que té més zeros, qualsevol càlcul que s'hi faça és més econòmic (en termes de nombre d'operacions). Per altra banda, si fem servir aquesta matriu per a resoldre un sistema d'equacions, la solució és immediata sense haver de fer cap càlcul addicional, perquè en cada equació només apareix una incògnita principal. Finalment, les formes esglaonades tenen moltes altres aplicacions, al marge de la resolució de sistemes, i per això, com més simples siguen, millor hi treballarem.

Algorisme de Gauss-Jordan

***Aquest algorisme calcula la forma esglaonada reduïda de la matriu* A.**

Partim de la matriu* R = A *i la transformarem fins que siga esglaonada reduïda

Pas 1 (esglaonament): Apliqueu l'algorisme de Gauss per a transformar R en una matriu esglaonada.

Pas 2 (reducció): Començant pel darrer pivot (el que hi ha més a la dreta), feu zeros per damunt de tots els pivots (de dreta a esquerra).

Pas 3: (normalització): Dividiu cada fila no nul·la pel seu pivot (per fer uns els elements principals).

Exemple 5.5.

Apliqueu l'algorisme de Gauss-Jordan per a trobar la forma esglaonada reduïda de la matriu ampliada del sistema lineal

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 2 \\ x_1 - x_2 + x_3 &= 1 \\ 2x_1 + x_2 \qquad &= 3 \\ 3x_2 - 2x_3 &= 1 \end{aligned}$$

Hem de trobar la forma esglaonada reduïda de la matriu

$$\left[\mathsf{A} \mid \vec{b}\right] = \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 1 & -1 & 1 & 1 \\ 2 & 1 & 0 & 3 \\ 0 & 3 & -2 & 1 \end{array}\right]$$

Primer pas: algorisme de Gauss

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 1 & -1 & 1 & 1 \\ 2 & 1 & 0 & 3 \\ 0 & 3 & -2 & 1 \end{array}\right] \xrightarrow{\mathsf{E}_{3,1}(-2)\mathsf{E}_{2,1}(-1)} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 2 & -1 \\ 0 & -3 & 2 & 1 \\ 0 & 3 & -2 & 1 \end{array}\right] \xrightarrow{\mathsf{E}_{3,2}(-1)\mathsf{E}_{4,2}(1)} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Aquesta matriu és esglaonada.

Segon pas: reducció Només hi ha dos pivots. De fet, l'únic que hem de fer en aquest pas és anul·lar un element de la matriu.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\mathsf{E}_{1,2}(2/3)} \left[\begin{array}{ccc|c} 1 & 0 & 1/3 & 4/3 \\ 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Tercer pas: normalització Finalment, hem de dividir la segona fila entre -3.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1/3 & 4/3 \\ 0 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\mathsf{E}_{2}(-1/3)} \left[\begin{array}{ccc|c} 1 & 0 & 1/3 & 4/3 \\ 0 & 1 & -2/3 & 1/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = \left[\mathsf{R} \mid \vec{c}\right]$$

Aquesta darrera matriu és la forma esglaonada reduïda que cercàvem. A més a més R és la forma esglaonada reduïda de la matriu A.

5.3. Discussió i resolució dels sistemes lineals

Suposem que volem estudiar la compatibilitat d'un sistema lineal qualsevol i, en cas que siga compatible, resoldre'l. Podem treballar d'aquesta manera:

- Primer de tot apliquem l'algorisme de Gauss per a obtenir una forma esglaonada de la matriu ampliada del sistema.
- Aleshores anomenarem *incògnites principals* (o *variables principals*) a aquelles incògnites que en alguna equació apareixen multiplicades per un element principal.

  Llavors,

  - Si la columna dels termes independents té algun pivot, aleshores el sistema és incompatible, perquè la fila corresponent a aquest pivot representa l'equació

    $$0x_1 + 0x_2 + \cdots + 0x_n = 1$$

  - En cas contrari el sistema és compatible i la solució es pot obtenir completant l'algorisme de Gauss-Jordan, aïllant en cada equació el seu element principal, i substituint les variables no principals per paràmetres. Per tant,

- Si totes les variables són principals, el sistema és compatible determinat.
- Si hi ha alguna variable no principal, el sistema és compatible indeterminat, i té un conjunt infinit de solucions que es pot expressar amb tants paràmetres com variables no principals té.

Podem formalitzar aquest raonament introduint el concepte de rang d'una matriu.

DEFINICIÓ 5.2.

El rang de la matriu A, $\operatorname{rang}\mathsf{A}$, *és el nombre d'uns principals de la seua forma esglaonada reduïda (o, equivalentment, el nombre de files no nul·les de qualsevol de les seues formes esglaonades).*[a]

[a] O, també, el nombre de pivots o de columnes pivot.

Amb aquesta definició, el teorema següent és immediat.

TEOREMA 5.1. (TEOREMA DE ROUCHÉ)

Considerem el sistema lineal amb n *incògnites* $\mathsf{A}\vec{x} = \vec{b}$.

- *El sistema és compatible si i només si* $\operatorname{rang}\mathsf{A} = \operatorname{rang}\left[\mathsf{A} \mid \vec{b}\right]$
- *Si el sistema és compatible, aleshores*
  - *El nombre d'incògnites principals és* $\operatorname{rang}\mathsf{A}$.
  - *El nombre de paràmetres necessaris per expressar la solució general és* $n - \operatorname{rang}\mathsf{A}$.
  - *El sistema és determinat si i només si* $\operatorname{rang}\mathsf{A} = n$.

Demostració: La prova es basa en el raonament anterior: el sistema és compatible si i només si en la darrera columna de la forma esglaonada reduïda de la matriu ampliada no hi ha cap pivot, és a dir, si i només si el rang de la matriu de coeficients i el de la matriu ampliada coincideixen. D'altra banda, el nombre de variables no principals és clarament $n - \operatorname{rang}\mathsf{A}$. ☐

Com que el rang d'una matriu es pot determinar si se'n coneix una forma esglaonada (encara que no siga reduïda), aquest teorema ens permet decidir el caràcter de qualsevol sistema lineal sense més que trobar-ne una forma esglaonada y el procés per discutir i resoldre el sistema serà aquest:

Discussió i resolució d'un sistema lineal amb n incògnites

- Apliqueu l'algorisme de Gauss a la matriu ampliada $\left[\mathsf{A} \quad \vec{b}\right]$ per obtenir-ne una forma esglaonada
- Si $\operatorname{rang}\mathsf{A} \neq \operatorname{rang}\left[\mathsf{A} \quad \vec{b}\right]$, llavors el sistema és incompatible.

  En cas contrari,

  Si $\operatorname{rang}\mathsf{A} = n$, llavors el sistema és determinat; si no, indeterminat.

  Per trobar les solucions,

  - Completeu l'algorisme de Gauss-Jordan per obtenir la forma esglaonada reduïda de la matriu ampliada
  - En el nou sistema lineal, aïlleu en cada equació la incògnita principal
  - Substituïu les incògnites no principals per paràmetres.

EXEMPLE 5.6.

Discussió i càlcul de les solucions del sistema $\left\{\begin{aligned} &&& - 2x_3 &&& + 7x_5 &= 12\\ 2x_1 &+ 4x_2 &&- 10x_3 &+ 6x_4 &&+ 12x_5 &= 28\\ 2x_1 &+ 4x_2 &&- 5x_3 &+ 6x_4 &&- 5x_5 &= -1\end{aligned}\right.$

En forma matricial, el sistema és $A\vec{x} = \vec{b}$ amb

$$A = \begin{bmatrix} 0 & 0 & -2 & 0 & 7 \\ 2 & 4 & -10 & 6 & 12 \\ 2 & 4 & -5 & 6 & -5 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 12 \\ 28 \\ -1 \end{bmatrix}$$

de manera que la matriu ampliada del sistema és

$$\left[\begin{array}{ccccc|c} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{array}\right]$$

Aplicant l'algorisme de Gauss obtenim una forma esglaonada d'aquesta matriu:

$$\left[\begin{array}{ccccc|c} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 0 & 0 & 1/2 & 1 \end{array}\right]$$

D'ací deduïm que els rangs de la matriu de coeficients i de la matriu ampliada són iguals a 3, així que el sistema és indeterminat (amb dos paràmetres).

Ara continuem l'algorisme de Gauss-Jordan i trobem que la forma esglaonada reduïda de la matriu ampliada és

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right]$$

En conseqüència, el sistema lineal $A\vec{x} = \vec{b}$ és equivalent a aquest altre:

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 7 \\ 1 \\ 2 \end{bmatrix} \qquad \boxed{\begin{array}{rcl} x_1 + 2x_2 \quad + 3x_4 & = & 7 \\ x_3 & = & 1 \\ x_5 & = & 2 \end{array}}$$

Així que, aïllant els elements principals, obtindrem

$$\begin{aligned} x_1 &= 7 - 2x_2 - 3x_4 \\ x_3 &= 1 \\ x_5 &= 2 \end{aligned}$$

i la solució general s'obté canviant les dues variables x_2 i x_4 per paràmetres, $x_2 = \alpha_1$, $x_4 = \alpha_2$,

$$\begin{aligned} x_1 &= 7 - 2\alpha_1 - 3\alpha_2 \\ x_2 &= \alpha_1 \\ x_3 &= 1 \\ x_4 &= \alpha_2 \\ x_5 &= 2 \end{aligned} \qquad \boxed{\vec{x} = \begin{bmatrix} 7 \\ 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} + \alpha_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}$$

5.3.1. L'algorisme de substitució regressiva

Si hem esglaonat la matriu ampliada d'un sistema lineal i aquest sistema és compatible, en comptes d'aplicar l'algorisme de Gauss-Jordan per trobar-ne la forma esglaonada reduïda, podem trobar la solució fent servir el mètode de substitució regressiva.

Algorisme de substitució regressiva

Aquest algorisme resol el sistema lineal (compatible) $Sx = \vec{b}$ ***quan*** S ***és una matriu esglaonada.***

Pas 1: Aïlleu la incògnita principal en cada equació.

Pas 2: Substituïu cada equació en totes les anteriors.

Pas 3: Substituïu les incògnites no principals per paràmetres.

☞ Les operacions que cal fer per a resoldre un sistema lineal mitjançant l'algorisme de Gauss-Jordan són exactament les mateixes que les que requereix l'ús combinat de l'algorisme de Gauss i el de substitució regressiva. Però en l'algorisme de Gauss-Jordan aquestes operacions es fan sempre a nivell matricial, mentre que la substitució regressiva funciona a nivell de les equacions; en aquest sentit, l'algorisme de Gauss-Jordan és preferible a la substitució.[4]

Exemple 5.7.

Resoleu, fent servir l'algorisme de substitució regressiva, del sistema lineal

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

Aquest sistema ja és esglaonat, i les incògnites principals són x_1, x_2 i x_4.

Pas 1: Aïllem la incògnita principal de cada equació:

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

Pas 2: Substituïm cada equació en totes les anteriors:

$$\left.\begin{aligned} x_1 &= 1 + x_2 - x_3 - 3x_4 \\ x_2 &= 2 + 2x_3 + x_4 \\ x_4 &= 2 \end{aligned}\right\} \xrightarrow{x_4=2} \left.\begin{aligned} x_1 &= 1 + x_2 - x_3 - 3\cdot 2 \\ x_2 &= 2 + 2x_3 + 2 \\ x_4 &= 2 \end{aligned}\right\} \left.\begin{aligned} x_1 &= -5 + x_2 - x_3 \\ x_2 &= 4 + 2x_3 \\ x_4 &= 2 \end{aligned}\right\}$$

$$\xrightarrow{x_2=4+2x_3} \left.\begin{aligned} x_1 &= -5 + (4 + 2x_3) - x_3 \\ x_2 &= 4 + 2x_3 \\ x_4 &= 2 \end{aligned}\right\} \left.\begin{aligned} x_1 &= -1 + x_3 \\ x_2 &= 4 + 2x_3 \\ x_4 &= 2 \end{aligned}\right\}$$

Pas 3: Substituïm l'incògnita x_3 per un paràmetre:

$$\begin{aligned} x_1 &= -1 + \alpha \\ x_2 &= 4 + 2\alpha \\ x_3 &= \alpha \\ x_4 &= 2 \end{aligned}$$

5.4. Exercicis

Exercici 5.1. Determineu quines són les matrius elementals E_1, E_2, E_3 que, en multiplicar-les per $\mathsf{A} = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$ ens proporcionen els resultats següents:

(a) $\mathsf{E}_1\mathsf{A} = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$ (b) $\mathsf{E}_2\mathsf{A} = \begin{bmatrix} 1 & -1 \\ 4 & 0 \end{bmatrix}$ (c) $\mathsf{E}_1\mathsf{A} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$

Exercici 5.2. En cada apartat expliqueu quin és l'efecte de la multiplicació indicada.

1. $\mathsf{E}_4(-2)\mathsf{A}$
2. $\mathsf{E}_{2,4}(3)\mathsf{A}$
3. $\mathsf{E}_{4,1}\mathsf{A}$
4. $\mathsf{E}_{2,1}(3)\mathsf{E}_{3,1}(1)\mathsf{E}_{4,1}(-1/2)\mathsf{A}$

[4] A més a més, en la pràctica, quan resolem un sistema lineal *a mà* és més probable que cometem algun error si utilitzem l'estratègia de substitució regressiva.

Exercici 5.3. Calculeu la forma esglaonada reduïda de la matriu

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 7 \\ 2 & 4 & -10 & 12 \\ 2 & 4 & -5 & -5 \end{bmatrix}$$

seguint els passos indicats: en cada pas feu les operacions elementals corresponents a les matrius que hi ha al damunt de la fletxa.

$$\mathsf{A} = \begin{bmatrix} 0 & 0 & -2 & 7 \\ 2 & 4 & -10 & 12 \\ 2 & 4 & -5 & -5 \end{bmatrix} \xrightarrow{\mathsf{E}_{1,2}} \xrightarrow{\mathsf{E}_{3,1}(-1)} \xrightarrow{\mathsf{E}_1(1/2)} \xrightarrow{\mathsf{E}_{3,2}(5/2)\mathsf{E}_{1,2}(-5/2)} \xrightarrow{\mathsf{E}_2(-1/2)} \xrightarrow{\mathsf{E}_{2,3}(7)\mathsf{E}_{1,3}(23)} \xrightarrow{\mathsf{E}_3(2)} \mathsf{R}$$

Exercici 5.4. Calculeu una forma esglaonada de la matriu

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 6 & 1 & 6 \\ 1 & 2 & 2 & 7 \end{bmatrix}$$

fent servir l'algorisme de Gauss.

Exercici 5.5. Trobeu la forma esglaonada reduïda de la matriu $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 6 & 1 & 6 \\ 1 & 2 & 2 & 7 \end{bmatrix}$. Discutiu i resoleu el sistema lineal

$$\begin{aligned} x_1 + 2x_2 \phantom{{}+2x_3} &= 1 \\ 3x_1 + 6x_2 + x_3 &= 6 \\ x_1 + 2x_2 + 2x_3 &= 7 \end{aligned}$$

Exercici 5.6. Discutiu i resoleu pel mètode de Gauss-Jordan els sistemes lineals

(a) $\begin{cases} 2x_1 + x_2 + x_3 = 4 \\ x_1 - x_2 + x_3 = 1 \\ 3x_1 - x_2 - x_3 = 1 \\ 6x_1 - x_2 + x_3 = 6 \end{cases}$ (b) $\begin{cases} x_1 - x_2 + x_3 = 0 \\ -2x_1 + 5x_2 + 2x_3 = 0 \\ -7x_1 + 7x_2 + x_3 = 0 \end{cases}$ (c) $\begin{cases} x_1 - x_2 + 2x_3 - x_4 = -1 \\ 2x_1 + x_2 - 2x_3 - 2x_4 = -2 \\ -x_1 + 2x_2 - 4x_3 + x_4 = 1 \\ 3x_1 - 3x_4 = -4 \end{cases}$

Exercici 5.7. Discutiu segons els valors dels paràmetres (i resoleu, si és possible) els següents sistemes lineals:

(a) $\begin{cases} 2x_1 + x_2 = a \\ 6x_1 + 3x_2 = b \end{cases}$ (b) $\begin{cases} x_1 + x_2 + x_3 = a \\ 2x_1 + x_2 + 2x_3 = b \\ 3x_2 + x_3 = c \end{cases}$ (c) $\begin{cases} mx_1 + x_2 + 2x_3 + x_4 = m \\ 2mx_1 + 3x_2 - x_3 - x_4 = 1 \\ 3mx_1 + 4x_2 + x_3 = m + 1 \end{cases}$

Exercici 5.8. (Una matriu inversa) Si $\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ i $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$,

1. Resoleu el sistema lineal
$$\mathsf{A}\vec{x} = \vec{u}$$
2. Trobeu la matriu quadrada B per a la qual la solució és $\vec{x} = \mathsf{B}\vec{u}$ i calculeu AB i BA.

Exercici 5.9. (a) Feu servir el teorema de Rouché per a justificar que un sistema lineal amb més incògnites que equacions no pot ser determinat. (b) Un sistema amb més equacions que incògnites pot ser determinat? Indeterminat? Incompatible?

6. L'equació matricial $AX = B$ i els sistemes homogenis

Acabem el primer capítol observant que l'algorisme de Gauss-Jordan es pot fer servir per a resoldre equacions lineals matricials o diversos sistemes lineals simultàniament.
A més a més estudiem els sistemes lineals homogenis ($A\vec{x} = \vec{0}$), la qual cosa ens porta a definir l'espai nul de la matriu A, i observem la relació que hi ha entre les solucions d'un sistema lineal i el sistema homogeni corresponent.

6.1. La resolució d'un sistema lineal revisitada

Si ja hem reduït el sistema $A\vec{x} = \vec{b}$ a $R\vec{x} = \vec{c}$, on R és la forma esglaonada reduïda de A, i si aquest sistema és compatible, a la lliçó anterior hem vist que, per resoldre'l, hem d'aïllar les incògnites principals i substituir les no principals per paràmetres. El que ens interessa ara és observar que, això d'aïllar les incògnites principals, encara ho podem fer a nivell matricial.

Considerem, per exemple, el sistema (esglaonat reduït)

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{R} \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \tag{6.1}$$

Per tal d'aïllar les incògnites principals, escriurem la matriu R com la suma d'una *part principal*, i una altra no principal,

$$R = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(la part principal conté les columnes principals). Aleshores, el sistema (6.1) es pot escriure com

$$\left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right) \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \vec{x}$$

Aquesta darrera igualtat conté, al costat esquerre, les incògnites principals i, al dret, les no principals:

$$\begin{aligned} x_1 &= 1 - 2x_2 - x_4 - x_6 \\ x_3 &= -1 + x_4 - x_6 \\ x_5 &= 2 - x_6 \end{aligned}$$

i la solució s'obté canviant per paràmetres les incògnites x_2, x_4 i x_6:

$$\begin{aligned} x_2 &= \alpha_1 \\ x_4 &= \alpha_2 \\ x_6 &= \alpha_3 \end{aligned} \Longrightarrow \begin{aligned} x_1 &= 1 - 2\alpha_1 - \alpha_2 - \alpha_3 \\ x_2 &= \alpha_1 \\ x_3 &= -1 + \alpha_2 - \alpha_3 \\ x_4 &= \alpha_2 \\ x_5 &= 2 - \alpha_3 \\ x_6 &= \alpha_3 \end{aligned}$$

Aquesta idea, aïllar les incògnites sobre la matriu R, serà útil en la propera secció.

6.2. L'equació matricial $\mathsf{AX} = \mathsf{B}$

L'algorisme de Gauss-Jordan podem fer-lo servir per a resoldre l'equació matricial $\mathsf{AX} = \mathsf{B}$ on B és una matriu (no únicament un vector). En aquest cas, la incògnita X també serà una matriu. A serà una matriu $m \times n$, X una $n \times p$ i, B, una matriu $m \times p$.

Per resoldre l'equació, construirem la matriu ampliada $\left[\mathsf{A} \mid \mathsf{B}\right]$, i hi trobarem la forma esglaonada reduïda $\left[\mathsf{R} \mid \mathsf{C}\right]$. Això ens permet calcular els rangs de A i $\left[\mathsf{A} \mid \mathsf{B}\right]$. Llavors, si $\operatorname{rang}\left[\mathsf{A} \mid \mathsf{B}\right] = \operatorname{rang}\mathsf{A}$ l'equació serà compatible, i si, a més, aquest rang coincideix amb el nombre de columnes de A, llavors hi haurà una sola solució.

Tot això ho podem resumir en una versió matricial del teorema de Rouché:

Teorema 6.1. (Teorema de Rouché per a equacions matricials)

Siguen A i B dues matrius $m \times n$ i $m \times p$, respectivament.

1. *L'equació matricial $\mathsf{AX} = \mathsf{B}$ és compatible si i només si $\operatorname{rang}\left[\mathsf{A} \mid \mathsf{B}\right] = \operatorname{rang}\mathsf{A}$.*
2. *L'equació matricial $\mathsf{AX} = \mathsf{B}$ és determinada si i només si $\operatorname{rang}\left[\mathsf{A} \mid \mathsf{B}\right] = \operatorname{rang}\mathsf{A} = n$.*

Exemple 6.1.

Resoleu l'equació

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \mathsf{X} = \begin{bmatrix} 1 & -3 & -1 \\ 2 & -6 & -2 \end{bmatrix}$$

Primer de tot, apliquem l'algorisme de Gauss-Jordan:

$$\left[\begin{array}{cc|ccc} 1 & 2 & 1 & -3 & -1 \\ 2 & 4 & 2 & -6 & -2 \end{array}\right] \xrightarrow{\mathsf{E}_{2,1}(-2)} \left[\begin{array}{cc|ccc} 1 & 2 & 1 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Ara ja podem assegurar que l'equació és compatible i indeterminada. A més a més, aquesta equació és equivalent a

$$\begin{aligned}
\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \mathsf{X} &= \begin{bmatrix} 1 & -3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \right) \mathsf{X} &= \begin{bmatrix} 1 & -3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\
\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathsf{X}_1 \\ \mathsf{X}_2 \end{bmatrix} &= \begin{bmatrix} 1 & -3 & -1 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathsf{X}_1 \\ \mathsf{X}_2 \end{bmatrix} \\
\mathsf{X}_1 &= \begin{bmatrix} 1 & -3 & -1 \end{bmatrix} - 2\mathsf{X}_2 \\
\mathsf{O} &= \mathsf{O}
\end{aligned}$$

Finalment, la solució s'obté assignant valors paramètrics a les incògnites no principals X_2. Cal que tinguem en compte que X és una matriu 2×3, així que les seues files contenen tres entrades. Per tant, posarem $\mathsf{X}_2 = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$ i tindrem

$$\left. \begin{aligned} \mathsf{X}_1 &= \begin{bmatrix} 1 & -3 & -1 \end{bmatrix} - 2\mathsf{X}_2 \\ \mathsf{X}_2 &= \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix} \end{aligned} \right] \Longrightarrow \mathsf{X} = \begin{bmatrix} 1 & -3 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -2\alpha_1 & -2\alpha_2 & -2\alpha_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \end{bmatrix}$$

L'aplicació més interessant de la resolució d'una equació matricial és el càlcul de la matriu inversa. D'això ens n'ocuparem al capítol 2.

6.2.1. Resolució simultània de sistemes lineals

Suposem que hem de resoldre diversos sistemes lineals que comparteixen la matriu de coeficients (és a dir, que només difereixen en els termes independents),

$$\mathsf{A}\vec{x} = \vec{b}_1 \qquad \mathsf{A}\vec{x} = \vec{b}_2 \qquad \ldots \qquad \mathsf{A}\vec{x} = \vec{b}_p$$

El conjunt de tots aquests sistemes és equivalent a l'equació matricial

$$\mathsf{AX} = \mathsf{B}$$

on $\mathsf{B} = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix}$, així que, l'única cosa que hem de fer és construir la matriu (super)ampliada $\left[\mathsf{A} \mid \mathsf{B}\right]$ i aplicar-hi l'algorisme de Gauss-Jordan.

Exemple 6.2.

Discussió i resolució simultània dels sistemes lineals.

$$\begin{aligned} x_1 + 2x_2 \qquad &= 0 \\ 2x_1 + 5x_2 - x_3 &= 0 \\ -x_1 - 2x_2 + x_3 &= 0 \end{aligned} \qquad \begin{aligned} x_1 + 2x_2 \qquad &= -1 \\ 2x_1 + 5x_2 - x_3 &= -5 \\ -x_1 - 2x_2 + x_3 &= 3 \end{aligned} \qquad \begin{aligned} x_1 + 2x_2 \qquad &= 4 \\ 2x_1 + 5x_2 - x_3 &= 9 \\ -x_1 - 2x_2 + x_3 &= -4 \end{aligned}$$

Construïm la matriu ampliada

$$\left[\ \mathsf{A} \mid \vec{b}_1 \quad \vec{b}_2 \quad \vec{b}_3\ \right] = \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 0 & -1 & 4 \\ 2 & 5 & -1 & 0 & -5 & 9 \\ -1 & -2 & 1 & 0 & 3 & -4 \end{array}\right]$$

Aplicant-hi l'algorisme de Gauss-Jordan, obtenim la forma esglaonada reduïda

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array}\right]$$

Els tres sistemes són determinats. I Les solucions són, respectivament, aquestes:

$$\vec{x} = (0,0,0) \qquad \vec{x} = (1,-1,2) \qquad \vec{x} = (2,1,0)$$

Exemple 6.3.

Trobeu les solucions generals dels sistemes

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 + x_2 + x_3 &= 3 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - x_3 &= 2 \\ x_1 + x_2 + x_3 &= 2 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

Aplicant l'algorisme de Gauss-Jordan a la matriu ampliada

$$\left[\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 2 & 0 \\ 1 & 1 & 1 & 3 & 2 & 0 \end{array}\right]$$

obtenim la forma esglaonada reduïda

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & 2 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

Resolem l'equació matricial

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathsf{X} &= \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathsf{X} &= \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathsf{X} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathsf{X}_1 \\ \mathsf{X}_2 \\ \mathsf{X}_3 \end{bmatrix} &= \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathsf{X}_1 \\ \mathsf{X}_2 \\ \mathsf{X}_3 \end{bmatrix} \\ \begin{bmatrix} \mathsf{X}_1 \\ \mathsf{X}_3 \end{bmatrix} &= \begin{bmatrix} 2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \mathsf{X}_2 \\ \mathsf{O} \end{bmatrix} \end{aligned}$$

És a dir,

$$\begin{aligned} x_{11} &= 2 - x_{21} & x_{12} &= 2 - x_{22} & x_{13} &= 0 - x_{23} \\ x_{31} &= 1 & x_{32} &= 0 & x_{33} &= 0 \end{aligned}$$

o bé

$$\begin{aligned} x_{11} &= 2-\alpha & x_{12} &= 2-\alpha & x_{13} &= 0-\alpha \\ x_{21} &= \alpha & x_{22} &= \alpha & x_{23} &= \alpha \\ x_{31} &= 1 & x_{32} &= 0 & x_{33} &= 0 \end{aligned}$$

així que les solucions són, respectivament,

$$\vec{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

☞ i es veu clarament que la part *paramètrica* de les tres solucions és exactament la mateixa.

Això passa sempre amb els sistemes que comparteixen la matriu de coeficients, com veurem a l'apartat següent.

6.3. Sistemes homogenis

Un sistema lineal és *homogeni* quan tots els termes independents són zeros. Per exemple, el sistema

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

és homogeni.

Els sistemes homogenis sempre són compatibles, atés que el vector $\vec{x} = \vec{0}$ és evidentment una solució ($\mathsf{A}\vec{0} = \vec{0}$). De fet, el que diu el teorema de Rouché, aplicat a sistemes homogenis és açò:

Corol·lari 6.2. (Teorema de Rouché per a sistemes homogenis)

Qualsevol sistema lineal homogeni $\mathsf{A}\vec{x} = \vec{0}$ *és compatible.*
A més a més, aquest sistema és determinat si i només si el rang de A *coincideix amb el nombre de columnes d'aquesta matriu (ço és, amb el nombre d'incògnites del sistema).* □

El conjunt de les solucions del sistema homogeni $\mathsf{A}\vec{x} = \vec{0}$ l'anomenem *espai nul* de la matriu A.

Definició 6.1. (Espai nul d'una matriu)

L'espai nul de la matriu A, $\operatorname{Nul}\mathsf{A}$, *és el conjunt de totes les solucions del sistema lineal* $\mathsf{A}\vec{x} = \vec{0}$.

L'espai nul té aquestes propietats:

Propietats 6.3.

1. *L'espai* $\operatorname{Nul}\mathsf{A}$ *conté sempre el vector zero.*
2. *Si* $\vec{x}_1$ *i* $\vec{x}_2$ *són elements de* $\operatorname{Nul}\mathsf{A}$ *i si en fem una combinació lineal qualsevol,* $\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2$*, aleshores aquesta combinació lineal també és un element de* $\operatorname{Nul}\mathsf{A}$.

Demostració:

1. $\mathsf{A}\vec{0} = \vec{0}$.
2. Si $\vec{x}_1$ i $\vec{x}_2$ són elements de $\operatorname{Nul}\mathsf{A}$, llavors $\mathsf{A}\vec{x}_1 = \mathsf{A}\vec{x}_2 = \vec{0}$, així que,

$$\mathsf{A}(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) = \alpha_1\mathsf{A}\vec{x}_1 + \alpha_2\mathsf{A}\vec{x}_2 = \alpha_1\vec{0} + \alpha_2\vec{0} = \vec{0} \quad \square$$

Propietat 6.4.

Si el sistema lineal $A\vec{x} = \vec{b}$ *és compatible i* $\vec{x}_0$ *n'és una solució particular, llavors la solució general d'aquest sistema és*

$$\underbrace{\vec{x}}_{\substack{\text{sol. general} \\ \text{de } \mathsf{A}\vec{x}=\vec{b}}} = \underbrace{\vec{x}_0}_{\substack{\text{sol. particular} \\ \text{de } \mathsf{A}\vec{x}=\vec{b}}} + \underbrace{\operatorname{Nul}\mathsf{A}}_{\substack{\text{sol. general} \\ \text{de } \mathsf{A}\vec{x}=\vec{0}}}$$

Demostració: Si $\vec{x}$ és solució de $\mathsf{A}\vec{x} = \vec{b}$, llavors $\mathsf{A}(\vec{x}_0 - \vec{x}) = \mathsf{A}\vec{x}_0 - \mathsf{A}\vec{x} = \vec{b} - \vec{b} = \vec{0}$, així que $\vec{x}_0 - \vec{x} \in \operatorname{Nul}\mathsf{A}$ i, per tant, $\vec{x} = \vec{x}_0 + (\vec{x} - \vec{x}_0) \in \vec{x}_0 + \operatorname{Nul}\mathsf{A}$.

Recíprocament, si $\vec{x} \in \vec{x}_0 + \operatorname{Nul}\mathsf{A}$, és a dir, si hi ha un vector $\vec{x}_1 \in \operatorname{Nul}\mathsf{A}$ de manera que $\vec{x} = \vec{x}_0 + \vec{x}_1$ llavors, $\mathsf{A}\vec{x} = \mathsf{A}(\vec{x}_0 + \vec{x}_1) = \mathsf{A}\vec{x}_0 + \mathsf{A}\vec{x}_1 = \vec{b} - \vec{0} = \vec{b}$, així que $\vec{x}$ és una solució de $\mathsf{A}\vec{x} = \vec{b}$. □

☞ Això justifica l'observació que féiem a la fi de l'apartat anterior: les solucions dels tres sistemes

$$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ x_1 + x_2 + x_3 &= 3 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - x_3 &= 2 \\ x_1 + x_2 + x_3 &= 2 \end{aligned} \qquad \begin{aligned} x_1 + x_2 - x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

tenen la mateixa *part paramètrica*; l'espai nul de la matriu de coeficients és

$$\operatorname{Nul}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \left\langle \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

i les solucions són

$$\vec{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + \operatorname{Nul} \mathsf{A} \qquad \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \operatorname{Nul} \mathsf{A} \qquad \vec{x} = \operatorname{Nul} \mathsf{A}$$

6.4. Exercicis

6.4.1. Equacions matricials

Exercici 6.1. Resoleu l'equació

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix} \mathsf{X} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & -2 \end{bmatrix}$$

Exercici 6.2. (Una matriu inversa)

(a) Resoleu l'equació

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \mathsf{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b) Siguen $\vec{x}_1$ i $\vec{x}_2$ les dues columnes de la solució de l'apartat anterior. Proveu que el sistema lineal

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} a \\ b \end{bmatrix}$$

és compatible determinat i que la solució és $\vec{x} = a\vec{x}_1 + b\vec{x}_2$.

(c) Resoleu els sistemes lineals següents:

$$\text{(a)} \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \text{(b)} \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{(c)} \quad \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \vec{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

6.4.2. Sistemes homogenis i espais nuls

Exercici 6.3. Trobeu els espais nuls de les matrius següents

$$\text{(a)} \quad \mathsf{A} = \begin{bmatrix} 2 & 0 & -4 \\ 3 & -3 & -6 \\ 1 & 0 & -2 \end{bmatrix} \qquad \text{(b)} \quad \mathsf{B} = \begin{bmatrix} 5 & 0 & -4 \\ 3 & 0 & -6 \\ 1 & 0 & 1 \end{bmatrix} \qquad \text{(c)} \quad \mathsf{C} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

$$\text{(d)} \quad \mathsf{D} = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 3 & 7 \\ 0 & 0 & -1 \end{bmatrix} \qquad \text{(e)} \quad \mathsf{E} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{(f)} \quad \mathsf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 3 & 0 \end{bmatrix}$$

Exercici 6.4. (a) Resoleu el sistema homogeni

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

(b) Trobeu l'espai nul de la matriu

$$\begin{bmatrix} 1 & 2 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(c) Trobeu el conjunt de tots els vectors que són ortogonals als dos vectors del conjunt

$$A = \{(1, 1, -1, 0), (1, 2, 1, 1)\}$$

Exercici 6.5. Considerem la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ i el vector $\vec{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

(a) Determineu l'espai nul de A.

(b) Comproveu que $\vec{x}_p = (1, 0, 0, 1)$ és una solució particular del sistema lineal $\mathsf{A}\vec{x} = \vec{b}$.

(c) Sense fer cap més càlcul, determineu la solució general del sistema lineal $\mathsf{A}\vec{x} = \vec{b}$.

6.4.3. Cadenes de Màrkov

Exercici 6.6. (Matrius estocàstiques i vectors estacionaris) Un vector $\vec{u}$ és *estocàstic* si totes les seues coordenades són no negatives i sumen 1; per exemple, $\vec{u}_0 = (0{,}20, 0{,}45, 0{,}35)$ és estocàstic. Una matriu quadrada $n \times n$, A, és *estocàstica* si totes columnes són vectors estocàstics; per exemple, la matriu

$$\mathsf{M} = \begin{bmatrix} 0{,}85 & 0{,}15 & 0{,}12 \\ 0{,}10 & 0{,}75 & 0{,}15 \\ 0{,}05 & 0{,}10 & 0{,}73 \end{bmatrix}$$

és estocàstica.
(a) Proveu que el vector producte d'una matriu estocàstica per un vector estocàstic també és estocàstic. Un vector $\vec{u}$ és *estacionari* per a la matriu A si $\mathsf{A}\vec{u} = \vec{u}$. (b) Proveu que, si la matriu A és estocàstica, llavors hi ha vectors estacionaris distints del vector zero. (c) Trobeu tots els vectors estacionaris de la matriu M. (d) Trobeu un vector estacionari estocàstic de la matriu M.

La matriu M d'aquest exercici és la matriu de transició de la cadena de Màrkov que vam introduir a la lliçó 1, i el vector que obtingueu al darrer apartat representa la distribució de la població a llarg termini en aquell problema.

Capítol 2

Matrius

Continguts

7. El rang d'una matriu

Abans de tractar les matrius invertibles estudiarem la dependència lineal i trobarem la relació que hi ha entre la dependència lineal i el rang d'una matriu.
D'aquesta manera recuperem la definició clàssica del rang d'una matriu com el nombre de columnes (o files) linelament independents que té aquesta matriu.

7.1. Interpretació matricial dels algorismes d'esglaonament

L'algorisme de Gauss-Jordan (i, en general, tots els algorismes que consisteixen en l'aplicació successiva de diverses operacions elementals) són equivalents a la multiplicació de dues matrius: si la matriu B s'obté a partir de A premultiplicant-hi les matrius elementals $\mathsf{E}_1, \mathsf{E}_2, \ldots, \mathsf{E}_p$, és a dir, si

$$\mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1\mathsf{A} = \mathsf{B}$$

llavors, anomenant T al producte $\mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1$, és clar que $\mathsf{TA} = \mathsf{B}$. Aquesta matriu T podem calcular-la simultàniament amb B si fem sobre la matriu identitat les mateixes operacions que hem fet sobre A, perquè

$$\mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1\mathrm{I} = \mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1 = \mathsf{T}$$

així que si construïm la matriu ampliada $\left[\mathsf{A} \mid \mathrm{I}\right]$ i hi apliquem les operacions elementals que transformen A en B obtindrem B i T simultàniament:

$$\mathsf{E}_p \ldots \mathsf{E}_2\mathsf{E}_1\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1\mathsf{A} \mid \mathsf{E}_p \cdots \mathsf{E}_2\mathsf{E}_1\mathrm{I}\right] = \left[\mathsf{B} \mid \mathsf{T}\right]$$

Exemple 7.1.

Calculeu la forma esglaonada reduïda, R, de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \end{bmatrix}$ i la matriu T tal que $\mathsf{TA} = \mathsf{R}$.

El que hem de fer és aplicar l'algorisme de Gauss-Jordan a la matriu $\left[\mathsf{A} \mid \mathsf{I}\right]$:

$$\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|cc} 1 & 2 & -1 & 1 & 0 \\ 2 & 0 & 3 & 0 & 1 \end{array}\right]$$

$$\mathsf{E}_{2,1}(-2)\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|cc} 1 & 2 & -1 & 1 & 0 \\ 0 & -4 & 5 & -2 & 1 \end{array}\right]$$

$$\mathsf{E}_{1,2}(1/2)\mathsf{E}_{2,1}(-2)\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|cc} 1 & 0 & 3/2 & 0 & 1/2 \\ 0 & -4 & 5 & -2 & 1 \end{array}\right]$$

$$\mathsf{E}_2(-1/4)\mathsf{E}_{1,2}(1/2)\mathsf{E}_{2,1}(-2)\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|cc} 1 & 0 & 3/2 & 0 & 1/2 \\ 0 & 1 & -5/4 & 1/2 & -1/4 \end{array}\right] = \left[\mathsf{R} \mid \mathsf{T}\right]$$

així que

$$\underbrace{\begin{bmatrix} 0 & 1/2 \\ 1/2 & -1/4 \end{bmatrix}}_{\mathsf{T}=\mathsf{E}_2(-1/4)\mathsf{E}_{1,2}(1/2)\mathsf{E}_{2,1}(-2)} \underbrace{\begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} 1 & 0 & 3/2 \\ 0 & 1 & -5/4 \end{bmatrix}}_{\mathsf{R}=\mathsf{E}_2(-1/4)\mathsf{E}_{1,2}(1/2)\mathsf{E}_{2,1}(-2)\mathsf{A}} \qquad \square$$

Una conseqüència interessant del fet que la forma esglaonada reduïda R de la matriu A és el producte d'una altra matriu T per A és aquesta: *les files de* R *(o de qualsevol altra matriu que s'haja obtingut de* A *fent-hi operacions elementals) són combinacions lineals de les files de* A. A més a més, les files de T contenen els pesos d'aquestes combinacions lineals.[1]

En el cas de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 3 \end{bmatrix}$,

$$0\begin{bmatrix} 1 & 2 & -1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3/2 \end{bmatrix}$$

$$\frac{1}{2}\begin{bmatrix} 1 & 2 & -1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -5/4 \end{bmatrix}$$

[1]Recordem que les files d'un producte de dues matrius són combinacions lineals de les files de la segona matriu.

7.2. (IN)DEPENDÈNCIA LINEAL

La idea de dependència o independència lineal (d'un conjunt de vectors $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$) està relacionada amb l'equació

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_p\vec{u}_p = \vec{0}$$

Aquesta equació té una solució evident: $x_1 = 0, x_2 = 0, \dots, x_p = 0$, perquè

$$0\vec{u}_1 + 0\vec{u}_2 + \cdots + 0\vec{u}_p = \vec{0}$$

Ara bé, en alguns casos, aquesta és l'única solució i, en altres casos, n'hi ha d'altres. En el primer cas direm que els vectors $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p$ són linealment independents.

DEFINICIÓ 7.1.

El conjunt de vectors $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ és linealment independent (o lliure) si l'única solució de l'equació

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_p\vec{u}_p = \vec{0} \tag{7.1}$$

és $x_1 = x_2 = \cdots = x_p = 0$. En cas contrari, el conjunt és linealment dependent (o lligat).

Vegem-ne alguns exemples.

EXEMPLE 7.2.

Estudieu si el conjunt

$$\mathscr{S} = \{(1,2,0,1), (0,1,0,0), (-1,0,1,0)\}$$

és linealment dependent o independent.

L'equació vectorial

$$x_1\begin{bmatrix}1\\2\\0\\1\end{bmatrix} + x_2\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + x_3\begin{bmatrix}-1\\0\\1\\0\end{bmatrix} = \vec{0}$$

és equivalent al sistema lineal

$$\begin{bmatrix}1 & 0 & -1\\2 & 1 & 0\\0 & 0 & 1\\1 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

I, com que el rang de la matriu $\left[\begin{smallmatrix}1 & 0 & -1\\2 & 1 & 0\\0 & 0 & 1\\1 & 0 & 0\end{smallmatrix}\right]$ és 3, aquest sistema és determinat i té només la solució nul·la. En conseqüència, el conjunt $\mathscr{S}$ és linealment independent.

EXEMPLE 7.3.

Estudieu si el conjunt de vectors

$$\mathscr{S} = \{(1,0,-1), (2,1,0), (5,2,-1)\}$$

és linealment dependent o independent.

L'equació que cal investigar ara és

$$x_1(1,0,-1) + x_2(2,1,0) + x_3(5,2,-1) = (0,0,0) \tag{7.2}$$

O bé

$$\begin{bmatrix}1 & 2 & 5\\0 & 1 & 2\\-1 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

Esglaonem la matriu ampliada del darrer sistema.

$$\begin{bmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ -1 & 0 & -1 & 0 \end{bmatrix} \xrightarrow{\mathsf{E}_{3,1}(1)} \begin{bmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{bmatrix} \xrightarrow{\mathsf{E}_{3,2}(-2)} \begin{bmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Així que el rang de la matriu és 2 i el sistema lineal és indeterminat. En conseqüència, el conjunt $\mathscr{S}$ és linealment dependent. Per tal de confirmar-ho, acabem d'aplicar l'algorisme de Gauss-Jordan.

$$\begin{bmatrix} 1 & 2 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{E_{1,2}(-2)} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solució és $x_1 = -\alpha, x_2 = -2\alpha, x_3 = \alpha$, així que, elegint $\alpha = 1$,

$$-(1,0,-1) - 2(2,1,0) + (5,2,-1) = \vec{0}$$

la qual cosa ens proporciona una solució no nul·la de l'equació (7.2).

☞ Aquests exemples posen de manifest la relació fonamental entre el rang d'una matriu i la independència lineal: Si $\mathscr{S} = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ és un conjunt finit de vectors, representarem com $\mathsf{M}_{\mathscr{S}}$ la matriu que té aquests vectors com a columnes, $\mathsf{M}_{\mathscr{S}} = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_p \end{bmatrix}$. Llavors,

> $\mathscr{S}$ és linealment independent si i només si rang $\mathsf{M}_{\mathscr{S}} = p$ (el nombre de vectors de $\mathscr{S}$).

La resta d'aquesta lliçó la dedicarem a concretar millor aquest fet.

7.3. Relacions de dependència i combinacions lineals

Quan un conjunt de vectors $\mathscr{S}$ és linealment dependent, cada solució no nul·la de l'equació (7.1) ens proporciona una *relació de dependència entre els vectors del conjunt* $\mathscr{S}$. Així, al darrer exemple hem trobat que la forma esglaonada reduïda de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 2 \\ -1 & 0 & -1 \end{bmatrix}$ és $\mathsf{R} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ i que la solució general del sistema lineal $\mathsf{A}\vec{x} = \vec{0}$ és $x_1 = -\alpha, x_2 = -2\alpha, x_3 = \alpha$; en conseqüència, hi ha la *relació de dependència* següent:

$$-(1,0,-1) - 2(2,1,0) + (5,2,-1) = (0,0,0)$$

En aquesta relació podem aïllar qualsevol dels vectors per a escriure'l com a combinació lineal dels altres dos. Per exemple,

$$(5,2,-1) = (1,0,-1) + 2(2,1,0)$$

En general, sempre que un conjunt de vectors siga linealment dependent hi podrem escriure algun d'aquests vectors com a combinació lineal dels altres. Aquesta és una de les propietats més importants de la dependència lineal i, de fet, en justifica el nom: que un conjunt de vectors siga linealment dependent significa que algun vector *depén* dels altres.

Propietat 7.1.

> *El conjunt $\mathscr{S}$ és linealment dependent si i només si algun dels vectors d'aquest conjunt és combinació lineal dels altres.* ☐

Per acabar aquest apartat enunciarem algunes altres propietats de la dependència lineal.

Propietats 7.2.

Siga $\mathscr{S}$ un conjunt de vectors.

1. *Si el vector nul $\vec{0}$ és un element de $\mathscr{S}$, llavors $\mathscr{S}$ és linealment dependent.*
2. *El conjunt format per un sol vector no nul és linealment independent.*
3. *Si $\mathscr{T}$ és un subconjunt de $\mathscr{S}$ i $\mathscr{T}$ és linealment dependent, llavors $\mathscr{S}$ també és dependent. Aquesta propietat es pot formular, de manera equivalent, així*
4. *Si $\mathscr{T}$ és un subconjunt de $\mathscr{S}$ i $\mathscr{S}$ és linealment independent, llavors $\mathscr{T}$ també és independent.*

Totes aquestes propietats es poden provar fàcilment. □

7.4. El rang

Al capítol primer, hem definit el rang d'una matriu com el nombre d'uns principals de la seua forma esglaonada reduïda. Ara interpretarem aquest concepte en termes de la independència lineal entre les files o les columnes de la matriu.

7.4.1. Dependència lineal entre les columnes d'una matriu

Suposem que $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ són les columnes d'una matriu $\mathsf{A} \in \mathscr{M}_{m\times n}$. Per a estudiar si aquestes columnes són linealment independents haurem de discutir l'equació

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{0} \tag{7.3}$$

o, de manera equivalent, el sistema lineal

$$\mathsf{A}\vec{x} = \vec{0}$$

Per tant, com ja hem diu adés, *la condició necessària i suficient perquè les columnes de* A *siguen linealment independents és que* $\operatorname{rang}\mathsf{A} = n$

D'altra banda, com que les operacions elementals no canvien les solucions dels sistemes lineals, si la matriu B s'ha obtingut fent operacions elementals sobre A i si $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ són les columnes de B, l'equació (7.3) és equivalent a aquesta altra

$$x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_n\vec{b}_n = \vec{0}$$

☞ Això significa que les operacions elementals no canvien les relacions de dependència entre les columnes de les matrius.

Aquest fet ens permet trobar les relacions de dependència mitjançant l'algorisme de Gauss-Jordan, perquè en una matriu esglaonada reduïda aquestes relacions de dependència són immediates.

Exemple 7.4.

Trobeu les relacions de dependència entre les columnes de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 2 & 3 & 5 & 0 & 17 \\ -2 & 1 & -5 & -7 & -9 & 0 & -38 \\ 0 & 0 & 0 & 1 & 2 & 0 & 5 \\ -2 & 1 & -5 & -7 & -9 & 1 & -41 \\ -1 & 1 & -3 & -4 & -4 & 1 & -24 \end{bmatrix}$$

La forma esglaonada reduïda d'aquesta matriu és

$$\mathsf{R} = \begin{bmatrix} 1 & 0 & 2 & 0 & -1 & 0 & 2 \\ 0 & 1 & -1 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Com que la matriu R és esglaonada reduïda, podem identificar immediatament que les columnes principals són la 1, la 2, la 4 i la 6; a més a més, s'hi veuen clarament les relacions de dependència següents:

$$\begin{aligned}\vec{r}_3 &= 2\vec{r}_1 + (-1)\vec{r}_2\\ \vec{r}_5 &= (-1)\vec{r}_1 + 3\vec{r}_2 + 2\vec{r}_4\\ \vec{r}_7 &= 2\vec{r}_1 + 1\vec{r}_2 + 5\vec{r}_4 + (-3)\vec{r}_6\end{aligned}$$

☞ Les columnes no principals en una matriu esglaonada reduïda són combinacions lineals de les columnes principals *anteriors* a elles i, a més a més, els pesos d'aquestes combinacions lineals són, precisament, les entrades d'aquestes columnes.

Així doncs, com que les relacions de dependència entre les columnes de A són les mateixes que les que hi ha entre les columnes de R,

$$\vec{a}_3 = 2\vec{a}_1 + (-1)\vec{a}_2$$

(com que la tercera columna de R és $(2, -1, 0, 0)$, la tercera columna de A és 2 vegades la primera columna més -1 vegada la segona!) i, anàlogament,

$$\begin{aligned}\vec{a}_5 &= (-1)\vec{a}_1 + 3\vec{a}_2 + 2\vec{a}_4\\ \vec{a}_7 &= 2\vec{a}_1 + 1\vec{a}_2 + 5\vec{a}_4 + (-3)\vec{r}_6\end{aligned}$$

A més a més, resulta que el nombre màxim de columnes linealment independents de qualsevol matriu A és igual al nombre d'uns principals que té la seua forma esglaonada reduïda, és a dir,

El rang per columnes
El rang d'una matriu qualsevol coincideix amb el nombre màxim de columnes linealment independents que té aquesta matriu.

7.4.2. Dependència lineal entre les files d'una matriu

El problema de buscar relacions de dependència entre les files d'una matriu A ja l'hem resolt anteriorment; com que les files de qualsevol forma esglaonada S de A són combinacions lineals de les files de A, si el rang de A és menor que el nombre de files que té aquesta matriu, llavors les files nul·les de S ens proporcionen relacions de dependència entre les files de A.

Exemple 7.5.

Trobeu alguna relació de dependència entre les files de la matriu

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 2 & 7 & -3 \end{bmatrix}$$

Cercarem una forma esglaonada, S, de A i la matriu T tal que $TA = S$, fent operacions elementals sobre la matriu ampliada $\begin{bmatrix} A & I \end{bmatrix}$:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 2 & 7 & -3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{E_{3,1}(-2)} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 3 & -3 & -2 & 0 & 1 \end{array}\right] \xrightarrow{E_{3,2}(3)} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 & 3 & 1 \end{array}\right]$$

La matriu $S = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ és una forma esglaonada de A i podem obtenir-la premultiplicant A per la matriu $T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix}$.

Per tant, la tercera fila del producte

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 2 & 7 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

és

$$\begin{bmatrix} -2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 2 & 7 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

En definitiva,

$$-2 \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} + 3 \begin{bmatrix} 0 & -1 & 1 \end{bmatrix} + 1 \begin{bmatrix} 2 & 7 & -3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

És fàcil veure que si la forma esglaonada reduïda de A no té cap fila de zeros, aleshores no hi ha relacions de dependència entre les files de A, així que

El rang per files
El rang d'una matriu qualsevol coincideix amb el nombre de files linealment independents que té aquesta matriu.

En conclusió, el nombre de files independents i el nombre de columnes independents en una matriu qualsevol són iguals, i aquest nombre és, precisament, el rang de la matriu.

Teorema 7.3. (El rang d'una matriu)

En qualsevol matriu, el nombre de columnes que són linealment és el mateix que el de files independents. A més a més, aquest nombre coincideix amb el rang de la matriu. □

7.5. Exercicis

7.5.1. Interpretació matricial dels algorismes d'esglaonament

Exercici 7.1. Donada la matriu $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 4 \end{bmatrix}$, si R és la forma esglaonada reduïda de A calculeu una matriu T de manera que TA = R.

7.5.2. Independència lineal i relacions de dependència

Exercici 7.2. Estudieu si els conjunts següents són linealment dependents o independents.

(a) $A = \{(1,2,0),(1,-1,1),(1,5,-1)\}$

(b) $B = \{(1,2,0,1),(1,-1,1,1),(1,5,-1,0)\}$

Exercici 7.3. (a) Calculeu la forma esglaonada reduïda, R, de la matriu

$$A = \begin{bmatrix} 1 & 1 & 0 & 3 & 0 \\ -2 & -2 & 1 & -7 & 2 \\ 0 & 0 & 1 & -1 & 2 \end{bmatrix}$$

i trobeu la matriu T de manera que TA = R

(b) Quin és el rang de la matriu A?

(c) Trobeu alguna relació de dependència entre les columnes de la matriu A.

(d) Trobeu alguna relació de dependència entre les files de la matriu A.

Exercici 7.4. Proveu que qualsevol subconjunt de $\mathbb{R}^3$ amb quatre vectors és necessàriament linealment dependent.

Exercici 7.5. Determineu alguna relació de dependència entre les columnes de la matriu

$$A = \begin{bmatrix} -3 & 1 & -5 & 13 \\ 0 & 1 & 1 & 4 \\ 1 & 0 & 2 & -3 \end{bmatrix}$$

i calculeu el rang de A.

Exercici 7.6. Extraieu un subconjunt linealment independent del conjunt

$$A = \{(1,2,1,0),(1,1,0,1),(0,1,1,-1),(1,1,1,1)\}$$

que tinga el màxim nombre possible d'elements.

Exercici 7.7. Proveu que si el conjunt $A = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ és linealment independent llavors, $B = \{\vec{u}_1 + \vec{u}_2, \vec{u}_2 + \vec{u}_3, \vec{u}_1 + \vec{u}_3\}$ també és independent. Què podem dir quant a la independència lineal d'aquest altre conjunt: $C = \{\vec{u}_1 + \vec{u}_2, \vec{u}_2 + \vec{u}_3, \vec{u}_1 - \vec{u}_3\}$?

Exercici 7.8. Proveu que, si $\mathsf{A} \in \mathcal{M}_{m\times n}$ i $\mathsf{B} \in \mathcal{M}_{n\times p}$ llavors, $\operatorname{rang} \mathsf{AB} \leq \operatorname{rang} \mathsf{A}$ i $\operatorname{rang} \mathsf{AB} \leq \operatorname{rang} \mathsf{B}$.

8. La matriu inversa

Aquesta és la lliçó més important del curs. Determinar si una matriu és invertible i, en cas que ho siga, calcular-ne la inversa és una de les qüestions centrals de l'Àlgebra Lineal.
Al mateix temps, és la lliçó més simple! Amb tot el que sabem a prop de les matrius i els sistemes lineals, tant el problema de decidir si una matriu és invertible com el càlcul efectiu de la inversa són trivials.

8.1. Definició, propietats i càlcul d'inverses

Definició 8.1.

Direm que la matriu quadrada $\mathsf{A} \in \mathcal{M}_{n\times n}$ *és invertible si existeix una altra matriu* B *de manera que*

$$\mathsf{AB} = \mathrm{I} \tag{8.1}$$

Segons aquesta definició, que la matriu A siga invertible vol dir que l'equació $\mathsf{AX} = \mathrm{I}$ siga compatible. Vegem algun exemple de matrius que són invertibles i de matrius que no ho són.

Exemple 8.1.

Estudieu si la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ és invertible.

Es tracta de discutir i resoldre l'equació matricial

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \mathsf{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Així que l'únic que hem de fer és calcular la forma esglaonada reduïda de la matriu ampliada. Això ho podem fer simplement restant a la primera fila el doble de la segona: $\left[\mathsf{A} \mid \mathrm{I}\right]$:

$$\mathsf{E}_{1,2}(-2)\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 0 & 1 & 0 & 1 \end{array}\right]$$

Llavors, l'equació és determinada i la solució és

$$\mathsf{X} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

En conseqüència, la matriu A és invertible.

Exemple 8.2.

Estudieu si la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 3 & 0 \end{bmatrix}$ és invertible.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\mathsf{E}_{3,1}(-1)} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow{\mathsf{E}_{3,2}(-1)} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right]$$

Aquesta matriu és esglaonada, així que ja podem observar que el rang de A és 2, mentre que el rang de la matriu ampliada $\left[\mathsf{A} \mid \mathrm{I}\right]$ és 3, així que l'equació matricial és incompatible. En conseqüència, la matriu A no és invertible.

Els fets més importants a prop de les matrius invertibles són els següents:

PROPIETATS 8.1.

1. *Una matriu $n \times n$ A és invertible si i només si* $\operatorname{rang} \mathsf{A} = n$.
2. *Si la matriu A és invertible llavors l'equació $\mathsf{AX} = \mathrm{I}$ té una solució única.*

Demostració:

1. Com que el rang de la matriu I d'ordre n és n, és evident que $\left[\mathsf{A} \mid \mathrm{I}\right]$ també té rang n (té n columnes linealment independents). Llavors, el teorema de Rouché garanteix que el sistema $\mathsf{AX} = \mathrm{I}$ és compatible si i només si $\operatorname{rang} A = n$.
2. Això també és conseqüència del teorema de Rouché, perquè, per la propietat anterior si A és una matriu $n \times n$ invertible, aleshores el rang de A és n i, llavors, la solució del sistema lineal $\mathsf{AX} = \mathrm{I}$ és única. □

DEFINICIÓ 8.2.

Si la matriu A és invertible, llavors la matriu inversa de A és l'única matriu A^{-1} que verifica la igualtat

$$\mathsf{AA}^{-1} = \mathrm{I}$$

Observem que la propietat 8.1.1 podem reformular-la de moltes maneres interessants:

TEOREMA 8.2.

Siga A una matriu quadrada $n \times n$. Totes les afirmacions següents són equivalents:

1. *A és invertible*
2. *L'equació matricial $\mathsf{AX} = \mathrm{I}$ és compatible*
3. *El rang de A és n*
4. *Les files de A són linealment independents*
5. *Les columnes de A són linealment independents*
6. *La forma esglaonada reduïda de A és la matriu identitat*
7. *Per a qualsevol vector $\vec{b}$ el sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és compatible*
8. *Per a qualsevol vector $\vec{b}$ el sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és determinat*
9. *El sistema lineal $\mathsf{A}\vec{x} = \vec{0}$ és determinat.*
10. *L'espai nul de A és* $\operatorname{Nul} \mathsf{A} = \{\vec{0}\}$ □

Però encara ens falta provar una altra propietat important de les matrius invertibles: segons la defiició de la matriu inversa de A, $\mathsf{AA}^{-1} = \mathrm{I}$; però resulta que el producte $\mathsf{A}^{-1}\mathsf{A}$ també és igual a la identitat.[2]

PROPIETAT 8.3.

Donades les matrius quadrades d'ordre n, A i B, si $\mathsf{AB} = \mathrm{I}$ aleshores $\mathsf{BA} = \mathrm{I}$.

Demostració: Si $\mathsf{AB} = \mathrm{I}$ aleshores $\mathsf{B} = \mathsf{A}^{-1}$ i $\operatorname{rang} \mathsf{A} = n$; per tant, la forma esglaonada reduïda de A és I i es pot obtenir fent operacions elementals sobre A. Si fem aquestes operacions elementals sobre la matriu ampliada $\left[\mathsf{A} \mid \mathrm{I}\right]$ obtindrem

$$\mathsf{E}_p \dots \mathsf{E}_2\mathsf{E}_1 \left[\mathsf{A} \mid \mathrm{I}\right] = \left[\mathrm{I} \mid \mathsf{B}\right]$$

Això vol dir que

$$\mathsf{E}_p \dots \mathsf{E}_2\mathsf{E}_1\mathsf{A} = \mathrm{I}$$
$$\mathsf{E}_p \dots \mathsf{E}_2\mathsf{E}_1 = \mathsf{B}$$

de manera que substituint la segona equació en la primera obtenim $\mathsf{BA} = \mathrm{I}$. □

[2] De fet, la definició habitual de la matriu inversa exigeix les dues igualtats $\mathsf{AA}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathrm{I}$.

Per acabar, veurem que el producte de matrius invertibles també ho és.

PROPIETAT 8.4.

Siguen A_1 i A_2 dues matrius $n \times n$. Aleshores A_1 i A_2 són invertibles si i només si el producte $\mathsf{A}_1\mathsf{A}_2$ és invertible.
En aquest cas,

$$(\mathsf{A}_1\mathsf{A}_2)^{-1} = \mathsf{A}_2^{-1}\mathsf{A}_1^{-1}$$

Demostració: Suposem que les dues matrius A_1 i A_2 són invertibles. Llavors, si multipliquem $\mathsf{A}_1\mathsf{A}_2$ per $\mathsf{A}_2^{-1}\mathsf{A}_1^{-1}$ tindrem

$$\mathsf{A}_1 \underbrace{\mathsf{A}_2\mathsf{A}_2^{-1}}_{\mathrm{I}} \mathsf{A}_1^{-1} = \mathsf{A}_1\mathsf{A}_1^{-1} = \mathrm{I}$$

així que la inversa de $\mathsf{A}_1\mathsf{A}_2$ és $\mathsf{A}_2^{-1}\mathsf{A}_1^{-1}$.

Si ara suposem que $\mathsf{A}_1\mathsf{A}_2$ és invertible, multiplicant A_1 per la matriu $\mathsf{A}_2\,(\mathsf{A}_1\mathsf{A}_2)^{-1}$ tindrem

$$\mathsf{A}_1\Big(\mathsf{A}_2\,(\mathsf{A}_1\mathsf{A}_2)^{-1}\Big) = (\mathsf{A}_1\mathsf{A}_2)\,(\mathsf{A}_1\mathsf{A}_2)^{-1} = \mathrm{I}$$

així que A_1 és invertible i la seua inversa és $\mathsf{A}_2\,(\mathsf{A}_1\mathsf{A}_2)^{-1}$. Finalment, com que $\mathsf{A}_2 = \mathsf{A}_1^{-1}(\mathsf{A}_1\mathsf{A}_2)$, A_2 també és invertible. $\square$

COROL·LARI 8.5.

El producte $\mathsf{A}_1\mathsf{A}_2\ldots\mathsf{A}_p$ és invertible si i només si ho són totes les matrius $\mathsf{A}_1, \mathsf{A}_2, \ldots, \mathsf{A}_p$. En tal cas,

$$\Big(\mathsf{A}_1\mathsf{A}_2\ldots\mathsf{A}_p\Big)^{-1} = \mathsf{A}_p^{-1}\ldots\mathsf{A}_2^{-1}\mathsf{A}_1^{-1} \quad \square$$

☞ Pareu atenció al fet que la inversa del producte és el producte de les inverses *en l'ordre contrari.*

8.2. INVERSES DE LES MATRIUS ELEMENTALS

Com que les matrius elementals tenen sempre rang màxim, podem assegurar que totes les matrius elementals són invertibles. A més a més, és molt fàcil calcular les inverses d'aquestes matrius.

PROPIETAT 8.6.

Totes les matrius elementals són invertibles. A més a més, la inversa d'una matriu elemental és també elemental (i del mateix tipus). En concret,

1. *La inversa d'una matriu elemental del tipus permutació és la mateixa matriu:* $\mathsf{E}_{i,j}^{-1} = \mathsf{E}_{i,j}$
2. *La inversa d'una matriu elemental del tipus escalat amb factor d'escala α és la matriu del tipus escalat amb factor d'escala $1/\alpha$:* $\mathsf{E}_i(\alpha) = \mathsf{E}_i(1/\alpha)$
3. *La inversa d'una matriu elemental del tipus reducció amb paràmetre α és la matriu del tipus reducció amb paràmetre $-\alpha$:* $\mathsf{E}_{i,j}^{-1}(\alpha) = \mathsf{E}_{i,j}(-\alpha)$

Totes aquestes propietats es comproven sense més que fer les multiplicacions corresponents. Per exemple,

$$\mathsf{E}_{2,3}^{-1}(\alpha)\mathsf{E}_{2,3}(-\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\alpha \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \square$$

☞ Però el més interessant d'aquesta propietat és que la matriu inversa d'una matriu elemental és la matriu elemental que desfà l'operació elemental que feia la primera: l'operació inversa de sumar a una fila el doble d'una altra és restar a una fila el doble de l'altra; el contrari de multiplicar una fila per 2 és dividir-la per 2; i per neutralitzar la permutació de dues files cal tornar-les a permutar.

☞ Notem que l'algorisme de Gauss-Jordan és admissible, com a mètode per resoldre un sistema d'equacions lineals, perquè les matrius elementals són invertibles.

Del fet que les inverses de les matrius elementals són també matrius elementals podem deduir una nova caracterització de les matrius invertibles.

PROPIETAT 8.7.

La matriu quadrada A és invertible si i només si A és un producte de matrius elementals.

Demostració: Si A és un producte de matrius elementals, llavors A és invertible perquè les matrius elementals ho són. Recíprocament, hem vist adés que si B és la inversa de A, llavors B es pot obtenir com a producte de les matrius elementals que transformen A en la identitat: $\mathsf{B} = \mathsf{E}_p \dots \mathsf{E}_2\mathsf{E}_1$. Llavors,

$$\mathsf{A} = \mathsf{B}^{-1} = \mathsf{E}_1^{-1}\mathsf{E}_2^{-1} \dots \mathsf{E}_p^{-1}$$

que és un producte de matrius elementals. □

8.3. Teorema de caracterització de les matrius invertibles

La darrera propietat que hem demostrat afig un nou item a la llista d'equivalències del teorema 8.2. De fet, la invertibilitat és tan important en l'àlgebra lineal que en tots els temes d'aquest curs anem a trobar noves equivalències del fet que una matriu siga invertible.

Al final del curs haurem vist, i seran òbvies, totes aquestes equivalències (i alguna més):

TEOREMA 8.8.

Siga A una matriu quadrada d'ordre n. Totes les afirmacions següents són equivalents:

1. A *és invertible*
2. *L'equació matricial* $\mathsf{AX} = \mathrm{I}$ *és compatible*
3. *El rang de* A *és* n
4. *Les files de* A *són linealment independents*
5. *Les columnes de* A *són linealment independents*
6. *La forma esglaonada reduïda de* A *és la matriu identitat*
7. *Per a qualsevol vector* $\vec{b}$ *el sistema lineal* $\mathsf{A}\vec{x} = \vec{b}$ *és compatible*
8. *Per a qualsevol vector* $\vec{b}$ *el sistema lineal* $\mathsf{A}\vec{x} = \vec{b}$ *és determinat*
9. *El sistema lineal* $\mathsf{A}\vec{x} = \vec{0}$ *és determinat*
10. *L'espai nul de* A *és* $\mathrm{Nul}\,\mathsf{A} = \{\vec{0}\}$
11. *Existeix una matriu* B *de manera que* $\mathsf{BA} = \mathrm{I}$
12. A *és un producte de matrius elementals*
13. *El determinant de* A *no és zero*
14. *L'espai columna de* A *és* $\mathrm{Col}\,\mathsf{A} = \mathbb{K}^n$
15. *L'espai nul de la matriu transposada* A^t *és* $\mathrm{Nul}\,\mathsf{A}^t = \{\vec{0}\}$
16. *L'espai fila de* A *és* $\mathrm{Fil}\,\mathsf{A} = \mathbb{K}^n$
17. *El nucli de l'aplicació lineal* $f(\vec{x}) = \mathsf{A}\vec{x}$ *és* $\mathrm{Nuc}\,f = \{\vec{0}\}$
18. *L'aplicació lineal* $f(\vec{x}) = \mathsf{A}\vec{x}$ *és injectiva*
19. *L'aplicació lineal* $f(\vec{x}) = \mathsf{A}\vec{x}$ *és suprajectiva*
20. *L'aplicació lineal* $f(\vec{x}) = \mathsf{A}\vec{x}$ *és bijectiva*
21. *El nombre* 0 *no és valor propi de* A
22. *El nombre* 0 *no és valor singular de* A

8.4. Exercicis

Exercici 8.1. Determineu si les matriu següents són invertibles i, en cas que ho siguen, calculeu la inversa i escriviu les matrius com a producte de matrius elementals.

(a) $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$ (c) $C = \begin{bmatrix} 2 & 2 & 2 \\ 3 & 2 & 2 \\ 5 & 1 & 0 \end{bmatrix}$ (d) $D = \begin{bmatrix} 2 & 2 & 2 \\ 3 & 2 & 2 \\ 1 & 0 & 0 \end{bmatrix}$

(e) $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (f) $F = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (g) $G = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (h) $H = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

(i) $L = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 3 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (j) $M = \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$ (k) $N = \begin{bmatrix} 2 & -i \\ i & 3 \end{bmatrix}$ (l) $P = \begin{bmatrix} 1-i & 0 \\ 0 & 1+i \end{bmatrix}$

Exercici 8.2. Calculeu la inversa, si existeix, de la matriu $A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$

Exercici 8.3. Sense fer cap càlcul digueu quina és la matriu inversa de $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{bmatrix}$

Exercici 8.4. Sabent que les matrius A i B són invertibles, aïlleu la matriu X en l'expressió

$$BA^2XB = C - 2I$$

Exercici 8.5. Proveu que si $A = \begin{bmatrix} 2 & 6 & 4 \\ 1 & 0 & -2 \\ 3 & 7 & -2 \end{bmatrix}$ aleshores $A^3 - 8A - 32I = O$ i utilitzeu aquest resultat per a calcular la matriu inversa A^{-1}.

Exercici 8.6. **(Càlcul ràpid de la inversa d'una matriu 2×2)** (a) Calculeu el producte AB essent

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

(b) El nombre $|A| = a_{11}a_{22} - a_{12}a_{21}$ s'anomena *determinant* de la matriu A. Proveu que la matriu A és invertible si i només si el seu determinant és no nul. En el cas que A siga invertible, trobeu una fórmula per a la matriu inversa de A. (c) Feu servir el determinant per calcular inverses, si existeixen, de les matrius 2×2 de l'exercici 8.1.

Exercici 8.7. **(Inverses laterals)** Si A és una matriu $m \times n$ i B és una matriu $n \times m$ direm que B és una *inversa dreta* de A si $AB = I$. De la mateixa manera, direm que B és una *inversa esquerra* de A si $BA = I$.
(a) Quina condició s'ha de complir perquè la matriu A tinga alguna inversa dreta?
(b) Quina condició s'ha de complir perquè la inversa dreta de la matriu A siga única?
(c) Quina condició s'ha de complir perquè la matriu A tinga alguna inversa esquerra?
(d) Quina condició s'ha de complir perquè la inversa esquerra de la matriu A siga única?
(e) Quines condicions s'han de complir perquè la matriu A tinga inverses pels dos costats?

Exercici 8.8. Calculeu totes les matrius inverses esquerres i/o dretes, en cas que existisquen, de les matrius següents:

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

9. Matrius simètriques i matrius ortogonals

Les matrius simètriques i les matrius ortogonals tenen molta importància en tots els problemes relacionats amb la geometria de $\mathbb{R}^n$, per la seua relació amb el producte escalar.

Les matrius ortogonals, és a dir les matrius quadrades les columnes de les quals són ortonormals, representen les transformacions lineals que conserven angles i distàncies. I les entrades de la matriu simètrica A^tA són tots els productes escalars de les columnes de A. Les matrius simètriques també tenen molta importància, perquè apareixen en força aplicacions pràctiques (i també associades amb les matrius ortogonals quan es diagonalitzen).

Al llarg del curs, trobarem les aplicacions d'aquests tipus de matrius a la solució dels problemes de mínims quadrats, la diagonalització de les matrius simètriques i la factorització en valors i vectors singulars.

9.1. La matriu transposada

Transposar una matriu és canviar les seues files per columnes: construir una nova matriu les files de la qual són les columnes de la primera. La transposada de la matriu A la representem com A^t.

Per exemple, la matriu transposada de $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ és $A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

És evident que si A és una matriu $m \times n$, aleshores A^t serà una matriu $n \times m$. En particular, la matriu transposada d'un vector és una matriu fila, i la matriu transposada d'una matriu quadrada també és quadrada.

El rang de la matriu transposada coincideix amb el de la matriu original:

Propietat 9.1.

Per a qualsevol matriu A*,* $\operatorname{rang} A^t = \operatorname{rang} A$

Demostració: L'únic que hem de fer, per demostrar-ho, és recordar que el rang es pot interpretar, bé com el nombre de files linealment independents que té la matriu, bé com el de columnes linealment independents. Però com les files de A^t són les columnes de A, es clar que:

$$\operatorname{rang} A^t = \text{nombre de files independents de } A^t = \text{nombre de columnes independents de } A = \operatorname{rang} A \quad \square$$

L'*operació* de transposar commuta amb la de sumar o multiplicar per un escalar (és a dir, que la suma de les transposades és la transposada de la suma i la transposada del producte d'un escalar per una matriu és el producte de l'escalar per la transposada de la matriu). En canvi, amb el producte *s'inverteix l'ordre.* Tot això es pot provar molt fàcilment.

Propietats 9.2.

1) La transposada de la suma és la suma de les transposades:

$$(A + B)^t = A^t + B^t$$

2) La transposada del producte per un escalar és el producte de la transposada per l'escalar:

$$(\lambda A)^t = \lambda A^t$$

3) La transposada del producte és el producte de les transposades en l'ordre contrari:

$$(AB)^t = B^tA^t \quad \square$$

Finalment, l'*operació* d'invertir una matriu també commuta amb la transposició.

PROPIETAT 9.3.

Si A *és una matriu invertible, llavors la seua transposada també és invertible i la inversa de la transposada és la transposada de la inversa:*

$$(\mathsf{A}^t)^{-1} = \left(\mathsf{A}^{-1}\right)^t$$

Demostració: Per a provar-ho, multipliquem A^t per $(\mathsf{A}^{-1})^t$ i recordem que la transposada del producte és igual al producte de les transposades en ordre contrari:

$$\mathsf{A}^t\,(\mathsf{A}^{-1})^t = (\mathsf{A}^{-1}\mathsf{A})^t = \mathrm{I}^t = \mathrm{I} \quad \square$$

9.2. MATRIUS SIMÈTRIQUES I ANTISIMÈTRIQUES

Una matriu A és *simètrica* si $\mathsf{A}^t = \mathsf{A}$. Per exemple, les matrius

$$\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \qquad \mathsf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \qquad \mathsf{A} = \begin{bmatrix} 0 & 2 & e^3 \\ 2 & 0 & -1 \\ e^3 & -1 & \pi \end{bmatrix}$$

són simètriques. S'anomenen d'aquesta manera perquè s'hi observa una simetria respecte a la diagonal principal (tècnicament, $a_{ij} = a_{ji}$, per a tots els índex i i j).

☞ Les matrius simètriques són sempre matrius quadrades.

PROPIETATS 9.4.

1) *Si* A *i* B *són simètriques, llavors* $\mathsf{A}+\mathsf{B}$ *també és simètrica.*
2) *Si* A *és simètrica i* λ *és un escalar, llavors* $\lambda\mathsf{A}$ *també és simètrica.*
3) *Si* A *i* B *són simètriques, llavors* AB *és simètrica si i només si* A *i* B *commuten entre elles (és a dir, si* $\mathsf{AB} = \mathsf{BA}$*).*
4) *Si* A *és una matriu invertible i simètrica, llavors la seua inversa també és simètrica.*

☞ Fixeu-vos bé en el cas del producte: perquè el producte de dues matrius simètriques siga una matriu simètrica cal que aquestes dues matrius commuten.

Una matriu A és *antisimètrica* si $\mathsf{A}^t = -\mathsf{A}$. Per exemple, les matrius

$$\mathsf{A} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix} \qquad \mathsf{A} = \begin{bmatrix} 0 & 2 & -e^3 \\ -2 & 0 & 1 \\ e^3 & -1 & 0 \end{bmatrix}$$

són antisimètriques. Element a element, la condició perquè la matriu A siga antisimètrica és que $a_{ij} = -a_{ji}$, per a tots els índex i i j. Observem que la diagonal de les matrius antisimètriques és sempre nul·la, perquè $a_{ii} = -a_{ii} \Rightarrow a_{ii} = 0$.

9.3. LA MATRIU $\mathsf{A}^t\mathsf{A}$ I EL PRODUCTE ESCALAR REAL

Fent servir la transposició de matrius podem reinterpretar el producte escalar de dos vectors reals $\vec{u}$ i $\vec{v}$ com un producte de dues matrius, perquè

$$\vec{u}\cdot\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} = \vec{u}^t\vec{v}$$

$$\vec{u}\cdot\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \vec{u}^t\vec{v}$$

De manera més general, si A és una matriu real, quan multipliquem A^t per A estem fent els productes escalars de les files de A^t per les columnes de A; però les files de A^t són les columnes de A, així que la multiplicació $\mathsf{A}^t\mathsf{A}$ conté tots els productes escalars entre les columnes de A,

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} \vec{a}_1^{\,t} \\ \vec{a}_2^{\,t} \\ \cdots \\ \vec{a}_n^{\,t} \end{bmatrix} \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \cdot \vec{a}_1 & \vec{a}_1 \cdot \vec{a}_2 & \cdots & \vec{a}_1 \cdot \vec{a}_n \\ \vec{a}_2 \cdot \vec{a}_1 & \vec{a}_2 \cdot \vec{a}_2 & \cdots & \vec{a}_2 \cdot \vec{a}_n \\ \cdots & & & \\ \vec{a}_n \cdot \vec{a}_1 & \vec{a}_n \cdot \vec{a}_2 & \cdots & \vec{a}_n \cdot \vec{a}_n \end{bmatrix}$$

per aquest motiu, en tots els problemes que tenen relació amb el producte escalar (ortogonalitat, projeccions, distàncies, angles, valors singulars...) sempre cal treballar amb aquesta matriu.

Més endavant estudiarem les propietats de la matriu $\mathsf{A}^t\mathsf{A}$. De moment, observarem que és una matriu simètrica, perquè

$$\left(\mathsf{A}^t\mathsf{A}\right)^t = \mathsf{A}^t\left(\mathsf{A}^t\right)^t = \mathsf{A}^t\mathsf{A}$$

i la farem servir per definir les matrius ortogonals.

9.4. Matrius ortogonals

Una matriu quadrada real Q és ortogonal si $\mathsf{Q}^t\mathsf{Q} = \mathrm{I}$, és a dir, si és invertible i la inversa de Q és la transposada Q^t. Per exemple, la matriu identitat és ortogonal, perquè $\mathrm{I}^t\mathrm{I} = \mathrm{I}\,\mathrm{I} = \mathrm{I}$, i la matriu

$$\mathsf{Q} = \begin{bmatrix} 2/3 & 1/3 & -2/3 \\ -2/3 & 2/3 & -1/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix}$$

també és ortogonal, perquè

$$\mathsf{Q}^t\mathsf{Q} = \begin{bmatrix} 2/3 & -2/3 & 1/3 \\ 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 & -2/3 \\ -2/3 & 2/3 & -1/3 \\ 1/3 & 2/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Aquestes matrius s'anomenen *ortogonals*, perquè el conjunt de les seues columnes és ortonormal:[3] com que $\mathsf{Q}^t\mathsf{Q}$ conté tots els productes escalars entre les columnes de Q, si aquest producte és la identitat,

$$\mathsf{Q}^t\mathsf{Q} = \begin{bmatrix} \vec{q}_1 \cdot \vec{q}_1 & \vec{q}_1 \cdot \vec{q}_2 & \cdots & \vec{q}_1 \cdot \vec{q}_n \\ \vec{q}_2 \cdot \vec{q}_1 & \vec{q}_2 \cdot \vec{q}_2 & \cdots & \vec{q}_2 \cdot \vec{q}_n \\ \cdots & & & \\ \vec{q}_n \cdot \vec{q}_1 & \vec{q}_n \cdot \vec{q}_2 & \cdots & \vec{q}_n \cdot \vec{q}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

llavors la diagonal d'uns indica que totes les columnes són vectors unitaris i la resta d'entrades (zeros) que són mútuament ortogonals.

Propietat 9.5.

Una matriu quadrada real és ortogonal si i només si les seues columnes formen un conjunt ortonormal de vectors.

D'altra banda, la multiplicació $\mathsf{Q}\mathsf{Q}^t$ conté els productes escalars entre les files de Q, així que les files de les matrius ortogonals també són ortonormals.

Les matrius ortogonals són interessants per molts motius, especialment per les seues aplicacions geomètriques. Vegem-ne algunes.

Primer de tot, podem observar que, si la matriu Q és ortogonal, llavors la resolució del sistema d'equacions lineals $\mathsf{Q}\vec{x} = \vec{b}$ es redueix a una multiplicació:

$$\mathsf{Q}\vec{x} = \vec{b} \iff \mathsf{Q}^t\mathsf{Q}\vec{x} = \mathsf{Q}^t\vec{b} \iff \vec{x} = \mathsf{Q}^t\vec{b}$$

Si multipliquem una matriu ortogonal pels vectors $\vec{u}$ i $\vec{v}$, llavors el producte escalar dels nous vectors és el mateix que el dels vectors originals, perquè

$$\mathsf{Q}\vec{u} \cdot \mathsf{Q}\vec{v} = (\mathsf{Q}\vec{u})^t(\mathsf{Q}\vec{v}) = \vec{u}^{\,t}\mathsf{Q}^t\mathsf{Q}\vec{v} = \vec{u}^{\,t}\vec{v} = \vec{u} \cdot \vec{v}$$

Aquesta propietat s'expressa dient que *les matrius ortogonals conserven els productes escalars.* Com a conseqüència d'això, es pot provar fàcilment que aquestes matrius també conserven els angles, les longituds i les distàncies.

[3] Recordeu que un conjunt de vectors és ortonormal si tots són unitaris i cadascun dels vectors és ortogonal a tots els altres.

9.5.* Matrius hermítiques i unitàries

L'equivalent complex de les matrius simètriques i de les ortogonals són les matrius hermítiques i les unitàries. Com que el producte escalar dels vectors complexos $\vec{u}$ i $\vec{v}$ és

$$\vec{u}\cdot\vec{v} = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \cdots & \overline{u_n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix}$$

el paper que fa la matriu transposada en el cas real, quan treballem amb matrius complexes el fa la matriu *adjunta* o *transposada conjugada*, és a dir, la matriu

$$\mathsf{A}^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{m2}} \\ \cdots & & & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{mn}} \end{bmatrix}$$

que s'obte transposant A i canviant tots els elements pels seus complexos conjugats.

Definicions 9.1.

1. *Una matriu complexa* A *és hermítica (o autoadjunta) si* $\mathsf{A}^* = \mathsf{A}$.
2. *Una matriu complexa* A *és antihermítica si* $\mathsf{A}^* = -\mathsf{A}$.
3. *Una matriu complexa* U *és unitària si les seues columnes formen un conjunt ortonormal, és a dir, si* $\mathsf{U}^*\mathsf{U} = \mathrm{I}$.

Exemples 9.1.

La matriu $\mathsf{A} = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$ és hermítica, perquè $\mathsf{A}^* = \begin{bmatrix} \overline{1} & \overline{-i} \\ \overline{i} & \overline{2} \end{bmatrix} = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} = \mathsf{A}$.

La matriu $\mathsf{B} = \begin{bmatrix} i & -2 \\ 2 & i \end{bmatrix}$ és antihermítica, perquè $\mathsf{B}^* = \begin{bmatrix} \overline{i} & \overline{2} \\ \overline{-2} & \overline{i} \end{bmatrix} = \begin{bmatrix} -i & 2 \\ -2 & -i \end{bmatrix} = -\mathsf{B}$.

La matriu $\mathsf{U} = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ és unitària, perquè

$$\mathsf{U}^*\mathsf{U} = \begin{bmatrix} \overline{0} & \overline{-i} \\ \overline{i} & \overline{0} \end{bmatrix} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

9.6. Exercicis

9.6.1. La matriu transposada i les matrius simètriques i antisimètriques

Exercici 9.1. Proveu que, si A i B són matrius quadrades del mateix ordre i A és simètrica llavors, la matriu $\mathsf{B}^t\mathsf{A}\mathsf{B} + \mathsf{A}$ també és simètrica.

Exercici 9.2. Si la matriu A i B, del mateix ordre, són respectivament, simètrica i antisimètrica, què podem dir de $\mathsf{AB} + \mathsf{BA}$ i de $\mathsf{AB} - \mathsf{BA}$?

Exercici 9.3. (a) Proveu que la matriu $\mathsf{A}\mathsf{A}^t$ és una matriu simètrica (per a qualsevol matriu A, no necessàriament quadrada).

(b) Donada la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}$ calculeu $\mathsf{A}\mathsf{A}^t$ i $\mathsf{A}^t\mathsf{A}$ i comproveu que són matrius simètriques distintes.

(c) Si A és una matriu quadrada, és cert que $\mathsf{A}\mathsf{A}^t = \mathsf{A}^t\mathsf{A}$?

Exercici 9.4. Proveu que, si A és una matriu quadrada, llavors la matriu $\mathsf{A} + \mathsf{A}^t$ és una matriu simètrica i que la matriu $\mathsf{A} - \mathsf{A}^t$ és una matriu antisimètrica.

Proveu que, per a qualsevol matriu A,

$$\mathsf{A} = \frac{\mathsf{A} + \mathsf{A}^t}{2} + \frac{\mathsf{A} - \mathsf{A}^t}{2}$$

(en conseqüència, qualsevol matriu quadrada és la suma d'una matriu simètrica i una antisimètrica).

9.6.2. Matrius ortogonals

Exercici 9.5. Proveu que les matrius següents són ortogonals:

$$\mathsf{A} = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \qquad \mathsf{P} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathsf{B} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

Exercici 9.6. Proveu que les matrius de la forma

$$\mathsf{A}_\alpha = \begin{bmatrix} \cos\alpha & \cos(\alpha + \pi/2) \\ \sin\alpha & \sin(\alpha + \pi/2) \end{bmatrix}$$

són ortogonals i interpreteu aquest fet geomètricament.

Exercici 9.7. Resoleu el sistema

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ -x_1 - x_2 + x_3 + x_4 &= -6 \\ x_1 - x_2 - x_3 + x_4 &= -2 \\ -x_1 + x_2 - x_3 + x_4 &= 0 \end{aligned}$$

aprofitant el fet que la matriu B de l'exercici anterior és ortogonal.

Exercici 9.8. Proveu que el producte de dues matrius ortogonals també és una matriu ortogonal.

Exercici 9.9. Una matriu permutació és la matriu que resulta de reordenar arbitràriament les files de la matriu identitat. Per exemple, les matrius

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

són matrius permutació.

(a) Quantes matrius permutació d'ordre n hi ha?
(b) Totes les matrius permutació són elementals?
(c) Proveu que les matrius permutació són ortogonals.

Exercici 9.10. Proveu que si Q és una matriu ortogonal, llavors

1. La norma del vector $\mathsf{Q}\vec{u}$ coincideix amb al norma de $\vec{u}$.
2. L'angle entre els vectors $\mathsf{Q}\vec{u}$ i $\mathsf{Q}\vec{v}$ és el mateix que el que hi ha entre $\vec{u}$ i $\vec{v}$.

9.6.3. Matrius heremítiques i antihermítiques

Exercici 9.11. Suposem que B i C són les parts real i imaginària de A, respectivament (és a dir, B i C són reals i A = B + iC). Proveu les propietats següents:

(a) Si la matriu A és hermítica llavors, B és simètrica i C és antisimètrica.
(b) Si la matriu A és hermítica llavors, la diagonal de A és real.
(c) Si la matriu A és antihermítica si i només si iA és hermítica.
(d) Si A és antihermítica llavors, els elements de la diagonal de A són imaginaris purs.
(e) Si A és antihermítica llavors, B és antisimètrica i C és simètrica.

10. Matrius triangulars. Factoritzacions LU

En aquesta unitat dediquem una mica d'atenció a la factorització LU, que és la interpretació matricial de l'algorisme de Gauss i la base d'algun dels mètodes directes més habituals en l'àlgebra lineal numèrica.

10.1. Matrius diagonals i matrius triangulars

Les matrius diagonals són les més senzilles, almenys des del punt de vista de la facilitat amb que s'hi treballa: normalment, qualsevol propietat és immediata i qualsevol problema es resol sense cap dificultat, si treballem amb una matriu diagonal.

La *diagonal principal* d'una matriu $m \times n$ és la formada pels elements a_{11}, a_{22}, …

Definició 10.1.

Una matriu quadrada A *és diagonal si tots els elements que no es troben a la diagonal principal són nuls.*

És a dir, que la matriu A és diagonal si té aquesta forma:

$$\mathsf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Per exemple, les matrius següents són diagonals:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Com dèiem abans, tots els problemes amb matrius són trivials, si la matriu és diagonal: per exemple, el rang d'una matriu diagonal és el nombre d'entrades no nul·les a la seua diagonal, i la inversa de la matriu diagonal A existeix si no hi ha cap zero a la diagonal i, llavors,

$$\mathsf{A}^{-1} = \begin{bmatrix} 1/a_{11} & 0 & \cdots & 0 \\ 0 & 1/a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1/a_{nn} \end{bmatrix}$$

D'altra banda, la suma i el producte de dues matrius diagonals i la multiplicació d'una matriu diagonal per un escalar són matrius diagonals. Si la matriu d'un sistema d'equacions lineals és diagonal, llavors el podem resoldre molt fàcilment, perquè cada equació conté, com a molt, una incògnita:

$$\mathsf{A}\vec{x} = \vec{b} \iff \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \vec{x} = \vec{b} \iff \begin{array}{c} a_{11}x_1 = b_1 \\ a_{22}x_2 = b_2 \\ \cdots \\ a_{nn}x_n = b_n \end{array}$$

Si la matriu A és triangular, les coses ja no són trivials, però continuen sent bastant simples.

Definició 10.2.

Una matriu quadrada és triangular superior si tots els elements que es troben davall la diagonal principal són nuls.

És a dir, que la matriu A és triangular superior si té aquesta forma:

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Totes les matrius diagonals són triangulars superiors; però és clar que n'hi ha, de triangulars superiors, que no són diagonals. Com ara, aquestes:

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

És fàcil veure que una matriu triangular superior serà invertible si no té cap zero a la diagonal; però ara el càlcul de la inversa requereix una mica de treball.

EXEMPLE 10.1.

Calculeu la matriu inversa de

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{bmatrix}$$

Per tal de calcular la forma esglaonada reduïda de la matriu ampliada $\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & -3 & 0 & 0 & 1 \end{array}\right]$ només hem de fer els passos finals de l'algorisme de Gauss-Jordan, perquè la matriu ja és esglaonada:

$$\left[\mathsf{A} \mid \mathrm{I}\right] = \left[\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 1 & 0 \\ 0 & 0 & -3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\mathsf{E}_{1,3}(2/3)\mathsf{E}_{2,3}(-1/3)} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 2/3 \\ 0 & 2 & 0 & 0 & 1 & -1/3 \\ 0 & 0 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{\mathsf{E}_2(1/2)\mathsf{E}_3(-1/3)} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 2/3 \\ 0 & 1 & 0 & 0 & 1/2 & -1/6 \\ 0 & 0 & 1 & 0 & 0 & -1/3 \end{array}\right]$$

així que, $\mathsf{A}^{-1} = \begin{bmatrix} 1 & 0 & 2/3 \\ 0 & 1/2 & -1/6 \\ 0 & 0 & -1/3 \end{bmatrix}$. □

☞ La inversió d'una matriu invertible i triangular superior es redueix a la fase final de l'algorisme de Gauss-Jordan, perquè la matriu ja és, d'antuvi, esglaonada.

De la mateixa manera, si la matriu d'un sistema d'equacions lineals és triangular superior, normalment podrem resoldre'l fent servir únicament l'algorisme de substitució regressiva.

DEFINICIÓ 10.3.

Una matriu quadrada és triangular inferior si tots els elements que es troben per damunt de la diagonal principal són nuls.

És a dir, que la matriu A és triangular inferior si té aquesta forma:

$$\mathsf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

I les propietats d'aquestes matrius són anàlogues a les de les matrius triangulars superiors.

10.2. LA FACTORITZACIÓ LU

Gairebé tots els problemes de l'àlgebra lineal que es tracten amb matrius poden interpretar-se en termes de *factoritzacions* de matrius. Ja hem vist que l'esglaonament i tot allò que s'hi relaciona, com ara, els algorismes del tipus Gauss-Jordan, es poden interpretar com un producte de dues matrius. La factorització LU n'és un cas particular.

DEFINICIÓ 10.4.

Una factorització LU estricta de la matriu quadrada A *és la descomposició d'aquesta matriu com un producte* $\mathsf{A} = \mathsf{LU}$ *on*

(a) L *és una matriu triangular inferior*

(b) Totes les entrades de la diagonal principal de L *són iguals a* 1

(c) U *és una matriu triangular superior.*

Per exemple,

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3/2 \end{bmatrix}$$

és una factorització LU estricta.[4]

Hi ha matrius quadrades que no admeten una factorització LU estricta (és a dir, que no es poden escriure com a producte d'una matriu triangular inferior per una de triangular superior). Per exemple, la matriu $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ no té cap factorització LU estricta, perquè, si en tinguera,

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} \end{bmatrix}$$

i, llavors tindríem $u_{11} = 0$ i $1 = l_{21}u_{11} = 0$.

Per tal de determinar si existeix la factorització, i calcular-la, notem que, si existeix, la matriu L és sempre invertible, (perquè és triangular inferior i no té zeros a la diagonal); en conseqüència, si la matriu A admet una factorització LU, llavors L és un producte de matrius elementals. En altres paraules,

☞ La factorització LU estricta, si existeix, es pot obtenir fent servir l'algorisme de Gauss.

EXEMPLE 10.2.

Trobeu, si és possible, una factorització LU estricta de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & -2 \\ 2 & 6 & -2 \end{bmatrix}$$

Apliquem-hi l'algorisme de Gauss:

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & -2 \\ 2 & 6 & -2 \end{bmatrix} \xrightarrow{\mathsf{E}_{3,1}(-2)\mathsf{E}_{2,1}(1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 2 & -4 \end{bmatrix} \xrightarrow{\mathsf{E}_{3,2}(-1)} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{bmatrix}$$

Si anomenem U a aquesta matriu darrera tindrem

$$\mathsf{E}_{3,2}(-1)\mathsf{E}_{3,1}(-2)\mathsf{E}_{2,1}(1)\mathsf{A} = \mathsf{U}$$

o bé,

$$\mathsf{A} = \mathsf{E}_{2,1}(-1)\mathsf{E}_{3,1}(2)\mathsf{E}_{3,2}(1)\mathsf{U} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{\mathsf{L}} \underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{bmatrix}}_{\mathsf{U}} \quad \square$$

Aquest exemple ha funcionat correctament perquè només hem fet operacions elementals del tipus reducció, $\mathsf{E}_{i,j}(\alpha)$ amb $j > i$ (a una fila li sumem un múltiple d'una fila *anterior* a ella); com que aquestes matrius són totes triangulars inferiors amb només uns a la diagonal, el seu producte també té aquestes característiques. De fet, aquesta és la clau perquè puguem trobar una factorització LU estricta:

[4] El nom «LU» fa referència al fet que L és una matriu triangular inferior (***L**ower triangular matrix*) i U és triangular superior (***U**pper triangular matrix*).

☞ La condició que s'ha de complir perquè la matriu A admeta una factorització LU estricta és que siga possible transformar la matriu A en triangular superior fent únicament operacions elementals del tipus reducció $\mathsf{E}_{i,j}(\alpha)$ amb $j > i$ (sumar a una fila un múltiple d'una altra fila *anterior*).

L'exemple següent confirmarà aquesta condició.

EXEMPLE 10.3.

Trobeu, si és possible, una factorització LU estricta de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & -1 \\ 2 & 2 & 2 \end{bmatrix}$$

Con en l'exemple anterior, aplicarem l'algorisme de Gauss. En primer lloc, fem zeros en la primera columna:

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & -1 \\ 2 & 2 & 2 \end{bmatrix} \xrightarrow{\mathsf{E}_{3,1}(-2)\mathsf{E}_{2,1}(-1)} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -4 \\ 0 & -2 & -4 \end{bmatrix}$$

i ja no podem continuar esglaonant si no és que permutem les files segona i tercera:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & -4 \\ 0 & -2 & -4 \end{bmatrix} \xrightarrow{\mathsf{E}_{2,3}} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & -4 \end{bmatrix} = \mathsf{U}$$

D'aquesta manera hem aconseguit trobar una matriu U triangular superior que és equivalent per files a A, i no hi ha cap dificultat per a trobar la matriu L que transforma A en U:

$$\mathsf{A} = \mathsf{E}_{2,1}(1)\mathsf{E}_{3,1}(2)\mathsf{E}_{2,3}\mathsf{U} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}}_{\mathsf{L}} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & -4 \end{bmatrix}}_{\mathsf{U}}$$

però *aquesta matriu* L *no és triangular inferior* (perquè la matriu $\mathsf{E}_{2,3}$ no ho és; en realitat, L és el resultat de permutar les files d'una matriu triangular inferior). En conseqüència, la matriu A no admet una factorització LU estricta. Tot i això, als efectes pràctics, aquesta factorització és igualment útil.

DEFINICIÓ 10.5.

Una factorització LU (no estricta) de la matriu quadrada A *és la descomposició d'aquesta matriu com un producte* $\mathsf{A} = \mathsf{LU}$ *on*

(a) L *és una permutació de les files d'una matriu triangular inferior que té totes les entrades de la diagonal principal iguals a* 1*.*

(b) U *és una matriu triangular superior.*

☞ En aquest sentit més ampli, qualsevol matriu quadrada admet una factorització LU; que en aquesta factorització la matriu L siga o no triangular depén de que es puga o no aplicar el mètode de Gauss sense permutar files.[5]

10.2.1. APLICACIÓ DE LA FACTORITZACIÓ LU A LA RESOLUCIÓ DE SISTEMES LINEALS

Si hem de resoldre el sistema de n equacions lineals amb n incògnites $\mathsf{A}\vec{x} = \vec{b}$ i coneixem una factorització LU de A, podrem escriure el sistema en la forma $\mathsf{LU}\vec{x} = \vec{b}$ o, com que L és invertible, $\mathsf{U}\vec{x} = \mathsf{L}^{-1}\vec{b}$. Llavors, fent el canvi $\mathsf{L}^{-1}\vec{b} = \vec{y}$ ($\vec{b} = \mathsf{L}\vec{y}$) i el sistema es transforma en $\mathsf{U}\vec{x} = \vec{y}$:

$$\mathsf{A}\vec{x} = \vec{b} \Longleftrightarrow \mathsf{LU}\vec{x} = \vec{b} \Longleftrightarrow \begin{cases} \mathsf{L}\vec{y} = \vec{b} \\ \mathsf{U}\vec{x} = \vec{y} \end{cases}$$

així que podem procedir de la manera següent:

[5] Els autors més estrictes només admeten, lògicament, la factorització LU estricta.

Resolució del sistema $\mathsf{A}\vec{x} = \vec{b}$ a partir de $\mathsf{A} = \mathsf{LU}$:

(a) Resolem el sistema $\mathsf{L}\vec{y} = \vec{b}$

(b) Substituïm la solució que hem obtingut en $\mathsf{U}\vec{x} = \vec{y}$ i resolem aquest nou sistema d'equacions.

EXEMPLE 10.4.

Feu servir la factorització LU per a resoldre el sistema d'equacios lineals

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & -2 \\ 2 & 6 & -2 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$$

En l'exemple 10.2. hem calculat aquesta factorització LU de la matriu del sistema:

$$\mathsf{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}}_{\mathsf{L}} \underbrace{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{bmatrix}}_{\mathsf{U}}$$

de manera que resoldrem successivament els sistemes $\mathsf{L}\vec{y} = (0, 1, 4)$ i $\mathsf{U}\vec{x} = \vec{y}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} \iff \begin{Bmatrix} y_1 = 0 \\ y_2 = 1 + y_1 \\ y_3 = 4 - 2y_1 - y_2 \end{Bmatrix} \iff \begin{Bmatrix} y_1 = 0 \\ y_2 = 1 \\ y_3 = 3 \end{Bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \iff \begin{Bmatrix} x_1 = -2x_2 - x_3 \\ 2x_2 = 1 + x_3 \\ -3x_3 = 3 \end{Bmatrix} \iff \begin{Bmatrix} x_1 = 1 \\ x_2 = 0 \\ x_3 = -1 \end{Bmatrix} \quad \square$$

Notem que el fet que les dues matrius L i U siguen triangulars permet resoldre els dos sistemes amb molt poc esforç (simplement hi apliquem un algorisme de substitució *progressiva* i un altre de substitució regressiva).

EXEMPLE 10.5.

Feu servir la factorització LU per a resoldre el sistema d'equacions lineals

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & -1 \\ 2 & 2 & 2 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Aquesta matriu és la de l'exemple 10.3., que no admet una factorització LU estricta. La factorització que hem trobat és aquesta:

$$\mathsf{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}}_{\mathsf{L}} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & -4 \end{bmatrix}}_{\mathsf{U}}$$

Fent servir la mateixa tècnica, començarem per resoldre el sistema $\mathsf{L}\vec{y} = (1, 1, 0)$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \iff \begin{Bmatrix} y_1 & & & = 1 \\ y_1 & & + y_3 & = 1 \\ 2y_1 & + y_2 & & = 0 \end{Bmatrix}$$

Aquest sistema no és pròpiament triangular; però, com que la matriu L és una permutació d'una matriu triangular inferior, podem reordenar les equacions perquè ho siga:

$$\begin{aligned} y_1 \qquad\qquad &= 1 \\ 2y_1 + y_2 \qquad &= 0 \\ y_1 \qquad + y_3 &= 1 \end{aligned}$$

i el podem resoldre també per substitució progressiva:

$$\left.\begin{array}{ll} y_1 & = 1 \\ 2y_1 + y_2 & = 0 \\ y_1 \quad + y_3 & = 1 \end{array}\right\} \iff \left\{\begin{array}{l} y_1 = 1 \\ y_2 = -2y_1 \\ y_3 = 1 - y_1 \end{array}\right\} \iff \left\{\begin{array}{l} y_1 = 1 \\ y_2 = -2 \\ y_3 = 0 \end{array}\right\}$$

Finalment, resolem el sistema $\mathsf{U}\vec{x} = \vec{y}$:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -4 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \iff \left\{\begin{array}{l} x_1 = 1 - 2x_2 - 3x_3 \\ -2x_2 = -2 + 4x_3 \\ -4x_3 = 0 \end{array}\right\} \iff \left\{\begin{array}{l} x_1 = -1 \\ x_2 = 1 \\ x_3 = 0 \end{array}\right\} \quad \square$$

10.2.2. Càlcul ràpid de la matriu L

Com ja sabem, sempre que transformem una matriu A en una altra matriu B fent-hi operacions elementals per files, hi ha una matriu invertible T de manera que $\mathsf{TA} = \mathsf{B}$; aquesta matriu és el producte de les matrius elementals que fem servir en passar de A a B, i la manera més senzilla de trobar-la consisteix a fer sobre la identitat les mateixes operacions elementals que fem sobre A.

En el cas de la factorització LU, transformem A en U, així que existeix una matriu T de manera que $\mathsf{TA} = \mathsf{U}$, però el que ens interessa en realitat, és la matriu inversa $\mathsf{L} = \mathsf{T}^{-1}$. Això vol dir que el càlcul de U i L requereix aquests dos processos d'esglaonament:

(a) Càlcul de les matrius U i T, aplicant a la identitat les mateixes operacions elementals que a A. Esquemàticament,

$$\begin{bmatrix} \mathsf{A} & \mathrm{I} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathsf{U} & \mathsf{T} \end{bmatrix}$$

(b) Inversió de T. La manera més eficient és aquesta:

$$\begin{bmatrix} \mathsf{T} & \mathrm{I} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathrm{I} & \mathsf{L} \end{bmatrix}$$

Ara bé, en el cas que existisca una factorització LU estricta, és possible obtenir la matriu L simultàniament amb la matriu U i sense cap operació addicional, aplicant aquesta versió de l'algorisme de Gauss:

Factorització LU estricta

Aquest algorisme calcula una factorització LU estricta de la matriu $n \times n$ A, si existeix.

Partim de les matrius $\mathsf{U} = \mathsf{A}$ i $\mathsf{L} = \mathrm{I}$ i les transformarem fins que U siga triangular superior i L triangular inferior.

- Si $u_{11} = 0$, llavors és impossible trobar una factorització LU estricta.

  En cas contrari, feu zero, en la matriu U, tots els elements que hi ha davall u_{11}, sumant a cada fila a partir de la segona un múltiple adequat de la primera fila.

  Si a la fila i li sumeu la primera fila multiplicada per α canvieu, en la matriu L, l'element l_{i1} per $-\alpha$: $l_{i1} = -\alpha$.

- Si $u_{22} = 0$, llavors és impossible trobar una factorització LU estricta.

  En cas contrari, feu zero, en la matriu U, tots els elements que hi ha davall u_{22}, sumant a cada fila a partir de la tercera un múltiple adequat de la segona fila.

  Si a la fila i li sumeu la segona fila multiplicada per α canvieu, en la matriu L l'element l_{i2} per $-\alpha$: $l_{i2} = -\alpha$.

- Repetiu aquest procés amb els elements que hi ha davall $a_{33}, \dots, a_{n-1n-1}$

Al final del procés tindrem

$$\mathsf{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \cdots & & & & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix} \qquad \mathsf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

Exemple 10.6.

Trobeu una factorització LU estricta de la matriu

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -2 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ -1 & -2 & 2 & 0 \end{bmatrix}$$

Comencem amb

$$L = I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad U = A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -2 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ -1 & -2 & 2 & 0 \end{bmatrix}$$

i apliquem l'algorisme de Gauss per a transformar U en triangular superior, fent servir únicament operacions elemental del tipus reducció i modificant simultàniament la matriu L.

Primer pas: Eliminem en la primera columna de A

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ -2 & -2 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ -1 & -2 & 2 & 0 \end{bmatrix} \xrightarrow{E_{4,1}(1)E_{2,1}(2)} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

i canviem $l_{2,1} = -2$, $l_{4,1} = -1$:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

Segon pas: Perquè la matriu siga triangular, ja només cal eliminar un element:

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \xrightarrow{E_{4,3}(-1)} \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = U$$

i canviar $l_{4,3} = 1$:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 1 \end{bmatrix} \qquad \square$$

☞ Recordeu que aquest mètode *ràpid* només és vàlid quan únicament fem operacions elementals del tipus reducció i, a més, les fem sistemàticament, d'esquerra a dreta i eliminat per sota els pivots.

10.2.3. Pivotatge parcial

En teoria, la combinació de l'algorisme de Gauss amb el de substitució regressiva ens proporciona la solució exacta de qualsevol sistema lineal. Tot i això, quan el sistema es resol amb l'ajut d'un ordinador (que és l'única manera raonable de resoldre'l, si no és que es tracta d'un sistema amb molt poques equacions i incògnites) es poden produir alguns errors d'arrodoniment en els càlculs o en l'emmagatzematge de les variables; aquests errors es transmeten als càlculs posteriors i de vegades un petit error en un càlcul intermig pot acabar produint un error inadmissible en el resultat final.

Per exemple, considerem el sistema lineal

$$\begin{aligned} 0{,}0001x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2 \end{aligned}$$

La matriu dels coeficents d'aquest sistema és

$$A = \begin{bmatrix} 0{,}0001 & 1 \\ 1 & 1 \end{bmatrix}$$

i podem obtenir-ne una factorització LU fent-hi una sola operació elemental:

$$\mathsf{A} = \begin{bmatrix} 0{,}0001 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{2,1}(-10\,000)} \begin{bmatrix} 0{,}0001 & 1 \\ 0 & -9\,999 \end{bmatrix} = \mathsf{U}$$

$$\mathsf{L} = \mathsf{E}_{2,1}(10\,000) = \begin{bmatrix} 1 & 0 \\ 10\,000 & 1 \end{bmatrix}$$

Resolem el sistema $\mathsf{L}\vec{y} = (1,2)$ per substitució progressiva,

$$\left.\begin{aligned} y_1 &= 1 \\ 10\,000y_1 + y_2 &= 2 \end{aligned}\right\} \longrightarrow \left.\begin{aligned} y_1 &= 1 \\ y_2 &= 2 - 10\,000y_1 \end{aligned}\right\} \longrightarrow \begin{aligned} y_1 &= 1 \\ y_2 &= -9\,998 \end{aligned}$$

i el sistema $\mathsf{U}\vec{x} = \vec{y}$ per substitució regressiva,

$$\left.\begin{aligned} 0{,}0001x_1 + x_2 &= 1 \\ -9\,999x_2 &= -9\,998 \end{aligned}\right\} \longrightarrow \left.\begin{aligned} x_1 &= 10\,000(1 - x_2) \\ x_2 &= 9\,998/9\,999 \end{aligned}\right\} \longrightarrow \begin{aligned} x_1 &= 10\,000/9\,999 = 1{,}00010001\ldots \\ x_2 &= 9\,998/9\,999 = 0{,}99989998\ldots \end{aligned}$$

de manera que hem obtingut la solució exacta sense cap dificultat.

Ara bé, si treballem amb una màquina que només pot treballar amb tres xifres significatives i arrodoneix els resultats, en resoldre el sistema

$$\begin{aligned} x_1 &= 10\,000(1 - x_2) \\ x_2 &= 9\,998/9\,999 \end{aligned}$$

en comptes de $x_2 = 1{,}0001\ldots$ obtindrem $x_2 = 1$ i, en calcular x_1 tindrem $x_1 = 10\,000(1 - x_2) = 0$, de manera que l'error d'arrodoniment en x_2 es multiplica per $10\,000$ en un sol pas.

Evidentment, aquest problema deriva del fet que les magnituds dels coeficients del sistema són d'ordres molt distints, i que hem multiplicat x_2 per un nombre molt gran; llavors, com que l'error en x_2 era $\epsilon = 1 - x_2 \approx 0{,}0001$, en calcular x_1,

$$x_1 = 10\,000(1 - x_2) = 10\,000(1 - (1 - \epsilon)) = 10\,000\epsilon \approx 1$$

Per minimitzar els efectes de la transmissió de l'error, és convenient de fer servir l'algorisme de Gauss amb l'estratègia del *pivotatge parcial*, que consisteix a elegir en cada pas el pivot més gran (en valor absolut) entre tots els possibles. Els sistemes de càlcul numèric, com ara, l'Scilab, fan servir aquesta estratègia sempre que fan operacions d'esglaonament.

En el nostre exemple, a l'hora d'obtenir la descomposició LU, elegiríem la segona fila com a fila pivot:

$$\mathsf{A} = \begin{bmatrix} 0{,}0001 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{2,1}} \begin{bmatrix} 1 & 1 \\ 0{,}0001 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{2,1}(-0{,}0001)} \begin{bmatrix} 1 & 1 \\ 0 & 0{,}9999 \end{bmatrix} = \mathsf{U}$$

$$\mathsf{L} = \mathsf{E}_{2,1}\mathsf{E}_{2,1}(0{,}0001) = \begin{bmatrix} 0{,}0001 & 1 \\ 1 & 0 \end{bmatrix}$$

Però la nostra màquina només aprecia tres xifres significatives, així que obtindrà

$$\mathsf{U} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathsf{L} = \begin{bmatrix} 0{,}0001 & 1 \\ 1 & 0 \end{bmatrix},$$

En resoldre el sistema $\mathsf{L}\vec{y} = (1,2)$ obtindrem $y_1 = 2, y_2 = 1 - 0{,}0002 \approx 1$ que substituïts en $\mathsf{U}\vec{x} = \vec{y}$ ens donen

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

i obtindrem la solució aproximada $x_1 = 1,\ x_2 = 1$. Ara l'error d'arrodoniment no s'ha incrementat i hem obtingut un resultat bastant més admissible.

Com hem dit adés, els programes de càlcul numèric matricial avançats (com l'Scilab) apliquen la tècnica del pivotatge parcial sempre que han de fer servir l'algorisme de Gauss. Per això, quan aquests programes calculen una factorització LU, elegeixen sistemàticament els pivots més adequats, de manera que les matrius L que obtenen gairebé mai no són triangulars. Per exemple, en el cas de la matriu A de l'exemple 10.2., el procés seria el següent:

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & -2 \\ 2 & 6 & -2 \end{bmatrix} \xrightarrow{\mathsf{E}_{1,3}} \begin{bmatrix} 2 & 6 & -2 \\ -1 & 0 & -2 \\ 1 & 2 & 1 \end{bmatrix}$$

(intercanviem les files 1a i 3a, per tal de fer servir el millor pivot)

$$\xrightarrow{\mathsf{E}_{3,1}(-1/2)\mathsf{E}_{2,1}(1/2)} \begin{bmatrix} 2 & 6 & -2 \\ 0 & 3 & -3 \\ 0 & -1 & 2 \end{bmatrix}$$

(ara no cal intercanviar files, perquè el pivot és l'adequat)

$$\xrightarrow{\mathsf{E}_{3,2}(1/3)} \begin{bmatrix} 2 & 6 & -2 \\ 0 & 3 & -3 \\ 0 & 0 & 1 \end{bmatrix} = \mathsf{U}$$

Finalment, fent les operacions corresponents, obtenim la matriu L:

$$\mathsf{L} = \begin{bmatrix} 1/2 & -1/3 & 1 \\ -1/2 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

10.3. Exercicis

10.3.1. Matrius triangulars

Exercici 10.1. Calculeu la matriu inversa, si existeix, de les matrius següents (els nombres a, b, c i d són tots no nuls):

(a) $\mathsf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (b) $\mathsf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ (c) $\mathsf{C} = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ (d) $\mathsf{D} = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$

Exercici 10.2. (Substitució progressiva) Resoleu el sistema d'equacions lineals

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & -1 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

10.3.2. Factoritzacions LU

Exercici 10.3. Donada la matriu $\mathsf{A} = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & 1 \\ 0 & 4 & 5 \end{bmatrix}$ (a) Trobeu, si és possible, una factorització LU estricta de A. (b) Trobeu una factorització LU de A fent servir l'algorisme de Gauss amb pivotatge parcial.

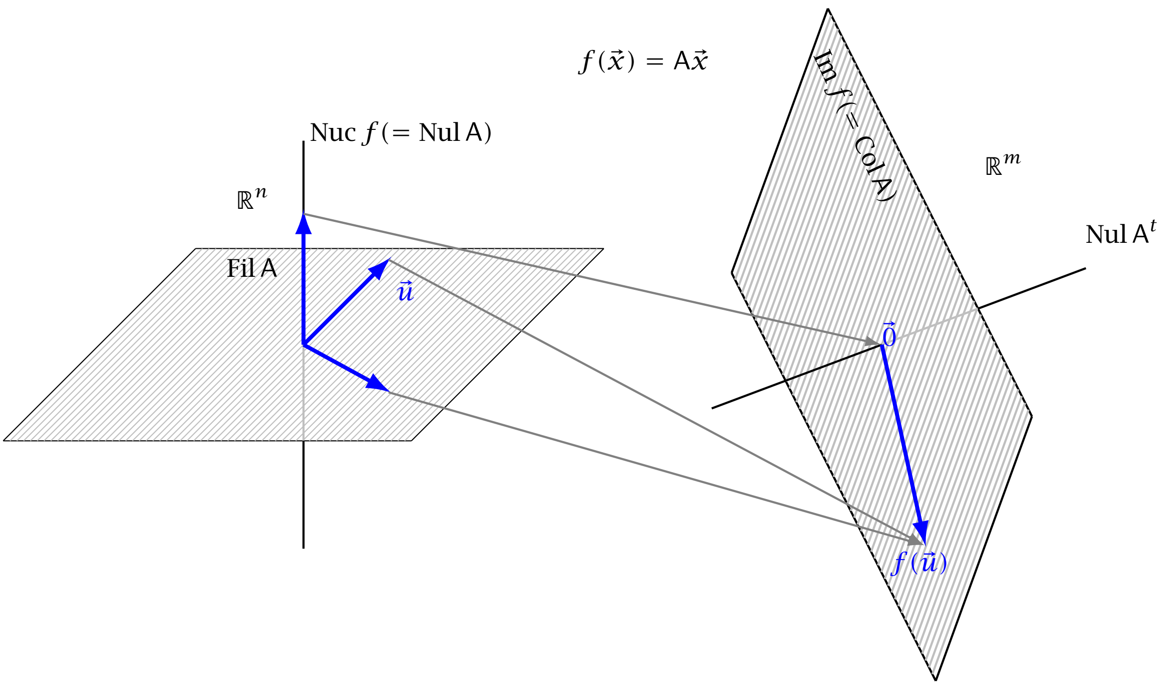
Exercici 10.4. Donada la matriu $\mathsf{A} = \begin{bmatrix} 1 & 3 & 2 \\ 1 & 3 & 0 \\ 3 & 3 & 2 \end{bmatrix}$ (a) Proveu que no és possible trobar una factorització LU estricta de A. (b) Trobeu una factorització LU de A (òbviament, no estricta).

Exercici 10.5. Trobeu una factorització LU de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -2 & 1 \\ 0 & -2 & -2 \end{bmatrix}$ i feu servir aquesta factorització per resoldre el sistema lineal $\mathsf{A}\vec{x} = (0, -6, -2)$.

Exercici 10.6. Trobeu una forma esglaonada de la matriu

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

fent servir l'algorisme de Gauss amb pivotatge parcial.



Segona part
Espais vectorials i aplicacions lineals

Capítol 3

Subespais de $\mathbb{R}^n$

Continguts

11. Els espais $\mathbb{R}^n$

En aquesta lliçó comencem a parlar d'*espais* i expliquem perquè l'espai $\mathbb{R}^3$ és de dimensió 3. Per a fer això, explicarem què és una base i veurem que totes les bases de $\mathbb{R}^n$ tenen exactament n vectors.
Tot seguit estudiarem els problemes de trobar les coordenades i canviar de base.

11.1. Bases i dimensió dels espais $\mathbb{R}^n$

$\mathbb{R}^3$ és un espai *de dimensió* 3. Això vol dir que si elegim tres vectors linealment independents llavors qualsevol altre vector d'aquest espai és combinació lineal d'aquests tres vectors.

Aquesta propietat és òbvia si elegim els vectors $\vec{e}_1 = (1,0,0)$, $\vec{e}_2 = (0,1,0)$ i $\vec{e}_3 = (0,0,1)$, perquè qualsevol vector $\vec{u} = (u_1, u_2, u_3)$ el podem escriure com

$$\vec{u} = u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1)$$

Però també és certa si elegim qualsevol altre conjunt de tres vectors, sempre que siguen linealment independents: si el conjunt $B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ és linealment independent, llavors el rang de la matriu

$$\mathsf{A} = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$$

és 3 i, en conseqüència, el sistema lineal $\mathsf{A}\vec{x} = \vec{u}$ té solució (única) per a qualsevol vector $\vec{u}$; en altres paraules, qualsevol vector de $\mathbb{R}^3$ és combinació lineal dels vectors de B.

Definició 11.1.

El conjunt B és una base de l'espai $\mathbb{R}^n$ si es compleixen les dues condicions següents:

(a) El conjunt B és linealment independent

(b) Si $\vec{u}$ és un vector de $\mathbb{R}^n$, llavors $\vec{u}$ és una combinació lineal dels vectors de B.

Quan un conjunt de vectors S compleix la segona condició d'aquesta definició, és a dir, quan qualsevol vector de $\mathbb{R}^n$ és combinació lineal dels vectors de S, diem que S *genera* $\mathbb{R}^n$ (o, que S *és generador* de $\mathbb{R}^n$). Un conjunt de vectors pot ser generador però no linealment independent, o ser linealment independent però no generador. Perquè siga una base cal que es complisquen totes dues condicions.

Ara discutirem *quants* vectors pot tenir el conjunt $S = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$, si és linealment independent i/o si és generador; d'alguna manera, *molts* vectors no poden ser independents i *pocs* vectors no poden generar l'espai.

Al conjunt S li associem la matriu $n \times p$

$$\mathsf{M}_S = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_p \end{bmatrix}$$

Aleshores, S és linealment independent si i només si el sistema $\mathsf{M}_S\vec{x} = \vec{0}$ és determinat; per tant, la condició necessària i suficient perquè S siga efectivament linealment independent és que el rang de M_S siga exactament p:

Si el conjunt S té exactament p vectors, llavors S és linealment independent si i només si rang $\mathsf{M}_S = p$.

D'altra banda, S genera $\mathbb{R}^n$ si l'equació $\mathsf{M}_S\vec{x} = \vec{b}$ és compatible per a qualsevol vector $\vec{b}$. En particular, si S és generador, totes les equacions

$$\mathsf{M}_S\vec{x} = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad \mathsf{M}_S\vec{x} = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathsf{M}_S\vec{x} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}$$

seran compatibles. Dit d'altra manera, l'equació matricial

$$\mathsf{M}_S\mathsf{X} = \mathrm{I}$$

és compatible, la qual cosa implica que el rang de M_S és n.

El conjunt S és generador si i només si $\operatorname{rang} \mathsf{M}_S = n$.

El teorema següent resumeix les condicions que han de complir els conjunts que són independents, generadors o bases:

Teorema 11.1.

Siga $S = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ un conjunt de vectors de $\mathbb{R}^n$.

(a) S és linealment independent si i només si $\operatorname{rang} \mathsf{M}_S = p$

(b) S genera l'espai $\mathbb{R}^n$ si i només $\operatorname{rang} \mathsf{M}_S = n$

En conseqüència,

(c) S és una base de $\mathbb{R}^n$ si i només si $\operatorname{rang} \mathsf{M}_S = n = p$. □

☞ Ben llegida, la condició (c) d'aquest teorema diu que S és una base de $\mathbb{R}^n$ si i només si la matriu M_S és invertible.[1]

Aquest teorema té moltes conseqüències importants:

(a) Un conjunt linealment independent no pot tenir més de n vectors.

(b) Un conjunt que siga generador de $\mathbb{R}^n$ no pot tenir menys de n vectors.

(c) Totes les bases de $\mathbb{R}^n$ tenen exactament n vectors (és a dir, que les bases de $\mathbb{R}^2$ tenen dos elements, les de $\mathbb{R}^3$ en tenen tres i les de $\mathbb{R}^7$ en tenen 7).

☞ Com que totes les bases de l'espai $\mathbb{R}^n$ tenen n elements, diem que *la dimensió de $\mathbb{R}^n$ és n*.

Definició 11.2. (Base canònica de $\mathbb{R}^n$)

La base canònica (o estàndard) de $\mathbb{R}^n$ és el conjunt

$$\mathscr{E}_n = \{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}$$

És clar que es tracta d'una base, perquè la matriu $\mathsf{M}_{\mathscr{E}_n}$ és la identitat (i perquè és clar que aquest conjunt és linealment independent i generador).

11.2. Coordenades

Per què imposem la condició que els vectors siguen independents, quan definim les bases? Si el conjunt S genera l'espai, llavors qualsevol vector es pot construir fent combinacions lineals amb els vectors de S, siga o no, aquest conjunt, linealment independent. Per exemple, el conjunts

$$B = \{(1, 1, 1), (1, 1, 0), (1, 0, 1)\} \qquad S = \{(1, 0, 1), (0, 1, 1), (1, 1, 1), (1 - 1, 1)\}$$

són, tots dos, generadors de $\mathbb{R}^3$, perquè $\operatorname{rang} M_B = \operatorname{rang} M_S = 3$. Aleshores, quin avantatge té B sobre S? Ho veurem clar quan expressem un vector qualsevol com a combinació lineal dels elements de B i de S i comparem els resultats.

Exemple 11.1.

Expresseu el vector $(3, 2, 1)$ com a combinació lineal dels elements de

$$B = \{(1, 1, 1), (1, 1, 0), (1, 0, 1)\}$$

i com a combinació lineals dels de

$$S = \{(1, 0, 1), (0, 1, 1), (1, 1, 1), (1 - 1, 1)\}$$

[1] La qual cosa afegeix una nova equivalència al teorema de caracterització de les matrius invertibles.

Simplement hem de resoldre els dos sistemes lineals les matrius ampliades dels quals són, respectivament,

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

Si calculem les formes esglaonades reduïdes obtindrem

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & -1 & 4 \end{bmatrix}$$

així que la solució del primer sistema és $\vec{x} = (0, 2, 1)$ i

$$(3, 2, 1) = 0(1, 1, 1) + 2(1, 1, 0) + 1(1, 0, 1)$$

la solució del segon sistema és $\vec{x} = (-1, -2, 4, 0) + \alpha(-2, 0, 1, 1)$, de manera que

$$(3, 2, 1) = (-1 - 2\alpha)(1, 0, 1) - 2(0, 1, 1) + (4 + \alpha)(1, 1, 1) + \alpha(1, -1, 1)$$

El fet que S no siga linealment independent fa que la solució del problema no siga única! El vector es pot escriure de moltes maneres diferents com a combinació lineal dels vectors de S. El que passa, en realitat, és que, *en S, hi sobren vectors*; podem generar tots els vectors de $\mathbb{R}^3$ amb, només, els tres primers vectors de S.

En definitiva, la independència lineal assegura que els vectors de l'espai es componen com a combinacions lineals dels elements de la base *de forma única*.

PROPIETAT 11.2.

Si B és una base de $\mathbb{R}^n$, llavors qualsevol vector de $\mathbb{R}^n$ s'expressa de forma única com a combinació lineal dels vectors de B. □

DEFINICIÓ 11.3. (COORDENADES D'UN VECTOR)

Si $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ és una base de $\mathbb{R}^n$ i el vector $\vec{u}$ es pot escriure com

$$\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_p\vec{u}_n$$

llavors, els pesos $\alpha_1, \alpha_2, \dots, \alpha_n$ d'aquesta combinació lineal són les coordenades del vector $\vec{u}$ respecte a la base B, la qual cosa s'expressa escrivint

$$\vec{u}_B = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

☞ La relació $\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_p\vec{u}_n$ significa que $\vec{u} = \mathsf{M}_B\vec{u}_B$.

Vegem-ne un exemple: adés hem vist que el conjunt $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 1)\}$ és una base de $\mathbb{R}^3$ i que el vector $\vec{u} = (3, 2, 1)$ es pot expressar d'aquesta manera:

$$(3, 2, 1) = 0(1, 1, 1) + 2(1, 1, 0) + 1(1, 0, 1)$$

Així, doncs, les coordenades de $\vec{u}$ respecte a la base B són 0, 2 i 1, és a dir,

$$\vec{u}_B = (0, 2, 1)$$

EXEMPLE 11.2.

Quines són les coordenades del vector $\vec{u} = (u_1, u_2, \dots, u_n)$ respecte a la base canònica $\mathcal{E}_n$?

Com que

$$\vec{u} = (u_1, u_2, \dots, u_n) = u_1(1, 0, \dots, 0) + u_2(0, 1, \dots, 0) + \cdots + u_n(0, 0, \dots, 1)$$

les coordenades respecte a la base canònica són les components del vector:

$$\vec{u}_{\mathcal{E}} = (u_1, u_2, \dots, u_n) \quad \square$$

EXEMPLE 11.3.

> Proveu que el conjunt $B = \{(1,2),(2,-1)\}$ és una base de $\mathbb{R}^2$ i trobeu les coordenades del vector $\vec{u} = (0,5)$ respecte aquesta base.

Aquest conjunt és base de $\mathbb{R}^2$ perquè $\operatorname{rang} \mathsf{M}_B = 2$. Les coordenades del vector $\vec{u}$ les podem obtenir resolen el sistema lineal $\mathsf{M}_B\vec{x} = \vec{u}$, la matriu ampliada del qual és $\begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 5 \end{bmatrix}$; aplicant-hi l'algorisme de Gauss-Jordan obtenim la forma esglaonada reduïda $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}$, així que $\vec{u}_B = (2,-1)$ (és a dir, $\vec{u} = 2(1,2) - 1(2,-1)$). □

11.3. LA MATRIU DE CANVI DE BASE

Si $B_1 = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ i $B_2 = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ són bases de $\mathbb{R}^n$ i $\vec{u}$ és un vector qualsevol, quina relació hi ha entre les coordenades de $\vec{u}$ respecte a cadascuna d'aquestes bases? En altres paraules, si $\vec{u}_{B_1} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ i $\vec{u}_{B_2} = (\beta_1, \beta_2, \ldots, \beta_n)$, quina relació hi ha entre unes i altres coordenades?

La resposta és òbvia: el que sabem és que

$$\vec{u} = \beta_1\vec{v}_1 + \beta_2\vec{v}_2 + \cdots + \beta_n\vec{v}_n = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n$$

és a dir,

$$\mathsf{M}_{B_2}\vec{u}_{B_2} = \mathsf{M}_{B_1}\vec{u}_{B_1}$$

Per tant,

$$\vec{u}_{B_2} = \left(\mathsf{M}_{B_2}\right)^{-1}\mathsf{M}_{B_1}\vec{u}_{B_1}$$

DEFINICIÓ 11.4.

> *Si B_1 i B_2 són dues bases de $\mathbb{R}^n$, la matriu del canvi de base de B_1 a B_2 és la matriu* $\mathsf{M}_{B_1B_2} = \left(\mathsf{M}_{B_2}\right)^{-1}\mathsf{M}_{B_1}$.

☞ Observeu que, quan la base B_2 és la canònica, llavors la matriu de canvi de basa coincideix amb la matriu M_{B_1} ($\mathsf{M}_{B_1\mathscr{E}_n} = \mathsf{M}_{B_1}$), perquè la matriu associada a la base canònica és la identitat.

EXEMPLE 11.4.

> Trobeu la matriu de canvi de la base de $\mathbb{R}^2$ $B_1 = \{(1,1),(1,-1)\}$ a la base $B_2 = \{((0,1),(2,3)\}$ i les coordenades respecte a la base B_2 del vector $\vec{u}$ sabent que $\vec{u}_{B_1} = (1,1)$.

La manera més eficient de calcular la matriu $\mathsf{M}_{B_1B_2} = \left(\mathsf{M}_{B_2}\right)^{-1}\mathsf{M}_{B_1}$ consisteix a aplicar l'algorisme de Gauss-Jordan a la matriu $\left[\mathsf{M}_{B_2} \mid \mathsf{M}_{B_1}\right]$, perquè $\left(\mathsf{M}_{B_2}\right)^{-1}\left[\mathsf{M}_{B_2} \mid \mathsf{M}_{B_1}\right] = \left[\mathrm{I} \,\middle|\, \left(\mathsf{M}_{B_2}\right)^{-1}\mathsf{M}_{B_1}\right]$:

$$\begin{bmatrix} 0 & 2 & 1 & 1 \\ 1 & 3 & 1 & -1 \end{bmatrix} \xrightarrow{\text{Gauss-Jordan}} \begin{bmatrix} 1 & 0 & -1/2 & -5/2 \\ 0 & 1 & 1/2 & 1/2 \end{bmatrix}$$

La matriu de canvi de la base B_1 a la base B_2 és $\mathsf{M}_{B_1B_2} = \dfrac{1}{2}\begin{bmatrix} -1 & -5 \\ 1 & 1 \end{bmatrix}$. I les coordenades de $\vec{u}$ respecte a la base B_2,

$$\vec{u}_{B_2} = \mathsf{M}_{B_1B_2}\vec{u}_2 = \frac{1}{2}\begin{bmatrix} -1 & -5 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

11.4. BASES ORTOGONALS I ORTONORMALS

Si el conjunt $B = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$, a més de ser base de $\mathbb{R}^n$, és ortogonal, llavors el càlcul de les coordenades d'un vector respecte aquesta base se simplifica notablement: en comptes de resoldre un sistema lineal, únicament haurem de fer uns pocs productes escalars. Per veure-ho, suposem que

$\vec{u}_B = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, és a dir, que $\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n$, llavors els productes escalars de $\vec{u}$ pels elements de B són

$$\begin{aligned}
\vec{u}\cdot\vec{u}_1 &= \alpha_1\vec{u}_1\cdot\vec{u}_1 + \underbrace{\alpha_2\,\vec{u}_2\cdot\vec{u}_1}_{0} + \cdots + \underbrace{\alpha_n\,\vec{u}_n\cdot\vec{u}_1}_{0} = \alpha_1\vec{u}_1\cdot\vec{u}_1\\
\vec{u}\cdot\vec{u}_2 &= \underbrace{\alpha_1\,\vec{u}_1\cdot\vec{u}_2}_{0} + \alpha_2\vec{u}_2\cdot\vec{u}_2 + \cdots + \underbrace{\alpha_n\,\vec{u}_n\cdot\vec{u}_2}_{0} = \alpha_2\vec{u}_2\cdot\vec{u}_2\\
&\ldots\\
\vec{u}\cdot\vec{u}_n &= \underbrace{\alpha_1\,\vec{u}_1\cdot\vec{u}_n}_{0} + \underbrace{\alpha_2\,\vec{u}_2\cdot\vec{u}_n}_{0} + \cdots + \alpha_n\vec{u}_n\cdot\vec{u}_n = \alpha_n\vec{u}_n\cdot\vec{u}_n
\end{aligned}$$

En conseqüència,

$$\alpha_1 = \frac{\vec{u}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}, \quad \alpha_2 = \frac{\vec{u}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}, \quad \ldots, \quad \alpha_n = \frac{\vec{u}\cdot\vec{u}_n}{\vec{u}_n\cdot\vec{u}_n}$$

☞ El vector de coordenades de $\vec{u}$ respecte la base B és

$$\boxed{\vec{u}_B = \left(\frac{\vec{u}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}, \frac{\vec{u}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}, \ldots, \frac{\vec{u}\cdot\vec{u}_n}{\vec{u}_n\cdot\vec{u}_n}\right)}$$

Exemple 11.5.

Trobeu les coordenades del vector $\vec{u} = (0,5)$ respecte la base $B = \{(1,2),(2,-1)\}$.

Es tracta de la mateixa base (i el mateix vector $\vec{u}$) de l'exemple 11.3., però ara observem que la base B és ortogonal (perquè $(1,2)\cdot(2,-1) = 0$), així que

$$\vec{u}_B = \left(\frac{(0,5)\cdot(1,2)}{(1,2)\cdot(1,2)}, \frac{(0,5)\cdot(2,-1)}{(2,-1)\cdot(2,-1)}\right) = \left(\frac{10}{5}, \frac{-5}{5}\right) = (2,-1)$$

☞ Si la base $B = \{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$ és ortonormal, llavors tots els vectors seus són unitaris i el càlcul de les coordenades encara se simplifica una mica més:

$$\boxed{\vec{u}_B = (\vec{u}\cdot\vec{q}_1, \vec{u}\cdot\vec{q}_2, \ldots, \vec{u}\cdot\vec{q}_n) = \begin{bmatrix}\vec{q}_1^{\,t}\\ \vec{q}_2^{\,t}\\ \cdots\\ \vec{q}_n^{\,t}\end{bmatrix}\vec{u} = \mathsf{M}_B^t\vec{u}}$$

Al marge de la simplicitat en el càlcul de les coordenades, les bases ortonormals tenen el valor afegit de respectar les propietats geomètriques dels vectors: si B és una base ortonormal i $\vec{u}$, i $\vec{v}$ són dos vectors qualssevol, llavors $\vec{u}\cdot\vec{v} = \vec{u}_B\cdot\vec{v}_B$, perquè

$$\vec{u}\cdot\vec{v} = \mathsf{M}_B^t\vec{u}_B\cdot\mathsf{M}_B^t\vec{v}_B = \vec{u}_B^{\,t}\mathsf{M}_B\mathsf{M}_B^t\vec{v}_B = \vec{u}_B^{\,t}\vec{v}_B = \vec{u}_B\cdot\vec{v}_B$$

($\mathsf{M}_B\mathsf{M}_B^t = \mathrm{I}$ perquè la matriu M_B és ortogonal) i d'ací es dedueix que, si B és una base ortonormal, llavors podem calcular longituds, angles i distàncies fent servir el vector de coordenades respecte a la base B.

11.5. Exercicis

Exercici 11.1. Feu servir la definició de base per a determinar si els conjunts següents són o no bases de l'espai $\mathbb{R}^2$.

(a) $B_1 = \{(0,-4),(1,0)\}$ (b) $B_2 = \{(1,-1),(-2,2)\}$ (c) $B_3 = \{(1,-1),(2,1),(1,1)\}$

Exercici 11.2. Justifiqueu si els conjunts següents són o no linealment independents, generadors o bases de $\mathbb{R}^3$.

(a) $S_1 = \{(1,1,1),(0,1,1),(0,0,1)\}$ (b) $S_2 = \{(1,1,1),(0,1,1),(2,5,5)\}$
(c) $S_3 = \{(1,1,1),(0,1,1),(0,0,1),(2,5,5)\}$ (d) $S_4 = \{(1,1,1),(2,5,5)\}$

EXERCICI 11.3. Vertader o fals (justifiqueu la resposta):
(a) Si S és un conjunt de tres vectors en $\mathbb{R}^4$ llavors S és linealment independent.
(b) Si S és un conjunt de cinc vectors en $\mathbb{R}^4$ llavors S és linealment dependent.
(c) Si S és un conjunt de tres vectors en $\mathbb{R}^4$ llavors S no és generador.
(d) Si S és un conjunt de cinc vectors en $\mathbb{R}^4$ llavors S és generador.
(e) Si S és un conjunt de quatre vectors en $\mathbb{R}^4$ llavors S és base.
(f) Si S és un conjunt linealment independent de quatre vectors en $\mathbb{R}^4$ llavors S és base.
(g) Si S és un conjunt de quatre vectors en $\mathbb{R}^4$ i S genera $\mathbb{R}^4$, llavors S és base.

EXERCICI 11.4. (a) Proveu que els conjunts $B_1 = \{(1,1), (-1,2)\}$ i $B_2 = \{(1,1), (0,1)\}$ són bases de $\mathbb{R}^2$. (b) Calculeu els vectors de coordenades ($\vec{u}_{B_1}$ i $\vec{u}_{B_2}$) del vector $\vec{u} = (1,7)$ respecte a cada una d'aquestes bases. (c) Trobeu el vector $\vec{v}$ sabent que les coordenades d'aquest vector respecte a la base B_1 són $\vec{v}_{B_1} = (2,2)$.

EXERCICI 11.5. Trobeu les matrius de canvi de base $\mathsf{M}_{B_1B_2}$ i $\mathsf{M}_{B_2B_1}$ on B_1 i B_2 són les bases de $\mathbb{R}^2$ de l'exercici anterior.

EXERCICI 11.6. Calculeu les coordenades del vector $\vec{u} = (1,2,-1,-2)$ respecte a la base

$$B = \left\{\frac{1}{2}(1,-1,1,-1), \frac{1}{2}(1,-1,-1,1), \frac{1}{2}(1,1,-1,-1), \frac{1}{2}(1,1,1,1)\right\}$$

12. Subespais de $\mathbb{R}^n$

Definició 12.1.

Un subespai de $\mathbb{R}^n$ és un conjunt F, de vectors n-dimensionals, que té aquestes dues propietats:

(a) El vector $\vec{0}$ és un element de F:

$$\vec{0} \in F$$

(b) Si fem qualsevol combinació lineal amb dos elements de F el vector que en resulta també és un element de F:

$$\text{Si } u_1, u_2 \in F \text{ llavors } \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 \in F$$

(per a dos escalars α_1 i α_2 qualssevol).

Primer de tot estudiarem quins conjunts són subespais de $\mathbb{R}^2$ i de $\mathbb{R}^3$ i quins no ho són.

12.1. Subespais de $\mathbb{R}^2$

Per començar amb un exemple senzill, observarem que el conjunt

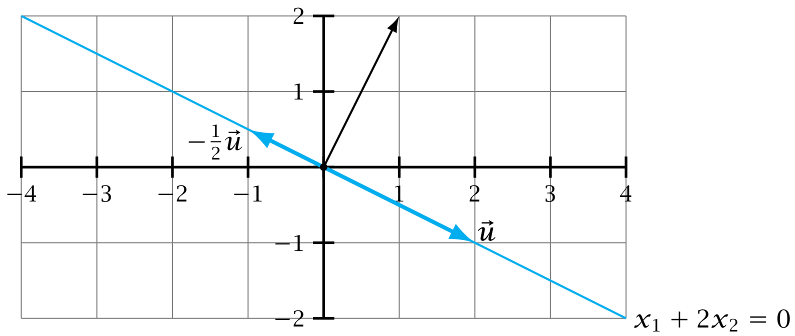
$$F = \{\alpha(2,-1) : \alpha \in \mathbb{R}\}$$

és un subespai de $\mathbb{R}^2$. Des del punt de vista geomètric, aquest conjunt F és la recta d'equació $x_1+2x_2 = 0$ o, amb més precisió,

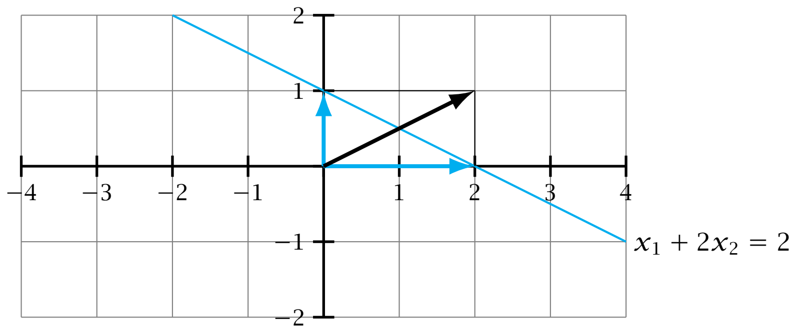
$$F = \{(x_1, x_2) : x_1 + 2x_2 = 0\}$$

F és subespai de $\mathbb{R}^2$ perquè qualsevol combinació lineal de dos vectors d'aquesta recta es troba sobre la mateixa recta. Com que tots els vectors de F són múltiples del vectors $\vec{u} = (2,-1)$ diem que aquest vector *genera* el subespai F.

Observeu que tots els vectors de F són ortogonals al vector $(1,2)$.



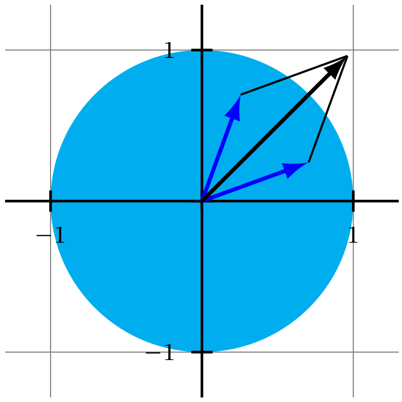
En canvi, el conjunt de les solucions de l'equació $x_1 + 2x_2 = 2$ no és un subespai de $\mathbb{R}^2$, perquè el vector zero no és solució de l'equació (la recta no conté l'origen de coordenades) i, també perquè aquest conjunt no compleix la segona condició per a ser subespai: la combinació de dos vectors de la recta no sempre es troba sobre aquesta recta.



Un altre subconjunt de $\mathbb{R}^2$ que tampoc no és subespai és el cercle

$$G = \{(x_1, x_2) \colon\ x_1^2 + x_2^2 \leq 1\}$$

Aquest cercle si que conté el vector zero, però no totes les combinacions lineals.



Un altre subespai és el que conté únicament el vector zero: $\{\vec{0}\}$, perquè qualsevol combinació lineal que hi fem serà nul·la:

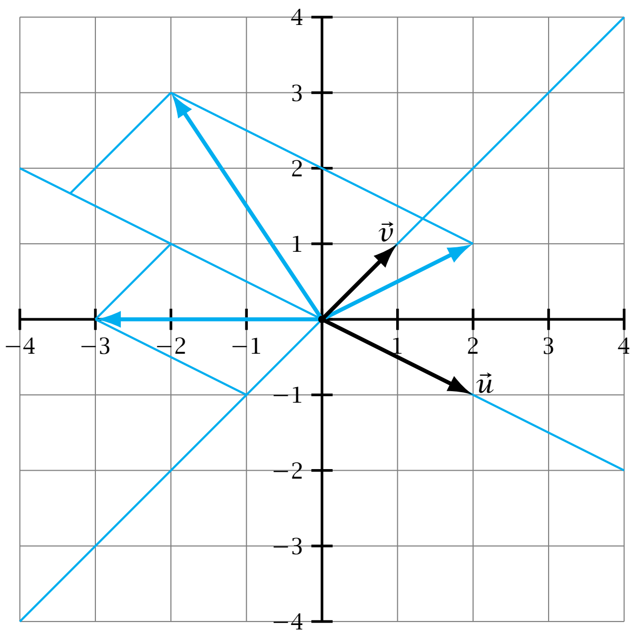
$$\alpha_1\vec{0} + \alpha_2\vec{0} = \vec{0}$$

Aquest és, evidentment, el subespai més petit de $\mathbb{R}^2$. Per contra, el *més gran* és el mateix $\mathbb{R}^2$!

Per a determinar *tots* els subespais de $\mathbb{R}^2$ només ens hem de fixar en aquest detall: si F conté algun vector $\vec{u}$ no nul, llavors qualsevol vector que siga múltiple de $\vec{u}$, $\alpha\vec{u}$, també ha de ser un element de F, de manera que si F no es redueix al vector zero, llavors F conté, almenys, la recta definida per $\vec{u}$:

$$\vec{u} \in F \Rightarrow \{\alpha\vec{u} \colon\ \alpha \in \mathbb{R}\} \subset F$$

Si tots els vectors de F són múltiples de $\vec{u}$, llavors, F és aquesta recta. En cas contrari, és a dir, si hi ha algun altre vector en F, $\vec{v}$, que no és múltiple de $\vec{u}$, aleshores qualsevol vector que siga combinació lineal de $\vec{u}$ i $\vec{v}$, $\alpha_1\vec{u} + \alpha_2\vec{v}$, serà també un element de F. En aquest cas, F coincideix amb $\mathbb{R}^2$, perquè la matriu $\begin{bmatrix}\vec{u} & \vec{v}\end{bmatrix}$ té rang 2 i, en conseqüència, qualsevol vector és combinació lineal de $\vec{u}$ i $\vec{v}$.



En resum,

Els únics subespais vectorials de $\mathbb{R}^2$ són aquests:

- El subespai zero, $\{\vec{0}\}$
- Si F no és nul i no conté dos vectors linealment independents, llavors, F és una recta que passa per l'origen. Si $\vec{u} = (a, b)$ és un vector no nul d'aquest subespai F, llavors

$$F = \{\alpha(a, b) : \alpha \in \mathbb{R}\} = \langle (a, b) \rangle$$

  L'equació d'aquesta recta és $-bx_1 + ax_2 = 0$, així que també podem escriure el subespai F com

$$F = \{(x_1, x_2) : -bx_1 + ax_2 = 0\}$$

  i, finalment, hem d'observar que F és el conjunt dels vectors ortogonals al vector $(-b, a)$.
- Si F conté dos vectors linealment independents, llavors F coincideix amb $\mathbb{R}^2$.

☞ La recta d'equació $ax_1 + bx_2 + c = 0$ només és un subespai si passa per l'origen, és a dir, si $c = 0$.

12.2. Subespais de $\mathbb{R}^3$

En $\mathbb{R}^3$, a banda dels subespais òbvis, $\{\vec{0}\}$ i el mateix $\mathbb{R}^3$, hi són subespais les rectes i els plans que contenen el zero. Com en el cas bidimensional, si el subespai F no es redueix al vector zero i no hi ha cap parell de vectors linealment independents, llavors elegint-hi un vector no nul $\vec{u}$ deduirem que els únics vectors de F són els múltiples d'aquest vector, i F és una recta que conté l'origen:

$$F = \{\alpha\vec{u} : \alpha \in \mathbb{R}\}$$

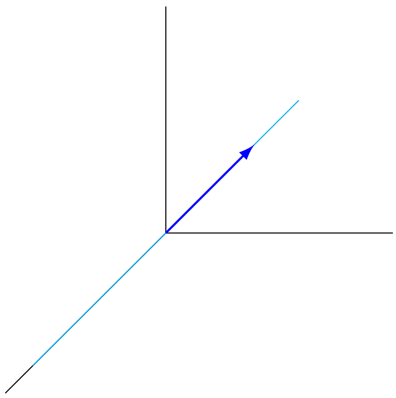
Per exemple, $F_1 = \{\alpha(1, 2, 1) : \alpha \in \mathbb{R}\}$ és un subespai de $\mathbb{R}^3$.

En canvi, si en F podem trobar dos vectors $\vec{u}$ i $\vec{v}$ linealment independents (però no tres), llavors tots els vectors de F seran combinació lineal de $\vec{u}$ i $\vec{v}$,

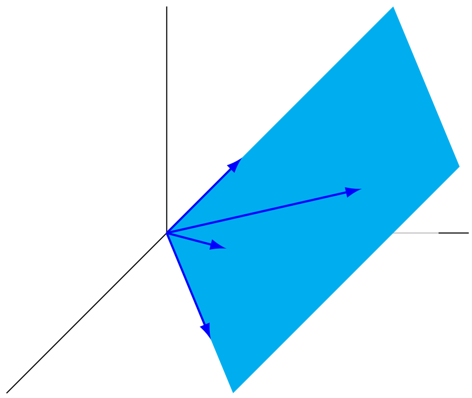
$$F = \{\alpha_1\vec{u} + \alpha_2\vec{v} : \alpha_1, \alpha_2 \in \mathbb{R}\}$$

Per exemple, $F_2 = \{\alpha_1(1, 2, 1) + \alpha_2(0, 1, 2) : \alpha_1, \alpha_2 \in \mathbb{R}\}$ és un subespai de $\mathbb{R}^3$.

Finalment, si F conté tres vectors independents, llavors $F = \mathbb{R}^3$.



Un únic vector independent: una recta



Dos vectors independents: un pla

12.3. Dimensió i bases dels subespais de $\mathbb{R}^n$

La definició de base es pot estendre a qualsevol subespai de $\mathbb{R}^n$.

Definició 12.2. (Base d'un subespai)

Siga F un subespai de $\mathbb{R}^n$. El conjunt B és una base de F si es compleixen les tres condicions següents:

(a) El conjunt B és linealment independent

(b) L'embolcall lineal de B és F.

Ara ja estem en condicions de descriure tots els subespais de $\mathbb{R}^n$; només hem de fer dues o tres precisions prèvies.

Propietats 12.1.

Siga F un subespai de $\mathbb{R}^n$. Es compleixen les propietats següents:

(a) Si F no és el subespai nul, aleshores existeix alguna base de F

(b) el nombre d'elements d'una base de F no és més gran que n.

Demostració: Si F no és el subespai nul, hi existirà algun vector, $\vec{u}_1$ distint del vector zero; si aquest vector per si sol no és base de F llavors hi ha algun altre vector de F, $\vec{u}_2$ que no és combinació lineal de $\vec{u}_1$, de manera que $\{\vec{u}_1, \vec{u}_2\}$ és linealment independent; si aquest conjunt de dos vectors tampoc no és base, hi podriem afegiu-ne un altre, i considerar el conjunt linealment independent $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.

Aquest procés ha d'acabar necessàriament en un màxim de n passos, perquè si arribem a construir un conjunt de n vectors linealment independents, aquests vectors generen l'espai complet $\mathbb{R}^n$. $\square$

☞ El subespai nul *no té cap base*, perquè no té subconjunts linealment independents.

La propietat més important és que totes les bases de F tenen el mateix nombre d'elements:

Propietat 12.2.

Si F és un subespai no nul de $\mathbb{R}^n$, llavors totes les bases de F tenen el mateix nombre d'elements.

Demostració: Suposem que $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ i $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$ són bases de F. El que volem provar és que $p = q$.

Per ser B_1 una base, tots els vectors de B_2 són combinacions lineals dels de B_1, de manera que l'equació $\mathsf{M}_{B_1}\mathsf{X} = \mathsf{M}_{B_2}$ és compatible. Per tant,

$$\operatorname{rang}\begin{bmatrix}\mathsf{M}_{B_1} & \mathsf{M}_{B_2}\end{bmatrix} = \operatorname{rang}\mathsf{M}_{B_1} = p$$

Però com que B_2 també és base, els vectors de B_1 són combinacions lineals dels de B_2, l'equació $\mathsf{M}_{B_2}\mathsf{X} = \mathsf{M}_{B_1}$ és compatible i

$$\operatorname{rang}\begin{bmatrix}\mathsf{M}_{B_2} & \mathsf{M}_{B_1}\end{bmatrix} = \operatorname{rang}\mathsf{M}_{B_2} = q$$

així que $p = q$. $\square$

Definició 12.3. (Dimensió d'un subespai)

Si F és el subespai nul, la dimensió de F és zero. En cas contrari, el nombre d'elements d'una base qualsevol de F és la dimensió d'aquest subespai, $\dim F$.

Tenint en compte aquestes propietats és clar que els subespais de $\mathbb{R}^n$ es poden classificar atenent a la dimensió de la manera següent: si F és un subespai de $\mathbb{R}^n$ llavors, f és, o bé el subespai nul, o bé un subespai de dimensió 1, o un de dimensió 2..., o un de dimensió $n-1$, o l'espai total $\mathbb{R}^n$. Els subespais de dimensió 1 s'anomenen *rectes*, els de dimensió 2, *plans*, i els de dimensió $n-1$, *hiperplans.*

☞ L'equació $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ representa un *hiperplà* en l'espai n-dimensional. Però aquest hiperplà només és un subespai si passa per l'origen, és a dir, si $b = 0$.

Per exemple, en $\mathbb{R}^4$ hi ha subespais de dimensió 1 (rectes), 2 (plans), i 3 (hiperplans), a banda dels subespais trivials $\{\vec{0}\}$ i $\mathbb{R}^4$.

DEFINICIÓ 12.4. (COORDENADES EN UN SUBESPAI)

Si $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ és una base del subespai F i el vector $\vec{u}$ es pot escriure com

$$\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_p\vec{u}_p$$

llavors, els pesos $\alpha_1, \alpha_2, \dots, \alpha_p$ d'aquesta combinació lineal són les coordenades del vector $\vec{u}$ respecte a la base B, la qual cosa s'expressa escrivint

$$\vec{u}_B = (\alpha_1, \alpha_2, \dots, \alpha_p)$$

☞ El vector $\vec{u}$ té n components (perquè és un element de $\mathbb{R}^n$) però p coordenades respecte a la base B, perquè la dimensió del subespai F és p.

☞ La relació $\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_p\vec{u}_p$ significa que $\vec{u} = \mathsf{M}_B\vec{u}_B$.

EXEMPLE 12.1.

Comproveu que el conjunt $B = \{(1,0,0), (0,1,-1)\}$ és una base del subespai de $\mathbb{R}^3$, $F = \{(x_1, x_2, x_3)\ :\ x_2 + x_3 = 0\}$ i trobeu les coordenades respecte aquesta base del vector $\vec{w} = (2, -3, 3)$.

Anomenem $\vec{u} = (1,0,0)$ i $\vec{v} = (0,1,-1)$ els vectors de B. Hem de comprovar tres coses:

(a) B és un subconjunt de F, perquè $u_2 + u_3 = 0 + 0 = 0$ i $v_2 + v_3 = 1 - 1 = 0$.

(b) Qualsevol vector de F és combinació lineal del vectors de B:

$$\text{Si } y + z = 0 \text{ llavors } (x, y, z) = (x, y, -y) = x(1,0,0) + y(0,1,-1)$$

(c) El conjunt B és linealment independent: Si $\alpha_1\vec{u} + \alpha_2\vec{v} = \vec{v}$ llavors $\alpha_1(1,0,0) + \alpha_2(0,1,-1) = (0,0,0)$ així que $(\alpha_1, \alpha_2, -\alpha_2) = (0,0,0)$ i $\alpha_1 = \alpha_2 = 0$.

Les coordenades respecte aquesta base del vector $\vec{w}$ són 2 i -3, perquè

$$\vec{w} = (2, -3, 3) = 2(1,0,0) - 3(0,1,-1)$$

Simbòlicament escriurem $\vec{w}_B = (2, -3)$.

☞ El vector $\vec{w}$ d'aquest exemple és tridimensional, però el vector de coordenades $\vec{w}_B$ només té dos components, perquè estem en un subespai de dimensió 2.

12.4. EXERCICIS

12.4.1. SUBESPAIS

EXERCICI 12.1. Estudieu si els conjunts de vectors següents són subespais de l'espai $\mathbb{R}^2$ o $\mathbb{R}^3$ o si no ho són. Descriviu geomètricament aquests conjunts.

(a) $F_1 = \{(a, 0)\ :\ a \in \mathbb{R}\}$

(b) $F_2 = \{(a - b, 2a + b)\ :\ a, b \in \mathbb{R}\}$

(c) $F_3 = \{(a - b, b - c, c - a)\ :\ a, b, c \in \mathbb{R}\}$

(d) $F_4 = \{(a + 1, a, 0)\ :\ a \in \mathbb{R}\}$

(e) $F_5 = \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_1 - x_2 = 0, x_1 - x_3 = 1\}$

(f) $F_6 = \{(a, b)\ :\ a, b \in \mathbb{R}, a \leq b\}$

(g) $F_7 = \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_1 - x_2 = 0, x_1 - x_3 = 0\}$

(h) $F_8 = \{(1, 2, 0), (0, -1, 1)\}$

EXERCICI 12.2. Proveu que el conjunt $F = \{(2a, 3a)\ :\ a \in \mathbb{R}\}$ és un subespai de $\mathbb{R}^2$, representeu-lo gràficament i descriviu-lo (1) com el conjunt de totes les combinacions lineals d'un vector, (2) com la solució general d'una equació lineal, i (3) com el conjunt de tots els vectors ortogonals a un vector.

EXERCICI 12.3. Proveu que els conjunts següents són subespais de $\mathbb{R}^3$ i descriviu-los (1) com el conjunt de totes les combinacions lineals d'un conjunt de vectors, (2) com la solució general d'un sistema d'equacions lineals, i (3) com el conjunt de tots els vectors ortogonals a un conjunt de vectors.

(a) $F_1 = \{(a, 2a, 3a)\ :\ a \in \mathbb{R}\}$

(b) $F_2 = \{(a, a + b, b)\ :\ a, b \in \mathbb{R}\}$

12.4.2. Bases i coordenades

Exercici 12.4. Trobeu una base de cadascun dels subespais de $\mathbb{R}^4$ següents

(a) $F = \{(a-b, b-c, c-d, d-a) \colon a, b, c, d \in \mathbb{R}\}$
(b) $G = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \colon x_1 - x_2 = 0; x_3 + x_4 = 0\}$

12.4.3. Embolcalls lineals

Exercici 12.5. Trobeu bases dels subespais següents:

(a) $\langle (1,1,0), (2,-1,1), (3,0,1) \rangle$ (b) $\langle (1,1,0), (2,-1,1), (3,0,0) \rangle$

Exercici 12.6. Siga S un subconjunt no buit de $\mathbb{R}^n$. Proveu que

(a) $\langle S \rangle$ és un subespai de $\mathbb{R}^n$

(b) si F és un subespai de $\mathbb{R}^n$ i $S \subset F$, llavors $\langle S \rangle \subset F$.

Exercici 12.7. És evident que, si el vector $\vec{u}$ no és un element de S, llavors $\langle S \rangle \subset \langle S \cup \{\vec{u}\} \rangle$. Poden ser iguals, $\langle S \rangle$ i $\langle S \cup \{\vec{u}\} \rangle$? En cas afirmatiu, quina condició s'ha de complir perquè no ho siguen?

Exercici 12.8. (Canvi de base en un subespai)
(a) Proveu que els dos conjunts $B_1 = \{(1,1,1), (1,-1,0)\}$ i $B_2 = \{(2,0,1), (0,2,1)\}$ són bases del subespai de $\mathbb{R}^3$ $F = \{(x+y, x-y, x) \colon x, y \in \mathbb{R}\}$. (b) Comproveu que el vector $\vec{u} = (3,1,2)$ és un element de F i calculeu les coordenades de $\vec{u}$ respecte a cadascuna d'aquestes bases. (c) Trobeu una matriu $\mathsf{M}_{B_1B_2}$ de manera que, per a qualsevol vector $\vec{v} \in F$, $\vec{v}_{B_2} = \mathsf{M}_{B_1B_2}\vec{v}_{B_1}$. d) Comproveu amb el vector $\vec{u}$ el resultat que heu obtingut.

13. Els quatre subespais deduïts d'una matriu

Qualsevol matriu $m \times n$ A defineix quatre subespais importants, dos de $\mathbb{R}^n$ i altres dos de $\mathbb{R}^m$, formats per les combinacions lineals de les columnes de A (l'espai columna), per les de les files (l'espai fila), i per les solucions dels sistemes lineals $\mathsf{A}\vec{x} = \vec{0}$ (l'espai nul) i $\mathsf{A}^t\vec{x} = \vec{0}$ (l'espai nul esquerre).
Per mostrar l'interès dels quatre subespais, demostrarem que qualsevol subespai de $\mathbb{R}^n$ és l'espai columna d'una matriu i, també, l'espai nul d'una altra matriu. En altres paraules, qualsevol subespai es pot interpretar com un dels quatre subespais, si s'elegeix la matriu adequada.

13.1. Els quatre subespais

Una matriu $m \times n$ (o els sistemes lineals associats a aquesta matriu) defineix quatre subspais vectorials especials: dos subespais de $\mathbb{R}^m$ i altres dos de $\mathbb{R}^n$. El rang de la matriu determina les dimensions dels quatre subespais.

13.1.1. L'espai columna

Fins ara, hem estudiat els sistemes del tipus $\mathsf{A}\vec{x} = \vec{b}$ considerant $\vec{b}$ com un vector donat i mirant de decidir si el sistema té alguna solució i, en tal cas, de trobar aquesta solució. Ara ens ho mirarem des del costat de la matriu A i ens demanarem quins són els vectors $\vec{b}$ per als quals el sistema té alguna solució.

Això defineix el subespai columna de la matriu.

Si A és una matriu $m \times n$, quins són els vectors $\vec{b}$ per als quals el sistema d'equacions lineals

$$\mathsf{A}\vec{x} = \vec{b}$$

és compatible?

Amb tot el que ja sabem a prop de les matrius i els sistemes lineals, la resposta és immediata, especialment si interpretem el producte $\mathsf{A}\vec{x}$ de la manera correcta: com que la pregunta és si existeix $(x_1, x_2, \ldots, x_n)$ de manera que $x_1\mathsf{A}_1 + x_2\mathsf{A}_2 + \cdots + x_n\mathsf{A}_n = \vec{b}$, podem afirmar que el sistema té solució si i només si *el vector $\vec{b}$ és combinació lineal de les columnes de* A.

Definició 13.1.

Anomenem espai columna de la matriu A*,* Col A *al conjunt de totes les combinacions lineals de les columnes de* A*.*

☞ L'espai columna és l'embolcall lineal de les columnes de la matriu i, també, el conjunt de vectors per als quals el sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és compatible.

L'espai columna és un subespai de $\mathbb{R}^m$ (noteu que les columnes de la matriu A tenen m components); la dimensió d'aquest subespai és el rang de A (recordem que el rang és precisament el nombre de columnes linealment independents).

$$\boxed{\dim \operatorname{Col} \mathsf{A} = \operatorname{rang} \mathsf{A}}$$

Si les columnes de A són linealment independents, llavors són una base de l'espai columna. Però si no ho són, caldrà *eliminar* les columnes que calga fins que en quede un subconjunt independent. Això es pot fer fàcilment a partir de la forma esglaonada reduïda (o qualsevol forma esglaonada) de la matriu: si S és una forma esglaonada de la matriu A, llavors les columnes de A que corresponen a les columnes principals de S formen una base de l'espai columna.

Exemple 13.1.

Trobeu una base de l'espai columna de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ -1 & 2 & 1 & 2 & 0 \\ 1 & -2 & 2 & 1 & 0 \\ 4 & -8 & 0 & -4 & 1 \end{bmatrix}$$

La forma esglaonada reduïda d'aquesta matriu és

$$\mathsf{R} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Com que les columnes principals de la la matriu R són la primera, la tercera i la cinquena, les columnes corresponents de la matriu A són una base de l'espai columna de A:

$$\boxed{B_{\mathrm{Col}\,A} = \{(1,-1,1,4),(0,1,2,0),(0,0,0,1)\}}$$

Observeu que les columnes segona i quarta de R indiquen que la segona columna de A és -2 vegades la primera i que la quarta columna de A és la tercera menys la primera,

$$\mathsf{A}_2 = -2\mathsf{A}_1, \qquad \mathsf{A}_4 = -\mathsf{A}_1 + \mathsf{A}_3$$

per això, aquestes columnes són innecessàries. □

☞ L'espai columna de R *no és el mateix* que l'espai columna de A: les combinacions lineals de les files de R són les mateixes que les de les files de A, però amb les columnes no passa el mateix. Les columnes de una forma esglaonada de A ens indiquen quines columnes de A convé elegir per a trobar una base, però *no són* la base.

13.1.2. L'espai nul

L'espai nul de A també és un subespai vectorial (en aquest cas, de $\mathbb{R}^n$). Recordem la definició d'aquest conjunt.

Definició 13.2.

Anomenem espai nul de la matriu A *al conjunt de totes les solucions del sistema lineal* $\mathsf{A}\vec{x} = \vec{0}$.

Les propietats 6.3., que vam veure a la lliçó 6., demostren que l'espai nul és un subespai de $\mathbb{R}^n$. Per determinar la dimensió d'aquest subespai només cal recordar que les solucions del sistema lineal depenen de tants paràmetres com columnes no principals tenen les formes esglaonades de la matriu. En conseqüència, aquesta dimensió és $n - \operatorname{rang}\mathsf{A}$:

$$\boxed{\dim \mathrm{Nul}\,\mathsf{A} = n - \operatorname{rang}\mathsf{A}}$$

Exemple 13.2.

Trobeu una base de l'espai nul de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ -1 & 2 & 1 & 2 & 0 \\ 1 & -2 & 2 & 1 & 0 \\ 4 & -8 & 0 & -4 & 1 \end{bmatrix}$$

Aquesta matriu és la mateixa de l'exemple anterior. La forma esglaonada reduïda és

$$\mathsf{R} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

així que el sistema lineal $\mathsf{A}\vec{x} = \vec{0}$ és equivalent a $\mathsf{R}\vec{x} = \vec{0}$. Resolent-lo obtindrem

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \vec{x} = \vec{0}$$

$$\left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) \vec{x} = \vec{0}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \vec{x} = - \begin{bmatrix} 0 & -2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \vec{x}$$

$$\begin{aligned} x_1 &= 2x_2 + x_4 \\ x_3 &= -x_4 \\ x_5 &= 0 \end{aligned}$$

i la solució es pot expressar com

$$\vec{x} = \alpha_1(2,1,0,0,0) + \alpha_2(1,0,-1,1,0)$$

Per tant, podem escollir aquesta base:

$$\boxed{B_{\mathrm{Nul}\,\mathsf{A}} = \{(2,1,0,0,0), (1,0,-1,1,0)\}}$$

Com esperavem, l'espai nul és un subespai de $\mathbb{R}^5$ i la dimensió d'aquest subespai és $5-\operatorname{rang}\mathsf{A} = 2$. □

13.1.3. L'espai fila

Si les combinacions lineals de les columnes de la matriu A determinen un subespai de $\mathbb{R}^m$, òbviament les combinacions lineals de les files de A determinaran un subespai de $\mathbb{R}^n$.

Definició 13.3.

Anomenem espai fila de la matriu A*,* Fil *A al conjunt de totes les combinacions lineals de les files de* A*.*

Com que el nombre de files linealment independents de la matriu A coincideix amb el rang d'aquesta matriu, la dimensió de l'espai fila també és el rang de A:

$$\boxed{\dim \operatorname{Fil}\mathsf{A} = \operatorname{rang}\mathsf{A}}$$

Per a trobar una base de l'espai fila també podem fer servir l'esglaonament de la matriu A. Però hi ha una diferència important respecte al cas de l'espai columna: l'espai fila de A si que coincideix amb l'espai fila de les formes esglaonades S de A, perquè les operacions elementals es fan per files: les files de A són combinacions de les files de S i, recíprocament, les files de S són combinacions de les de A.

Exemple 13.3.

Trobeu una base de l'espai fila de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ -1 & 2 & 1 & 2 & 0 \\ 1 & -2 & 2 & 1 & 0 \\ 4 & -8 & 0 & -4 & 1 \end{bmatrix}$$

Com que la forma esglaonada reduïda és

$$\mathsf{R} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

podem elegir la base

$$\boxed{b_{\mathrm{Fil\,A}} = \{(1,-2,0,-1,0),(0,0,1,1,0),(0,0,0,0,1)\}}$$

que té tres vectors, com era d'esperar, perquè la dimensió de l'espai fila ha de ser igual al rang de A. □

Atés que les tres primeres files de R són independents, podem afirmar que les tres primeres files de A són una base de Fil A? Per què?

És evident que l'espai fila d'una matriu és el mateix que l'espai columna de la matriu transposada,

$$\boxed{\mathrm{Fil\,A} = \mathrm{Col\,A}^t}$$

així que, alternativament, podríem calcular una base de l'espai fila esglaonant la matriu transposada, en comptes de la matriu A.

Exemple 13.4.

Trobeu una base de l'espai fila de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ -1 & 2 & 1 & 2 & 0 \\ 1 & -2 & 2 & 1 & 0 \\ 4 & -8 & 0 & -4 & 1 \end{bmatrix}$$

fent servir l'esglaonament de la matriu transposada A^t.

La forma esglaonada reduïda de la matriu

$$\mathsf{A}^t = \begin{bmatrix} 1 & -1 & 1 & 4 \\ -2 & 2 & -2 & -8 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

és

$$\mathsf{R} = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

d'on deduïm que les columnes independents de la matriu A^t són la primera, la segona i la quarta (la tercera és el triple de la primera més el dobles de la segona). Així que podem escollir, com a base de $\mathrm{Fil\,A} = \mathrm{Col\,A}^t$,

$$\boxed{B'_{\mathrm{Fil\,A}} = \{(1,-2,0,-1,0),(-1,2,1,2,0),(4,-8,0,-4,1)\}}$$

13.1.4. L'espai nul esquerre

El quart espai és l'espai nul de A^t.

Definició 13.4.

Anomenem espai nul esquerre de la matriu A *a l'espai nul de la matriu transposada* A^t*, és a dir, al conjunt de les solucions del sistema lineal* $\mathsf{A}^t\vec{x} = \vec{0}$.

Es tracta d'un subespai de $\mathbb{R}^m$, de dimensió $m - \mathrm{rang\,A}$. El nom de subespai nul *esquerre* es justifica pel fet que l'equació $\mathsf{A}^t\vec{x} = \vec{0}$ es pot escriure també com $\vec{x}^{\,t}\mathsf{A} = \vec{0}^t$. Per exemple, si $\mathsf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \end{bmatrix}$, el sistema

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

el podem escriure, alternativament, com

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

així que l'espai nul *esquerre* és el conjunt dels vectors que anulen la matriu *per l'esquerra*.
Exemple 13.5.

Trobeu una base de l'espai nul esquerre de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & -2 & 0 & -1 & 0 \\ -1 & 2 & 1 & 2 & 0 \\ 1 & -2 & 2 & 1 & 0 \\ 4 & -8 & 0 & -4 & 1 \end{bmatrix}$$

La forma esglaonada reduïda de la matriu

$$\mathsf{A}^t = \begin{bmatrix} 1 & -1 & 1 & 4 \\ -2 & 2 & -2 & -8 \\ 0 & 1 & 2 & 0 \\ -1 & 2 & 1 & -4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

és

$$\mathsf{R} = \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

així que la incògnita no principal és la tercera. Les solucions del sistema $\mathsf{A}^t\vec{x} = \vec{0}$ són

$$x_1 = -3\alpha, \quad x_2 = -2\alpha, \quad x_3 = \alpha, \quad x_4 = 0$$

o bé

$$\vec{x} = \alpha(-3, -2, 1, 0)$$

i la base que busquem es aquesta:

$$\boxed{B_{\mathrm{Nul}\,\mathsf{A}^t} = \{(-3, -2, 1, 0)\}}$$

que té $4 - \operatorname{rang} \mathsf{A} = 4 - 3 = 1$ vector. $\square$

13.2. Ortogonalitat i sistemes lineals

Direm que dos conjunts de vectors de $\mathbb{R}^n$, A i B, són ortogonals si tots els vectors de A són ortogonals a tots els de B, i anomenarem *espai ortogonal de* A, $A^\perp$, al conjunt de tots els vectors que són ortogonals a A. El nom d'*espai* ortogonal està justificat per la propietat següent:
Propietat 13.1.

Siga A un conjunt qualsevol de vectors de $\mathbb{R}^n$. El conjunt $A^\perp$ és un subespai de $\mathbb{R}^n$.

Demostració: Com que el vector zero és ortogonal a qualsevol altre vector, tenim que $\vec{0} \in A^\perp$. L'altra condició que s'ha de complir és que si elegim dos vectors qualssevol en $A^\perp$, $\vec{u}_1$ i $\vec{u}_2$ i dos escalars α_1 i α_2, llavors la combinació lineal $\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2$ també es troba en $A^\perp$; ara bé, que $\vec{u}_1$ i $\vec{u}_2$ siguen elements de $A^\perp$ vol dir que, per a qualsevol $\vec{a} \in A$, $\vec{u}_1 \cdot \vec{a} = 0$ i $\vec{u}_2 \cdot \vec{a} = 0$. En conseqüència,

$$(\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2) \cdot \vec{a} = \alpha_1(\vec{u}_1 \cdot \vec{a}) + \alpha_2(\vec{u}_2 \cdot \vec{a}) = 0$$

així que $\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 \in A^\perp, \ \forall \vec{a} \in A$ $\square$

Si A és un conjunt finit, diguem $A = \{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m\}$, posant aquest vectors com a columnes de la matriu M_A, $\mathsf{M}_A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_m \end{bmatrix}$, tindrem que un vector $\vec{x}$ és ortogonal a A si i només si $\mathsf{M}_A^t\vec{x} = \vec{0}$

(perquè els components del producte $\mathsf{M}_A^t\vec{x}$ són els productes escalars dels vectors de A pel vector $\vec{x}$). En conseqüència,

> Si A és un conjunt finit, llavors l'ortogonal de A és l'espai nul esquerre de la matriu M_A: $A^\perp = \operatorname{Nul}\mathsf{M}_A^t$.

de manera que el càlcul de $A^\perp$ es redueix a la resolució d'un sistema lineal.

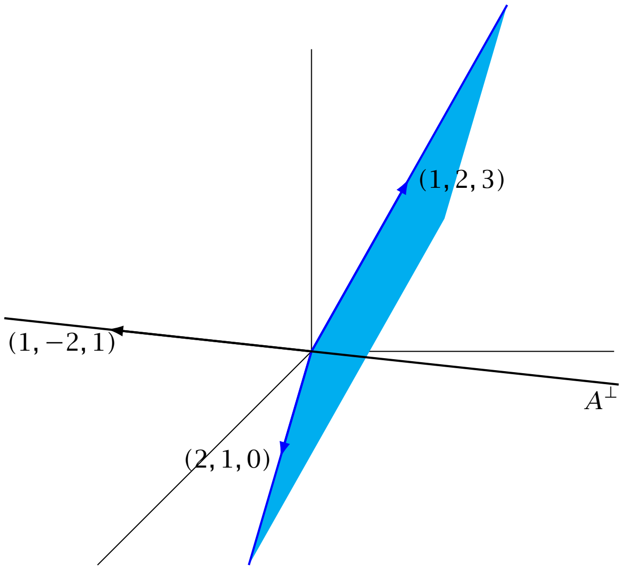
Exemple 13.6.

> Calculeu de l'espai ortogonal del conjunt $A = \{(1,2,3),(2,1,0)\}$

Construïm la matriu M_A^t, posant-hi com a files els vectors del conjunt A: $\mathsf{M}_A^t = \begin{bmatrix}1&2&3\\2&1&0\end{bmatrix}$ i resolem el sistema homogeni $\mathsf{M}_A^t\vec{x} = \vec{0}$:

La forma esglaonada reduïda de la matriu $\left[\begin{array}{ccc|c}1&2&3&0\\2&1&0&0\end{array}\right]$ és $\left[\begin{array}{ccc|c}1&0&-1&0\\0&1&2&0\end{array}\right]$, així que

$$A^\perp = \{\alpha(1,-2,1) : \lambda \in \mathbb{R}\}$$



Observeu que, com que $A^\perp$ és ortogonal als vectors $(1,2,3)$ i $(2,1,0)$, també és ortogonal a totes les combinacions lineals d'aquests dos vectors, de manera que $A^\perp$ és la recta ortogonal al pla generat pels vectors de A, que no és altra cosa que l'espai columna de la matriu $\mathsf{M}_A = \begin{bmatrix}1&2\\2&1\\3&0\end{bmatrix}$. □

13.2.1. Ortogonalitat dels quatre subespais

Recordem esquemàticament el que sabem a prop dels quatre subespais associats a la matriu A:

> Suposem que A és una matriu $m \times n$ de rang r. Aleshores,
>
> 1. Col A i Nul A^t són subespais de $\mathbb{R}^m$ i
>   - $\dim \operatorname{Col}\mathsf{A} = r$
>   - $\dim \operatorname{Nul}\mathsf{A}^t = m - r$
> 2. Fil A i Nul A són subespais de $\mathbb{R}^n$ i
>   - $\dim \operatorname{Fil}\mathsf{A} = r$
>   - $\dim \operatorname{Nul}\mathsf{A} = n - r$

Ara podem afegir una nova propietat de tipus geomètric: els dos subespais nuls són els ortogonals dels espais fila i columna.

Propietat 13.2.

Si A és una matriu real llavors, l'ortogonal de l'espai fila és l'espai nul.

Demostració: Bàsicament, la prova consisteix a observar que, com que cada equació del sistema $A\vec{x} = \vec{0}$ és el producte escalar d'una fila de A pel vector $\vec{x}$ igualat a zero, els elements de l'espai nul de A són ortogonals a les files de A. □

☞ En conseqüència, Si A és una matriu real llavors, l'ortogonal de l'espai columna és l'espai nul esquerre.

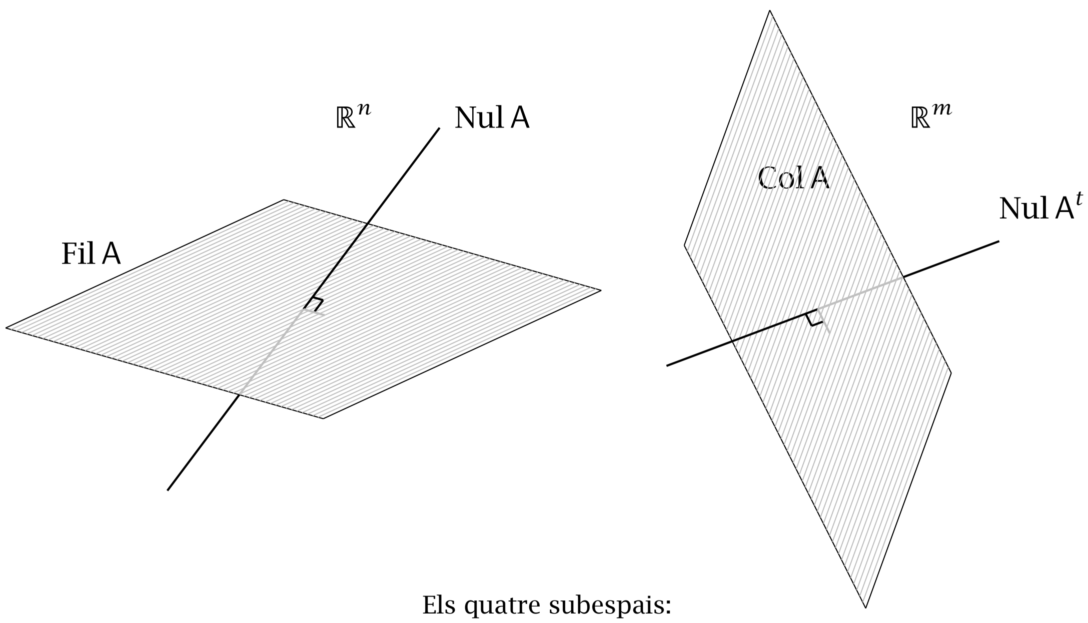
El teorema següent resumeix tot el que hem dit a prop dels quatre subespais.

Teorema 13.3. (Teorema fonamental de l'Àlgebra Lineal. Versió provisional)

Suposem que A *és una matriu* $m \times n$ *de rang* r. *Aleshores,*

1. $\operatorname{Col} A$ *i* $\operatorname{Nul} A^t$ *són subespais de* $\mathbb{R}^m$ *i*
   - $\dim \operatorname{Col} A = r$
   - $\dim \operatorname{Nul} A^t = m - r$
   - *Aquests dos subespais són cadascun l'ortogonal de l'altre.*
2. $\operatorname{Fil} A$ *i* $\operatorname{Nul} A$ *són subespais de* $\mathbb{R}^n$ *i*
   - $\dim \operatorname{Fil} A = r$
   - $\dim \operatorname{Nul} A = n - r$
   - *Aquests dos subespais són cadascun l'ortogonal de l'altre.* □

El gràfic següent mostra totes aquestes propietats (l'espai nul *esquerre* és el de la dreta!).

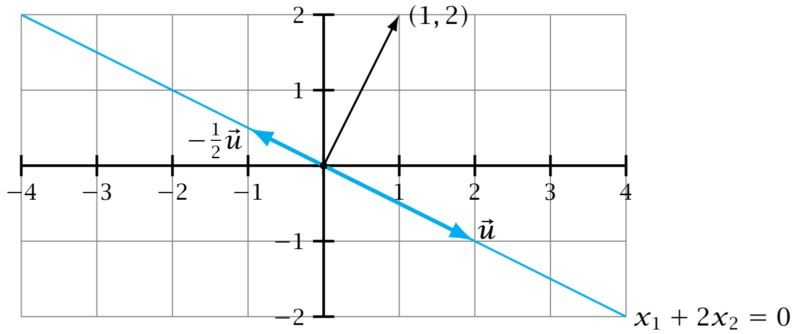


Els quatre subespais:
Interpretació geomètrica del Teorema Fonamental

13.3. Subespais i matrius. Les equacions d'un subespai

Hi ha dues maneres típiques de presentar un subespai: mostrant-ne una base (o un conjunt generador) o bé un sistema d'equacions la solució del qual és el subepai. Això es pot fer perquè *qualsevol subespai de $\mathbb{R}^n$ és l'espai columna d'una matriu i l'espai nul d'una altra.* Per exemple, el subespai (de $\mathbb{R}^2$) $F = \{\alpha(2,-1) : \alpha \in \mathbb{R}\}$ és l'espai columna de la matriu $\left[\begin{smallmatrix}2\\-1\end{smallmatrix}\right]$ i també l'espai nul de la matriu $\begin{bmatrix}1 & 2\end{bmatrix}$:

$$F = \operatorname{Col}\begin{bmatrix}2\\-1\end{bmatrix} = \{(1,2)\}^\perp = \operatorname{Nul}\begin{bmatrix}1 & 2\end{bmatrix}$$



Vegem com podem trobar aquestes matrius:

- El més senzill és representar un subespai com l'espai columna d'una matriu: si B és una base de F llavors,

$$F = \operatorname{Col} \mathsf{M}_B$$

- D'altra banda, com que l'ortogonal de l'espai columna és l'espai nul esquerre, $F^\perp = (\operatorname{Col} \mathsf{M}_B)^\perp = \operatorname{Nul} M_B^t$, si trobem una base B' de $\operatorname{Nul} M_B^t$, tindrem que

$$F = (F^\perp)^\perp = B'^\perp = \operatorname{Nul} \mathsf{M}_{B'}$$

  (les columnes de la matriu $\mathsf{M}_{B'}$ són solucions del sistema lineal $M_B\vec{x} = \vec{0}$).

Exemple 13.7.

> Trobeu una base del subespai
>
> $$F = \{(a+b, a-b, a) : a, b \in \mathbb{R}\}$$
>
> i espresseu F com l'espai columna d'una matriu i l'espai nul d'una altra matriu.

Primer de tot, trobarem una base de F. Si observem que

$$F = \{a(1,1,1) + b(1,-1,0) : a, b \in \mathbb{R}\}$$

com que els vectors $(1,1,1)$ i $(1,-1,0)$ són linealment independents, el conjunt

$$B_1 = \{(1,1,1), (1,-1,0)\}$$

és base de F. En conseqüència,

$$\boxed{F = \operatorname{Col}\begin{bmatrix}1 & 1\\1 & -1\\1 & 0\end{bmatrix}}$$

Tot seguit trobarem una base de $F^\perp = \operatorname{Nul}\begin{bmatrix}1 & 1 & 1\\1 & -1 & 0\end{bmatrix}$, resolent el sistema lineal

$$\begin{bmatrix}1 & 1 & 1\\1 & -1 & 0\end{bmatrix}\vec{x} = \begin{bmatrix}0\\0\end{bmatrix}$$

La solució general d'aquest sistema és $\vec{x} = \alpha(-1/2, -1/2, 1)$, així que

$$\boxed{F = \text{Nul}\begin{bmatrix}-1/2 & -1/2 & 1\end{bmatrix}}$$

Que un subespai coincidisca amb l'espai nul d'una matriu vol dir que podem expressar aquest subespai con el conjunt de solucions d'un sistema d'equacions lineals.

Així, com que el subespai F del darrer exemple és l'espai nul de la matriu $\mathsf{A}_2 = \begin{bmatrix}-1/2 & -1/2 & 1\end{bmatrix}$ tindrem que

$$\boxed{F = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : -\frac{1}{2}x_1 - \frac{1}{2}x_2 + x_3 = 0\right\}}$$

Direm que $-\frac{1}{2}x_1 - \frac{1}{2}x_2 + x_3 = 0$ és l'equació del subespai F.

En resum, tal com dèiem al principi d'aquest apartat, tenim dues maneres bàsiques de descriure un subespai:

- Com l'embolcall lineal d'un conjunt de vectors. És a dir, donant-ne un conjunt generador o, millor, una base. Si amb els vectors d'aquest conjunt generador hi construïm una matriu A_1, llavors el subespai és l'espai columna $\text{Col}\,\mathsf{A}_1$.
- Com l'ortogonal d'un conjunt (finit) de vectors. Si construïm una matriu A_2 posant aquests vectors com a files, llavors el subespai és l'espai nul $\text{Nul}\,\mathsf{A}_2$, és a dir, el conjunt de solucions del sistema d'equacions $\mathsf{A}_2\vec{x} = \vec{0}$.

Les files de A_2 són una base del subespai ortogonal a les columnes de A_1.

13.4. Exercicis

13.4.1. Els quatre subespais

Exercici 13.1. Trobeu (si és possible) bases dels quatre subespais associats a les matrius següents:

$$\text{(a)}\quad \mathsf{M}_1 = \begin{bmatrix}1 & 0 & 1 & -1\\ 0 & 1 & 1 & 0\\ 1 & 0 & 1 & 1\end{bmatrix} \qquad \text{(b)}\quad \mathsf{M}_2 = \begin{bmatrix}1 & 0 & 1 & -1\\ 0 & 1 & 1 & 0\\ 1 & 1 & 2 & -1\end{bmatrix} \qquad \text{(c)}\quad \mathsf{M}_3 = \begin{bmatrix}1 & 0 & 1 & -1\\ 2 & 0 & 2 & -2\\ -1 & 0 & -1 & 1\end{bmatrix}$$

Exercici 13.2. La matriu A té 7 files i 5 columnes i el seu rang és 4. Quines són les dimensions dels quatre subespais associats a aquesta matriu?

Exercici 13.3. (a) Quina condició ha de complir la matriu $m \times n$ A perquè l'espai columna de A siga $\mathbb{R}^m$?

(b) I perquè l'espai nul siga $\{\vec{0}\}$?

(c) Si A és una matriu $n \times n$ invertible, quins són els quatre subespais associats a A?

Exercici 13.4. Donada la matriu

$$\mathsf{A} = \begin{bmatrix}1 & 0 & -1 & -2\\ 0 & 1 & 0 & 1\\ 2 & 0 & -2 & -4\end{bmatrix}$$

(a) Calculeu la forma esglaonada reduïda R de la matriu A, la matriu T tal que $\mathsf{R} = \mathsf{TA}$ i la matriu inversa $\mathsf{L} = \mathsf{T}^{-1}$ (tot això requereix únicament una operació elemental).

(b) Trobeu els quatre subespais deduïts de A.

(c) Quina relació hi ha entre l'espai columna de A i les columnes de L? Per què?

(d) Quina relació hi ha entre l'espai nul de A^t i les files de T? Per què?

(e) Sabent que

$$\mathsf{A} = \underbrace{\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 5 & 0 & 1\end{bmatrix}}_{\mathsf{L}}\underbrace{\begin{bmatrix}1 & 3 & 0 & 5\\ 0 & 0 & 1 & 16\\ 0 & 0 & 0 & 0\end{bmatrix}}_{\mathsf{R}}$$

determineu bases dels quatre subespais associats a A sense calcular explícitament la matriu A.

Exercici 13.5. Trobeu els valors del paràmetre λ per als quals l'espai nul de la matriu $\mathsf{A}_\lambda = \left[\begin{smallmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{smallmatrix}\right]$ *no és* el subespai $\{\vec{0}\}$.

13.4.2. Ortogonalitat i sistemes lineals

Exercici 13.6. Trobeu l'ortogonal del conjunt $A = \{(1,0,1,0),(0,1,0,1)\}$. Quina relació hi ha entre A i $(A^\perp)^\perp$?

Exercici 13.7. Quina condició s'ha de complir perquè $A = (A^\perp)^\perp$?

Exercici 13.8. Si F i G són dos subconjunts de $\mathbb{R}^n$ i $F \subset G$, quina relació hi ha entre $F^\perp$ i $G^\perp$?

Exercici 13.9. Quina relació hi ha entre els conjunt A i B, si $A^\perp = B^\perp$?

13.4.3. Subespais i matrius

Exercici 13.10. Expresseu cadascun dels subespais següents com l'espai columna d'una matriu i com l'espai nul d'una altra matriu.

(a) $F_1 = \langle (1,-1,2) \rangle$
(b) $F_2 = \langle (1,1,0),(-2,0,1) \rangle$
(c) $F_3 = \{(a, a+b, a+2b, -a) : a, b \in \mathbb{R}\}$
(d) $F_4 = \{(x_1,x_2,x_3,x_4) : x_1 + x_2 = 0, x_2 - x_3 - x_4 = 0\}$

Exercici 13.11. Trobeu una base i les equacions de cadascun dels subespais de l'exercici anterior.

Exercici 13.12. Trobeu les equacions del subespai $F = \langle (1,1,1,1),(0,1,1,0) \rangle$ i una base del subespai $G = \{(x_1,x_2,x_3,x_4) : x_1 + x_2 - x_3 - x_4 = 0\}$.

14. Intersecció i suma de subespais. Suma directa

La intersecció de dos (o més de dos) subespais també és un subespai. En canvi, la unió no sempre ho és. Per això introduïm el concepte de *suma* de subespais (el mínim subespai que conté la unió).

14.1. Intersecció de dos subespais

La intersecció de dos subespais també és un subespai.[2] Per comprovar-ho, serà suficient que observem que si elegim les matrius A_1 i A_2 de manera que $F = \operatorname{Nul} A_1$ i $G = \operatorname{Nul} A_2$, llavors les condicions que s'han de complir perquè el vector $\vec{u}$ es trobe a la intersecció $F \cap G$ són aquestes:

$$\vec{u} \in F \cap G \iff \mathsf{A}_1\vec{u} = \vec{0}, \qquad \mathsf{A}_2\vec{u} = \vec{0}$$

Però aquestes dues condicions es poden expressar com una sola, si fem servir la matriu $\left[\begin{smallmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \end{smallmatrix}\right]$:

$$\vec{u} \in F \cap G \iff \begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \end{bmatrix} \vec{u} = \vec{0}$$

Però això vol dir que

$$\vec{u} \in F \cap G \iff \vec{u} \in \operatorname{Nul} \begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \end{bmatrix}$$

En altres paraules, $F \cap G$ és l'espai nul de la matriu $\left[\begin{smallmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \end{smallmatrix}\right]$, així que, efectivament, és un subespai. $\square$

☞ A banda de demostrar que la intersecció de dos subespais també és un subespai, hem trobat un algorisme senzill per calcular la intersecció de dos subespais:

Càlcul de la intersecció de dos subespais F i G

(a) Expresseu F i G com a espais nuls: $F = \operatorname{Nul} \mathsf{A}_1$, $G = \operatorname{Nul} \mathsf{A}_2$

(b) La intersecció és $F \cap G = \operatorname{Nul} \begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \end{bmatrix}$

Exemple 14.1.

Trobeu el subespai intersecció dels subespais

$$F = \langle (1,1,0,1), (1,2,1,1) \rangle$$
$$G = \langle (1,1,1,1), (1,3,3,1) \rangle$$

(Primera solució) Primer de tot expressarem els dos subespais com a espais nuls. Com que

$$F = \operatorname{Col} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

l'ortogonal de F és

$$F^{\perp} = \operatorname{Nul} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 \end{bmatrix} = \operatorname{Nul} \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} = \langle (1,-1,1,0), (-1,0,0,1) \rangle$$

així que

$$F = \operatorname{Nul} \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$$

De manera anàloga trobem que

$$G = \operatorname{Nul} \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$

[2] La intersecció de més de dos subespais (encara en el cas d'infinits subespais) també és un subespai.

Per tant, la intersecció és

$$F \cap G = \mathrm{Nul}\begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \langle (0,1,1,0) \rangle \quad \square$$

(Solució alternativa) Com que, en aquest exemple, coneixíem una base de cada subespai, hem hagut de trobar els subespais ortogonals per tal de poder expressar F i G com a espais nuls. Si volem evitar aquests càlculs, podem raonar de la següent manera: Si el vector $\vec{u}$ es troba a la intersecció $F \cap G$, llavors,

– per una banda, $\vec{u} \in F$, així que existeixen dos escalars x_1 i x_2 de manera que

$$\vec{u} = x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{x}, \qquad \text{on } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

– per una altra, per ser un vector de G,

$$\vec{u} = a_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + a_2 \begin{bmatrix} 1 \\ 3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \vec{a}, \qquad \text{on } a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Reunint aquestes dues condicions, tindrem que

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \vec{a}$$

Aquesta igualtat la podem interpretar com un sistema d'equacions lineals en el qual les incògnites són les x. Restant a la segona i a la quarta equacions d'aquest sistema la primera tindrem

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \vec{a}$$

i, restant la segona de la tercera,

$$\left[\begin{array}{cc} 1 & 1 \\ 0 & 1 \\ \hline 0 & 0 \\ 0 & 0 \end{array}\right] \vec{x} = \left[\begin{array}{cc} 1 & 1 \\ 0 & 2 \\ \hline 1 & 1 \\ 0 & 0 \end{array}\right] \vec{a} = \left[\begin{array}{c} a_1 + a_2 \\ 2a_2 \\ \hline a_1 + a_2 \\ 0 \end{array}\right]$$

D'aquesta manera, la matriu del costat esquerre és esglaonada i el sistema serà compatible si $a_1 + a_2 = 0$ (o $a_1 = -a_2$). Per tant,

$$\vec{u} = -a_2(1,1,1,1) + a_2(1,3,3,1) = a_2(0,2,2,0)$$

En definitiva,

$$\boxed{F \cap G = \langle (0,1,1,0) \rangle} \quad \square$$

Aquest segon mètode és el més eficient si es vol trobar una base de la intersecció de dos subespais F i G dels quals es coneixen les bases $B_F = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ i $B_G = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q\}$: com que

$$F = \mathrm{Col}\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_p \end{bmatrix} = \mathrm{Col}\, \mathsf{M}_{B_F}$$
$$G = \mathrm{Col}\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_q \end{bmatrix} = \mathrm{Col}\, \mathsf{M}_{B_G}$$

un vector $\vec{u}$ que es trobe en el subespai intersecció serà combinació lineal dels vectors d'una base i també dels de l'altra:

$$\vec{u} = \mathsf{M}_{B_F}\vec{x} \qquad \vec{u} = \mathsf{M}_{B_G}\vec{a}$$

així que hem de trobar els valors dels paràmetres $\vec{a}$ perquè el sistema

$$\mathsf{M}_{B_F}\vec{x} = \mathsf{M}_{B_G}\vec{a} \tag{14.1}$$

siga compatible.

Com que el rang de M_{B_F} és p (el nombre de columnes que té, perquè són linealment independents), la forma esglaonada reduïda d'aquesta matriu és $\begin{bmatrix} \mathrm{I} \\ \mathsf{O} \end{bmatrix}$ (I és la identitat $p \times p$ i O és la matriu zero $(n-p) \times p$). En conseqüència, fent operacions elementals el sistema (14.1) queda d'aquesta manera:

$$\begin{bmatrix} \mathrm{I} \\ \mathsf{O} \end{bmatrix} \vec{x} = \begin{bmatrix} \mathsf{V}_1 \\ \mathsf{V}_2 \end{bmatrix} \vec{a}$$

Finalment, perquè aquest sistema siga compatible caldrà que

$$\mathsf{V}_2\vec{a} = \vec{0}$$

☞ La solució general d'aquest darrer sistema lineal ens proporciona les condicions que han de complir els paràmetres $\vec{a}$.

Esquemàticament,

Càlcul de la intersecció de dos subespais F i G
(a partir de les bases de F i G)

(a) Elegiu dues bases B_F (base de F) i B_G (base de G)

(b) Feu les operacions elementals convenients per a reduir la matriu $\left[\mathsf{M}_{B_F} \mid \mathsf{M}_{B_G}\right]$ a la forma

$$\left[\mathsf{M}_{B_F} \mid \mathsf{M}_{B_G}\right] \longrightarrow \left[\begin{array}{c|c} \mathsf{U} & \mathsf{V}_1 \\ \mathsf{O} & \mathsf{V}_2 \end{array}\right]$$

on U és invertible

(c) Resoleu el sistema homogeni

$$\mathsf{V}_2\vec{a} = \vec{0}$$

(d) La intersecció és el conjunt de tots els vectors de la forma

$$\vec{u} = \mathsf{M}_{B_G}\vec{a}$$

on $\vec{a}$ són les solucions del sistema que acabem de resoldre.

14.2. Suma de dos subespais

Al contrari que la intersecció, la unió $F \cup G$ dels subespais F i G no és sempre un subespai.

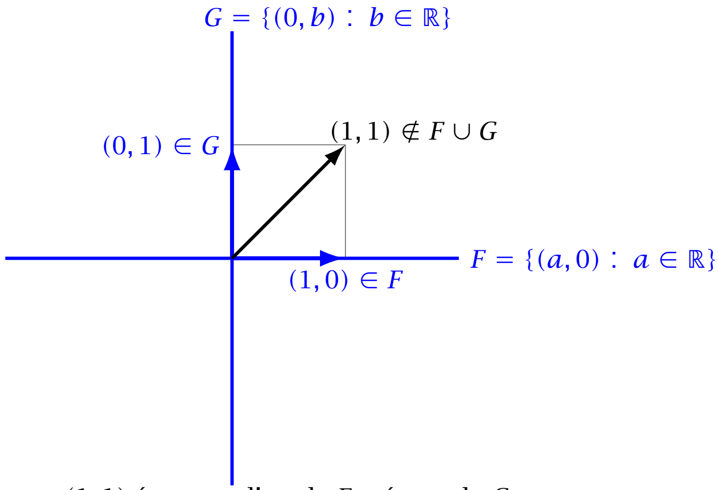
Exemple 14.2.

Siguen F i G els subespais de $\mathbb{R}^2$ definits com

$$F = \{(a, 0) \,:\, a \in \mathbb{R}\}$$
$$G = \{(0, b) \,:\, b \in \mathbb{R}\}$$

(F i G són les dues rectes coordenades). Proveu que $F \cup G$ no és un subespai de $\mathbb{R}^2$.

Si sumem els vectors $(1, 0)$ (que és de F) i $(0, 1)$ (de G) obtenim $(1, 0) + (0, 1) = (1, 1)$ que no és de F ni de G! $\square$



El vector $(1,1)$ és suma d'un de F més un de G
però no es troba a F ni a G

Com que, en general, la unió no arriba a ser subespai, hi afegirem els vectors que calga fins aconseguir un subespai.

Definició 14.1.

La suma dels subespais F i G és el conjunt de totes les sumes d'un element de F més un de G. Aquest conjunt el representarem com $F+G$:

$$F + G = \{\vec{u} + \vec{v} : \vec{u} \in F, \ \vec{v} \in G\}$$

Exemple 14.3.

La suma dels subespais $F = \{(a,0) : a \in \mathbb{R}\}$ i $G = \{(0,b) : b \in \mathbb{R}\}$ és l'espai $\mathbb{R}^2$.

Aquesta propietat és evident, perquè qualsevol vector de $\mathbb{R}^2$ el podem escriure com

$$\vec{x} = (x_1, x_2) = (x_1, 0) + (0, x_2) \quad \square$$

Per trobar un mètode de càlcul de la suma, notem que, si el conjunt B_F és base de F i B_G ho és de G, llavors la unió $B_F \cup B_G$ genera $F+G$. En altre paraules,

$$\text{si } B_F \text{ és base de } F \text{ i } B_G \text{ ho és de } G \text{ llavors, } F + G = \mathrm{Col}\begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}.$$

Aquesta expressió implica que la suma $F+G$ també és un subespai (de fet és el subespai *més petit* que conté a $F \cup G$, perquè un subespai que continga els vectors de F i els de G també haurà de contenir les sumes d'aquests vectors) i ens proporciona un algorisme que cercàvem:

Càlcul de la suma de dos subespais F i G

(a) Trobeu una base, B_F, de F i una de G, B_G

(b) La suma és $F + G = \mathrm{Col}\begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}$

Exemple 14.4.

Trobeu el subespai suma dels subespais

$$F = \langle (1,1,0,1),(0,1,1,0)\rangle$$
$$G = \langle (1,1,1,1),(1,3,3,0)\rangle$$

Els conjunts $B_F = \{(1,1,0,1),(0,1,1,0)\}$ i $B_G\{(1,1,1,1),(1,3,3,0)\}$ són bases de F i G, respectivament. Per tant,

$$F + G = \text{Col}\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 3 \\ 0 & 1 & 1 & 3 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Com que la forma esglaonada reduïda d'aquesta matriu és

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

els quatre vectors són linealment independents i la dimensió de $F+G$ és 4, de manera que $F+G = \mathbb{R}^4$. □

Exemple 14.5.

Trobeu el subespai suma dels subespais

$$F = \langle (1,1,0,1),(1,2,1,1)\rangle$$
$$G = \langle (1,1,1,1),(1,3,3,1)\rangle$$

Elegint les bases $\{(1,1,0,1),(1,2,1,1)\}$ i $\{(1,1,1,1),(1,3,3,1)\}$ tindrem

$$F + G = \text{Col}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

La forma esglaonada reduïda d'aquesta matriu és

$$\begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

així que els tres primers vectors formen una base de l'espai suma:

$$\boxed{B_{F+G} = \{(1,1,0,1),(1,2,1,1),(1,1,1,1)\}} \quad \square$$

Entre aquests dos exemples hi ha una diferència important: en el primer cas, la suma de dos subespais de dimensió dos és un subespai de dimensió quatre. En canvi, en el segon exemple, la suma *només* té dimensió tres. En els propers apartats estudiarem les dimensions de la suma i la intersecció i justificarem aquesta diferència.

14.3. Ortogonals dels espais suma i intersecció

Propietat 14.1. (L'ortogonal de l'espai suma)

L'ortogonal de l'espai suma és la intersecció dels ortogonals dels dos subespais:

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

Demostració: Si $F = \operatorname{Col} A_1$ i $G = \operatorname{Col} A_2$, com que en la lliçò anterior vam veure que l'ortogonal de l'espai columna és l'espai nul esquerre, tindrem que $F^\perp = \operatorname{Nul} \mathsf{A}_1^t$ i $G^\perp = \operatorname{Nul} \mathsf{A}_2^t$ així que

$$F^\perp \cap G^\perp = \operatorname{Nul} \begin{bmatrix} \mathsf{A}_1^t \\ \mathsf{A}_2^t \end{bmatrix}$$

D'altra banda, com que la suma és l'espai columna de la matriu $\begin{bmatrix} \mathsf{A}_1 & \mathsf{A}_2 \end{bmatrix}$, l'ortogonal de l'espai suma serà

$$(F+G)^\perp = \left(\operatorname{Col} \begin{bmatrix} \mathsf{A}_1 & \mathsf{A}_2 \end{bmatrix}\right)^\perp = \operatorname{Nul} \begin{bmatrix} \mathsf{A}_1 & \mathsf{A}_2 \end{bmatrix}^t = \operatorname{Nul} \begin{bmatrix} \mathsf{A}_1^t \\ \mathsf{A}_2^t \end{bmatrix}$$

de manera que, efectivament, els subespais $(F+G)^\perp$ i $F^\perp \cap G^\perp$ són iguals. □

Canviant en aquesta propietat F per $F^\perp$ i G per $G^\perp$ resultarà que

$$(F^\perp + G^\perp)^\perp = F \cap G$$

així que

Propietat 14.2. (L'ortogonal de l'espai intersecció)

L'ortogonal de l'espai intersecció és la suma dels ortogonals dels dos subespais:

$$(F \cap G)^\perp = F^\perp + G^\perp = \quad \square$$

14.4. La dimensió de la suma i la intersecció

Notem que tant la suma com la intersecció de dos subespais F i G es poden calcular a partir de la matriu que resulta de concatenar les matrius associades a una base de cada subespai, $\mathsf{M} = \begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}$, en un cas, calculant-ne l'espai columna i en l'altre a partir de l'espai nul d'una matriu deduïda d'aquesta. Per tant, és clar que les dimensions d'aquests dos espais han d'estar relacionades d'alguna manera, entre elles, i amb aquesta matriu.

14.4.1. La dimensió de la suma

El més senzill és calcular la dimensió de l'espai suma: Si B_F i B_G són bases de F i G, respectivament, com que $F + G$ és l'espai columna de la matriu M, és clar que la dimensió de $F + G$ és

$$\boxed{\dim(F+G) = \operatorname{rang} \begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}} \tag{14.2}$$

14.4.2. La dimensió de la intersecció

Recordem el que hem de fer per a calcular la intersecció $F \cap G$:

- Transformem, per mitjà d'operacions elementals, la matriu $\begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}$ a una de la forma

$$\begin{bmatrix} \mathsf{U} & \mathsf{V}_1 \\ \mathsf{O} & \mathsf{V}_2 \end{bmatrix}$$

  on U és invertible.

- Resolem el sistema homogeni

$$\mathsf{V}_2 \vec{a} = \vec{0}$$

- Multiplicant M_{B_G} per les solucions d'aquest sistema obtenim tots els vectors de la intersecció.

Per tant, la dimensió de l'espai intersecció serà el nombre d'incògnites no principals que tinga el sistema homogeni $\mathsf{V}_2 \vec{a} = \vec{0}$, és a dir,

$$\dim(F \cap G) = \dim(\operatorname{Nul} \mathsf{V}_2)$$

Per acabar de calcular aquesta dimensió ens hem de fixar en les dimensions de les matrius que hi intervenen. En la matriu

$$\begin{bmatrix} \mathsf{U} & \mathsf{V}_1 \\ \mathsf{O} & \mathsf{V}_2 \end{bmatrix}$$

tenim aquestes dimensions (suposem que p i q són, respectivament, les dimensions de F i G):

(a) U és una matriu $p \times p$ (perquè p és la dimensió de F).

(b) O és la matriu zero $(n-p) \times p$ (perquè estem en $\mathbb{R}^n$).

(c) V_1 és una matriu $p \times q$ (perquè q és la dimensió de G).

(d) V_2 és una matriu $(n-p) \times q$.

$$n\left\{ \overbrace{\left[\begin{array}{c|c} p \times p & p \times q \\ \hline (n-p) \times p & (n-p) \times q \end{array}\right]}^{p+q} \right.$$

Per tant,

$$\dim(F \cap G) = \dim(\operatorname{Nul} \mathsf{V}_2) = q - \operatorname{rang} \mathsf{V}_2$$

(perquè la dimensió de l'espai nul és el nombre de columnes menys el rang de la matriu). D'altra banda, el rang de V_2 és la diferència entre el rang de la matriu gran, M, i el rang del primer bloc, que és p, així que

$$\begin{aligned} \dim(F \cap G) &= \dim(\operatorname{Nul} \mathsf{V}_2) \\ &= q - \operatorname{rang} \mathsf{V}_2 \\ &= q - (\operatorname{rang} \mathsf{M} - p) \\ &= p + q - \operatorname{rang} \mathsf{M} \end{aligned}$$

És a dir,

$$\boxed{\dim(F \cap G) = \dim F + \dim G - \operatorname{rang} \begin{bmatrix} \mathsf{M}_{B_F} & \mathsf{M}_{B_G} \end{bmatrix}} \tag{14.3}$$

14.4.3. La fórmula de Grassman

Comparant les fórmules (14.2) i (14.3) obtenim el resultat següent:

Teorema 14.3. (Fórmula de Grassman)

Si F i G són dos subespais de $\mathbb{R}^n$, llavors

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G) \quad \square \tag{14.4}$$

Exemple 14.6.

Comproveu la fórmula de Grassman en el cas dels subespais de $\mathbb{R}^4$

$$\begin{aligned} F &= \langle (1,1,0,1), (1,2,1,1) \rangle \\ G &= \langle (1,1,1,1), (1,3,3,1) \rangle \end{aligned}$$

(a) Les dimensions de F i G són aquestes: $\dim F = 2$ i $\dim G = 2$.

(b) D'altra banda, en l'exemple 14.1. hem trobat que $F \cap G = \langle (0,1,1,0) \rangle$, així que $\dim(F \cap G) = 1$.

(c) i, en l'exemple 14.5., que $\{(1,1,0,1), (0,1,1,0), (1,1,1,1)\}$ és una base de $F + G$; per tant, $\dim(F + G) = 3$.

Llavors,

$$\dim F + \dim G - \dim(F \cap G) = 2 + 2 - 1 = 3 = \dim(F + G) \quad \square$$

14.5. Suma directa

Quan $F \cap G = \{\vec{0}\}$ la fórmula de Grassman es redueix a

$$\dim(F+G) = \dim F + \dim G$$

i es pot construir una base de la suma fent la unió d'una base de F i una altra de G. En aquest cas, es diu que la suma $F+G$ és directa.

Definició 14.2.

La suma dels subespais F i G és directa si $\dim(F+G) = \dim F + \dim G$ *(o, equivalentment, quan $F \cap G = \{\vec{0}\}$).*

☞ Quan es vol remarcar que la suma és directa s'escriu $F \oplus G$ (en comptes de $F+G$).

La propietat següent mostra que el fet que la suma dels subespais siga directa vol dir que aquests dos subespais són independents, d'una manera anàloga a la independència lineal de dos vectors.

Propietat 14.4.

Si F i G són dos subespais de $\mathbb{R}^n$, llavors les afirmacions següents són equivalents:

(a) La suma dels subespais F i G és directa.

(b) Si $\vec{u}_F$ i $\vec{u}_G$ són, respectivament, vectors de F i de G i si $\vec{u}_F + \vec{u}_G = \vec{0}$, llavors $\vec{u}_F = \vec{u}_G = \vec{0}$.

(c) Si $\vec{u}$ és un vector de la suma $F+G$, llavors existeix un únic vector $\vec{u}_F$ i un únic vector $\vec{u}_G$ de manera que

$$\vec{u} = \vec{u}_F + \vec{u}_G \quad \square$$

La clau d'aquesta propietat és la *unicitat*: encara que la suma no siga directa, qualsevol vector de $F+G$ és la suma d'un de F i un altre de G, però si la suma no és directa podem trobar diverses parelles de vectors amb aquesta propietat. El que fa especial la suma directa és que no pot haver dues parelles de vectors en F i en G amb aquesta propietat.

Exemple 14.7.

Comproveu que el vector $\vec{u} = (2,3,2,2)$ es un element de la suma $F+G$ dels dos subespais

$$F = \langle (1,1,0,1), (1,2,1,1) \rangle$$
$$G = \langle (1,1,1,1), (1,3,3,1) \rangle$$

És directa, la suma $F+G$?

Hem de comprovar que el vector $\vec{u}$ es troba en l'espai columna de la matriu

$$\mathsf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

és a dir, que el sistema $\mathsf{M}\vec{x} = \vec{u}$ és compatible.

Esglaonant la matriu ampliada $\begin{bmatrix} \mathsf{M} & \vec{u} \end{bmatrix}$ obtenim la forma esglaonada reduïda següent:

$$\begin{bmatrix} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

així que el sistema és compatible i, en conseqüència, el vector $\vec{u}$ es troba en la suma $F+G$. Per a expressar-lo com a combinació lineal de les columnes de M hem de trobar la solució del sistema:

$$x_1 = 2\alpha,\ x_2 = 1-2\alpha,\ x_3 = 1-\alpha,\ x_4 = \alpha$$

Per tant,

$$\vec{u} = \underbrace{2\alpha(1,1,0,1) + (1-2\alpha)(1,2,1,1)}_{\in F} + \underbrace{(1-\alpha)(1,1,1,1) + \alpha(1,3,3,1)}_{\in G} \tag{14.5}$$

on α és qualsevol nombre. Per tant, la suma no és directa. □

Entendrem correctament perquè hi ha més d'una solució, quan la suma no és directa, reordenant adequadament la combinació lineal (14.5):

$$\begin{aligned}\vec{u} =& 2\alpha(1,1,0,1) + (1-2\alpha)(1,2,1,1) + (1-\alpha)(1,1,1,1) + \alpha(1,3,3,1)\\ =&(1,2,1,1) + 2\alpha(1,1,0,1) - 2\alpha(1,2,1,1) + (1,1,1,1) - \alpha(1,1,1,1) + \alpha(1,3,3,1)\\ =& \underbrace{(1,2,1,1)}_{\in F} - \underbrace{2\alpha(0,1,1,0)}_{\in F\cap G} + \underbrace{(1,1,1,1)}_{\in G} + \underbrace{2\alpha(0,1,1,0)}_{\in F\cap G}\end{aligned}$$

Com que el vector $(0,1,1,0)$ es troba a tots dos subespais, podem sumar-lo i restar-lo a un vector de F i a un altre de G, de manera que obtindrem una nova combinació lineal que produeix el mateix resultat.

Exemple 14.8.

Comproveu que el vector $\vec{u} = (2,3,2,2)$ es un element de la suma $F + G$ dels dos subespais

$$\begin{aligned}F &= \langle(1,1,0,1),(0,1,1,0)\rangle\\ G &= \langle(1,1,1,1),(1,3,3,0)\rangle\end{aligned}$$

És directa, la suma $F + G$?

Hem de comprovar que el vector $\vec{u}$ es troba en l'espai columna de la matriu

$$\mathsf{M} = \begin{bmatrix}1&0&1&1\\1&1&1&3\\0&1&1&3\\1&0&1&0\end{bmatrix}$$

és a dir, que el sistema $\mathsf{M}\vec{x} = \vec{u}$ és compatible.

Esglaonant la matriu ampliada $\begin{bmatrix}\mathsf{M} & \vec{u}\end{bmatrix}$ obtenim la forma esglaonada reduïda següent:

$$\begin{bmatrix}1&0&0&0&1\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&0\end{bmatrix}$$

així que el sistema és compatible determinat i, en conseqüència, el vector $\vec{u}$ es troba en la suma $F + G$. Per a expressar-lo com a combinació lineal de les columnes de M hem de trobar la solució del sistema:

$$x_1 = 1,\ x_2 = 1,\ x_3 = 1,\ x_4 = 0$$

Per tant,

$$\begin{aligned}\vec{u} &= (1,1,0,1) + (0,1,1,0) + (1,1,1,1) + 0(1,3,3,1)\\ &= \underbrace{(1,2,1,1)}_{\in F} + \underbrace{(1,1,1,1)}_{\in G}\end{aligned}$$

i la suma és directa. □

14.5.1. Suma directa de diversos subespais

Les definicions de suma i suma directa es poden generalitzar a més de dos subespais.

Definició 14.3.

(a) La suma dels subespais F_1, F_2, …, F_r és el conjunt de totes les sumes d'un element de cadascun dels subespais:

$$F_1 + F_2 + \cdots + F_r = \{\vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_r\ :\ \vec{u}_1 \in F_1, \vec{u}_2 \in F_2, \dots, \vec{u}_r \in F_r\}$$

(b) Diem que aquesta suma és directa si $\dim(F_1 + F_2 + \cdots + F_r) = \dim F_1 + \dim F_2 + \cdots + \dim F_r$.

☞ Quan es vol remarcar que la suma és directa s'escriu $F_1 \oplus F_2 \oplus \cdots \oplus F_r$.

Les propietats 14.4. continuen sent vàlides:

Propietat 14.5.

Si $F_1, F_2, \ldots, F_r$ són subespais de $\mathbb{R}^n$, llavors les afirmacions següents són equivalents:

(a) La suma $F_1 + F_2 + \cdots + F_r$ és directa.

(b) Si $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$ són, respectivament, vectors de $F_1, F_2, \ldots, F_r$ i si $\vec{u}_1 + \vec{u}_2 + \cdots + \vec{u}_r = \vec{0}$, llavors $\vec{u}_1 = \vec{u}_2 = \cdots = \vec{u}_r = \vec{0}$.

(c) Si $\vec{u}$ és un vector de la suma, llavors existeix un únic vector en cadascun dels subespais, $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$ de manera que

$$\vec{u} = \vec{u}_1\vec{u}_2 + \cdots + \vec{u}_r \quad \square$$

☞ El fet que la intersecció de tots els subespais siga nul·la *no assegura* que la suma siga directa.

Exemple 14.9.

La intersecció dels tres subespais de $\mathbb{R}^3$

$$F_1 = \langle(1,0,0),(0,1,0)\rangle, F_2 = \langle(1,0,0),(0,0,1)\rangle, F_3 = \langle(0,1,0),(0,0,1)\rangle$$

és buida, però la suma no és directa, perquè $F_1 + F_2 + F_3 = \mathbb{R}^3$ i $\dim F_1 + \dim F_2 + \dim F_3 = 6$. $\square$

14.6. Exercicis

14.6.1. Suma i intersecció de subespais

Exercici 14.1. En cada apartat, trobeu una base del subespai intersecció dels dos subespais F i G donats.

(a) $F = \langle(1,1,0),(1,0,1)\rangle$, $G = \langle(0,1,1),(1,1,1)\rangle$

(b) $F = \{(x_1,x_2,x_3) : x_1 - x_2 = 0\}$, $G = \{(x_1,x_2,x_3) : x_3 = 0\}$

Exercici 14.2. En cada apartat, trobeu una base del subespai suma dels dos subespais F i G donats.

(a) $F = \langle(1,1,0),(1,0,1)\rangle$, $G = \langle(0,1,1),(1,1,1)\rangle$

(b) $F = \{(x_1,x_2,x_3,x_4) : x_1 - x_2 = 0, x_2 - x_3 = 0\}$, $G = \{(x_1,x_2,x_3,x_4) : x_1 + x_2 - 2x_3 + x_4 = 0, x_2 - x_3 = 0\}$

Exercici 14.3. (a) Trobeu les equacions del subespai de $\mathbb{R}^4$ $F = \langle(1,2,0,1),(2,3,0,3),(3,2,1,2)\rangle$. (b) Trobeu una base de $G = \{(x_1,x_2,x_3,x_4) \in \mathbb{R}^4 : x_1 - x_2 + x_3 + x_4 = 0\}$. (c) Proveu que $F + G = \mathbb{R}^4$. (d) Determineu la dimensió de $F \cap G$. (e) Trobeu una base de $F \cap G$.

14.6.2. La suma directa

Exercici 14.4. Determineu si la suma dels subespais de $\mathbb{R}^3$

$$F = \langle\{(1,1,0),(1,0,1),(1,2,-1)\}\rangle, \qquad G = \{(x_1,x_2,x_3) : x_1 = x_2 = x_3\}$$

és directa.

Exercici 14.5. Proveu que si E i F són dos subespais de $\mathbb{R}^n$ i E i F són ortogonals entre ells, llavors la suma $E + F$ és directa.

CAPÍTOL 4

ORTOGONALITAT I MÍNIMS QUADRATS

Continguts

15. Suma directa ortogonal i projeccions ortogonals

Si F és un subespai de $\mathbb{R}^n$, llavors $\mathbb{R}^n$ és la suma directa de F i $F^\perp$. En altres paraules, qualsevol vector $\vec{b}$ (de l'espai $\mathbb{R}^n$) es descompon en la suma d'un vector de F, $\vec{b}_F$, i un altre de $F^\perp$, $\vec{b}_{F^\perp}$; aquesta descomposició és única.

15.1. Suma directa ortogonal i projeccions ortogonals

El teorema que demostrem tot seguit mostra que la suma d'un subespai qualsevol i el seu ortogonal sempre és directa i, a més, a més, aquesta suma *ompli* tot l'espai.

Teorema 15.1.

Si F és un subespai qualsevol de $\mathbb{R}^n$, llavors

$$\mathbb{R}^n = F \oplus F^\perp$$

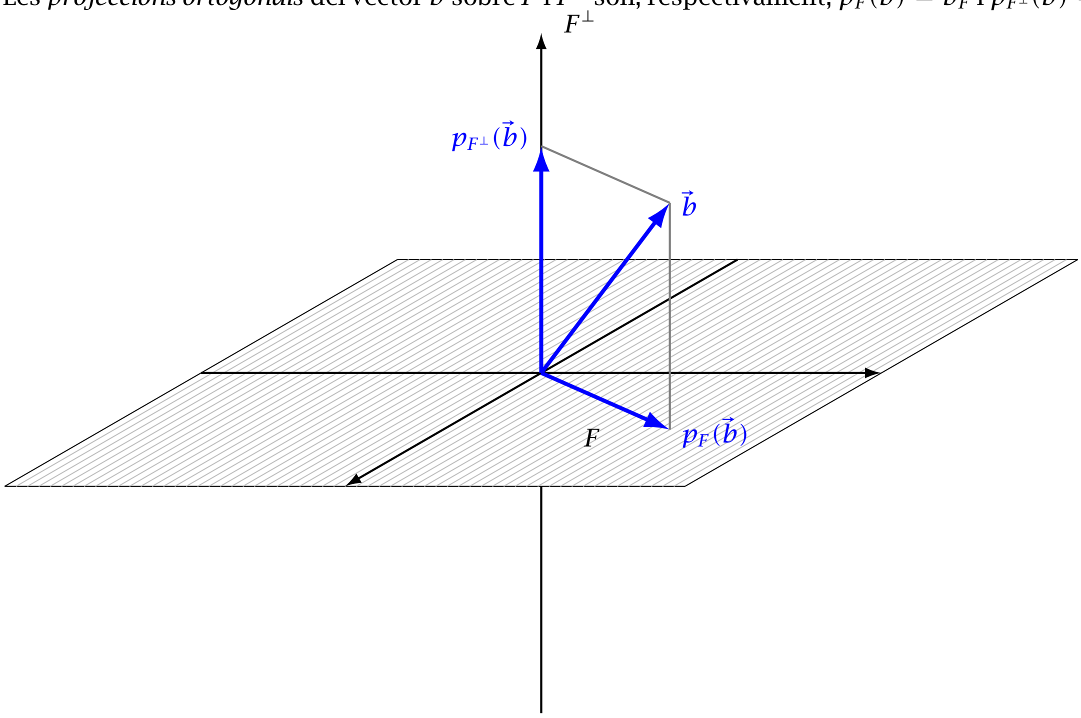
Per tant, qualsevol vector $\vec{b} \in \mathbb{R}^n$ es pot escriure, de forma única com a suma d'un vector de F i un altre de $F^\perp$:

$$\vec{b} = \vec{b}_F + \vec{b}_{F^\perp} \qquad \forall u \in \mathbb{R}^n$$

Demostració: Si $\vec{b}$ és un element de $F \cap F^\perp$, llavors $\vec{b}$ és ortogonal a ell mateix, de manera que $\vec{b} \cdot \vec{b} = \vec{0}$, així que $\vec{b} = \vec{0}$. Això demostra que $F \cap F^\perp = \{\vec{0}\}$ i la suma és directa, així que només ens falta provar que $\dim F + \dim F^\perp = n$. Però, com que $F = \operatorname{Col} \mathsf{A}$ per a alguna matriu A; llavors $F^\perp = \operatorname{Nul} \mathsf{A}^t$ i les dimensions de F i $F^\perp$ són, respectivament,

$$\left.\begin{array}{r} \dim F = \dim \operatorname{Col} \mathsf{A} = \operatorname{rang} \mathsf{A} \\ \dim F^\perp = \dim \operatorname{Nul} \mathsf{A}^t = n - \operatorname{rang} \mathsf{A} \end{array}\right\} \Rightarrow \dim F + \dim F^\perp = n \quad \square$$

☞ Les *projeccions ortogonals* del vector $\vec{b}$ sobre F i $F^\perp$ són, respectivament, $p_F(\vec{b}) = \vec{b}_F$ i $p_{F^\perp}(\vec{b}) = \vec{b}_{F^\perp}$.



Com a casos particulars obtenim que els espais $\mathbb{R}^m$ i $\mathbb{R}^n$ són la suma directa ortogonal dels quatre espais deduïts d'una matriu $m \times n$:

Corol·lari 15.2.

Si A és una matriu real $m \times n$, llavors

$$\mathbb{R}^m = \operatorname{Col} \mathsf{A} \oplus \operatorname{Nul} \mathsf{A}^t$$
$$\mathbb{R}^n = \operatorname{Fil} \mathsf{A} \oplus \operatorname{Nul} \mathsf{A} \quad \square$$

Aquest corol·lari completa el teorema fonamental de l'Àlgebra Lineal:

Teorema 15.3. (Teorema fonamental de l'Àlgebra Lineal (versió definitiva))

Suposem que A *és una matriu* $m \times n$ *de rang* r*. Aleshores,*

1. $\operatorname{Col}\mathsf{A}$ *i* $\operatorname{Nul}\mathsf{A}^t$ *són subespais de* $\mathbb{R}^m$ *i*
   - $\dim\operatorname{Col}\mathsf{A} = r$
   - $\dim\operatorname{Nul}\mathsf{A}^t = m - r$
   - *Aquests dos subespais són cadascun l'ortogonal de l'altre.*
   - $\mathbb{R}^m = \operatorname{Col}\mathsf{A} \oplus \operatorname{Nul}\mathsf{A}^t$
2. $\operatorname{Fil}\mathsf{A}$ *i* $\operatorname{Nul}\mathsf{A}$ *són subespais de* $\mathbb{R}^n$ *i*
   - $\dim\operatorname{Fil}\mathsf{A} = r$
   - $\dim\operatorname{Nul}\mathsf{A} = n - r$
   - *Aquests dos subespais són cadascun l'ortogonal de l'altre.*
   - $\mathbb{R}^n = \operatorname{Fil}\mathsf{A} \oplus \operatorname{Nul}\mathsf{A}$ □

15.2. Càlcul de la projecció ortogonal

Exemple 15.1.

Calculeu la projecció ortogonal del vector $\vec{b} = (2,3,2)$ sobre el pla

$$F = \langle (1,1,0), (2,3,1) \rangle$$

En primer lloc cercarem l'ortogonal de F:

$$\begin{aligned} F^\perp &= \operatorname{Nul}\begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} & \\ &= \operatorname{Nul}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} & \text{(Forma esglaonada reduïda)} \\ &= \langle (1,-1,1) \rangle & \text{(Solució del sistema homogeni)} \end{aligned}$$

El següent pas consisteix a expressar el vector $\vec{b}$ com a combinació lineal dels vectors de la base

$$B = \{(1,1,0), (2,3,1), (1,-1,1)\}$$

Esglaonant la matriu ampliada

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

obtenim

$$\begin{bmatrix} 1 & 0 & 0 & -5/3 \\ 0 & 1 & 0 & 5/3 \\ 0 & 0 & 1 & 1/3 \end{bmatrix}$$

de manera que

$$\vec{b} = \underbrace{-\frac{5}{3}(1,1,0) + \frac{5}{3}(2,3,1)}_{p_F(\vec{b})} + \underbrace{\frac{1}{3}(1,-1,1)}_{p_{F^\perp}(\vec{b})}$$

i la projecció ortogonal de $\vec{b}$ sobre F és el vector

$$\vec{p}_F(\vec{b}) = -\frac{5}{3}(1,1,0) + \frac{5}{3}(2,3,1) = \frac{5}{3}(1,2,1) \quad \square$$

En aquest exemple hem hagut de trobar una base de l'ortogonal de F i unir-la a la base de F per a construir una base de l'espai complet. Hi ha, però, un mètode alternatiu, que ens permet trobar les projeccions ortogonals, explotant l'ortogonalitat i sense necessitat de calcular cap base. El mètode es basa en aquests fets:

(a) Si $F = \operatorname{Col} \mathsf{A}$, llavors $p_F(\vec{b}) \in \operatorname{Col} \mathsf{A}$, així que $p_F(\vec{b}) = \mathsf{A}\vec{x}$ per a algun vector $\vec{x}$.

(b) La projecció de $\vec{b}$ sobre $F^\perp$ és $p_{F^\perp}(\vec{b}) = \vec{b} - p_F(\vec{b}) = \vec{b} - \mathsf{A}\vec{x}$, és a dir $\vec{b} - \mathsf{A}\vec{x} \in (\operatorname{Col} \mathsf{A})^\perp$.

(c) L'ortogonal de l'espai columna és l'espai nul esquerre: $(\operatorname{Col} \mathsf{A})^\perp = \operatorname{Nul} \mathsf{A}^t$

Així doncs, per a trobar la projecció ortogonal $p_F(\vec{b})$ podem cercar en primer lloc un vector $\vec{x}$ de manera que $\vec{b} - \mathsf{A}\vec{x} \in \operatorname{Nul} \mathsf{A}^t$ i, tot seguit, calcular $p_F(\vec{b})$ fent el producte $\mathsf{A}\vec{x}$. La condició clau és aquesta:

$$\vec{b} - \mathsf{A}\vec{x} \in \operatorname{Nul} \mathsf{A}^t$$

és a dir,

$$\mathsf{A}^t(\vec{b} - \mathsf{A}\vec{x}) = \vec{0}$$

o, millor,

$$\boxed{\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}} \tag{15.1}$$

L'expressió (15.1) és un sistema d'equacions lineals; si $\vec{x}$ és una solució d'aquest sistema llavors la projecció ortogonal de $\vec{b}$ sobre F és el vector $p_F(\vec{b}) = \mathsf{A}\vec{x}$.

☞ Les *equacions normals* en el problema de la projecció ortogonal són les equacion (15.1).

Algorisme de càlcul de la projecció ortogonal

Expresseu el subespai F com l'espai columna d'una matriu A. llavors,

(a) Calculeu els productes $\mathsf{A}^t\mathsf{A}$ i $\mathsf{A}^t\vec{b}$

(b) Resoleu el sistema de les equacions normals $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$

(c) Calculeu el vector $p_F(\vec{b}) = \mathsf{A}\vec{x}$.

Ara apliquem aquestes idees per a tornar a calcular la projecció ortogonal de l'exemple 15.1..

Exemple 15.2.

Calculeu la projecció ortogonal del vector $\vec{b} = (2,3,2)$ sobre el pla

$$F = \langle (1,1,0), (2,3,1) \rangle$$

El subespai F és l'espai columna de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}$.

(a) Calculem els productes $\mathsf{A}^t\mathsf{A}$ i $\mathsf{A}^t\vec{b}$:

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 14 \end{bmatrix} \qquad \mathsf{A}^t\vec{b} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix}$$

(b) Resolem el sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$:

$$\begin{bmatrix} 2 & 5 \\ 5 & 14 \end{bmatrix} \vec{x} = \begin{bmatrix} 5 \\ 15 \end{bmatrix} \iff \vec{x} = \begin{bmatrix} -5/3 \\ 5/3 \end{bmatrix}$$

(c) Calculem el vector $p_F(\vec{b}) = \mathsf{A}\vec{x}$:

$$p_F(\vec{b}) = \mathsf{A}\vec{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -5/3 \\ 5/3 \end{bmatrix} = \frac{5}{3}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Aquest algorisme se simplifica molt si fem servir bases ortogonals (o, millor encara, ortonormals). Si $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ és una base ortogonal de F i $\mathsf{A} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \end{bmatrix}$, llavors

(a) Els productes $\mathsf{A}^t\mathsf{A}$ i $\mathsf{A}^t\vec{b}$ són

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} \vec{v}_1\cdot\vec{v}_1 & 0 & \dots & 0 \\ 0 & \vec{v}_2\cdot\vec{v}_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \vec{v}_p\cdot\vec{v}_p \end{bmatrix} \qquad \mathsf{A}^t\vec{b} = \begin{bmatrix} \vec{v}_1\cdot\vec{b} \\ \vec{v}_2\cdot\vec{b} \\ \dots \\ \vec{v}_p\cdot\vec{b} \end{bmatrix}$$

(b) El sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$ serà

$$\begin{bmatrix} \vec{v}_1\cdot\vec{v}_1 & 0 & \dots & 0 \\ 0 & \vec{v}_2\cdot\vec{v}_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \vec{v}_p\cdot\vec{v}_p \end{bmatrix} \begin{bmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \dots \\ \vec{x}_p \end{bmatrix} = \begin{bmatrix} \vec{v}_1\cdot\vec{b} \\ \vec{v}_2\cdot\vec{b} \\ \dots \\ \vec{v}_p\cdot\vec{b} \end{bmatrix} \iff \begin{array}{c} (\vec{v}_1\cdot\vec{v}_1)\vec{x}_1 = \vec{v}_1\cdot\vec{b} \\ (\vec{v}_2\cdot\vec{v}_2)\vec{x}_2 = \vec{v}_2\cdot\vec{b} \\ \dots \\ (\vec{v}_p\cdot\vec{v}_p)\vec{x}_p = \vec{v}_p\cdot\vec{b} \end{array}$$

de manera que la solució d'aquest sistema lineal és

$$\vec{x} = \left(\frac{\vec{v}_1\cdot\vec{b}}{\vec{v}_1\cdot\vec{v}_1}, \frac{\vec{v}_2\cdot\vec{b}}{\vec{v}_2\cdot\vec{v}_2}, \dots, \frac{\vec{v}_p\cdot\vec{b}}{\vec{v}_p\cdot\vec{v}_p} \right)$$

(c) I la projecció ortogonal de $\vec{b}$ sobre F és

$$p_F(\vec{b}) = \mathsf{A}\vec{x} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \end{bmatrix} \begin{bmatrix} \dfrac{\vec{v}_1\cdot\vec{b}}{\vec{v}_1\cdot\vec{v}_1} \\ \dfrac{\vec{v}_2\cdot\vec{b}}{\vec{v}_2\cdot\vec{v}_2} \\ \dots \\ \dfrac{\vec{v}_p\cdot\vec{b}}{\vec{v}_p\cdot\vec{v}_p} \end{bmatrix}$$

$$\boxed{p_F(\vec{b}) = \frac{\vec{v}_1\cdot\vec{b}}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \frac{\vec{v}_2\cdot\vec{b}}{\vec{v}_2\cdot\vec{v}_2}\vec{v}_2 + \cdots + \frac{\vec{v}_p\cdot\vec{b}}{\vec{v}_p\cdot\vec{v}_p}\vec{v}_p} \tag{15.2}$$

☞ Si $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ és una base ortogonal de F, llavors la projecció ortogonal del vector $\vec{b}$ sobre F es calcula aplicant directament la fórmula (15.2).

☞ Si $B = \{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_p\}$ és una base ortonormal de F aquesta fórmula se simplifica una mica més:

$$\boxed{p_F(\vec{b}) = (\vec{q}_1\cdot\vec{b})\vec{q}_1 + (\vec{q}_2\cdot\vec{b})\vec{q}_2 + \cdots + (\vec{q}_p\cdot\vec{b})\vec{q}_p} \tag{15.3}$$

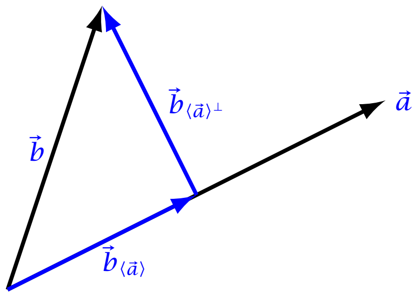
☞ La fórmula (15.3) podem expressar-la matricialment com

$$\begin{aligned} p_F(\vec{b}) &= (\vec{q}_1\cdot\vec{b})\vec{q}_1 + (\vec{q}_2\cdot\vec{b})\vec{q}_2 + \cdots + (\vec{q}_p\cdot\vec{b})\vec{q}_p \\ &= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_p \end{bmatrix} \begin{bmatrix} \vec{q}_1^{\,t}\vec{b} \\ \vec{q}_2^{\,t}\vec{b} \\ \dots \\ \vec{q}_p^{\,t}\vec{b} \end{bmatrix} \\ &= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_p \end{bmatrix} \begin{bmatrix} \vec{q}_1^{\,t} \\ \vec{q}_2^{\,t} \\ \dots \\ \vec{q}_p^{\,t} \end{bmatrix} \vec{b} = \mathsf{Q}\mathsf{Q}^t\vec{b} \end{aligned}$$

Si $B = \{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_p\}$ és una base ortonormal de F, llavors

$$\boxed{p_F(\vec{b}) = QQ^t\vec{b}} \tag{15.4}$$

15.2.1. Projecció ortogonal sobre una recta



Si el subespai F és una recta, $F = \langle\vec{a}\rangle$, llavors podem aplicar la fórmula (15.2) i obtindrem que la projecció ortogonal del vector $\vec{b}$ sobre la recta F és

$$p_F(\vec{b}) = \left(\frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}}\right)\vec{a}$$

Exemple 15.3.

Trobeu la projecció ortogonal del vector $\vec{b} = (1,3)$ sobre la recta generada per $\vec{a} = (4,2)$.

$$p_{\langle\vec{a}\rangle}(\vec{b}) = \left(\frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}}\right)\vec{a} = \left(\frac{(4,2)\cdot(1,3)}{(4,2)\cdot(4,2)}\right)(4,2) = \frac{10}{20}(4,2) = (2,1) \quad \square$$

15.3. Exercicis

Exercici 15.1. Calculeu l'espai ortogonal de $F = \langle\{(1,1,0,0),(0,1,1,0)\}\rangle$ i calculeu les projeccions ortogonals del vector $\vec{v} = (2,1,2,1)$ sobre F i sobre $F^\perp$.

Exercici 15.2. Calculeu les projeccions ortogonals del vector $\vec{v} = (2,1,-3)$ sobre el subespai $F = \langle(1,1,1),(2,1,-1),(1,0,-2)\rangle$ i sobre $F^\perp$.

Exercici 15.3. Comproveu que el conjunt $S = \{\frac{1}{15}(-5,14,-2), \frac{1}{15}(10,5,10)\}$ és ortonormal i calculeu la projecció ortogonal del vector $\vec{v} = (-3,2,-1)$ sobre $F = \langle S\rangle$.

Exercici 15.4. Siga $F = \mathrm{Col}\left(\frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix}\right)$. Calculeu la projecció del vector $\vec{u} = (1,1,1,1)$ sobre $F^\perp$.

Exercici 15.5. Siga F la recta generada pel vector $\vec{b} = (1,2,3)$.

(a) Calculeu les projeccions sobre F i sobre $F^\perp$ del vector $\vec{v} = (3,2,1)$.

(b) La funció $f(\alpha) = d((3,2,1), \alpha(1,2,3))$ mesura la distància del vector $(3,2,1)$ a cada punt de la recta F. Calculeu el mínim d'aquesta funció.

(c) Compareu els resultats dels dos apartats anteriors i interpreteu-los geomètricament.

16. Aproximació per mínims quadrats

De vegades un sistema lineal és incompatible, però ens pot interessar trobar el vector que està més a prop de ser-ne la solució. Per exemple, perquè el sistema deriva d'un problema físic que ha de tenir solució i, si no en té, és perquè les dades procedeixen de medicions que no són exactes.
En aquests casos es pot fer servir la tècnica dels mínims quadrats.

16.1. Propietats de la matriu $\mathsf{A}^t\mathsf{A}$

En el càlcul de la projecció ortogonal juga un paper fonamental la matriu $\mathsf{A}^t\mathsf{A}$ (cosa que no ens ha d'estranyar, tenint en compte que estem tractant un problema d'ortogonalitat, i les entrades d'aquesta matriu són els productes escalars de les columnes de A). I hi intervenen quatre vectors:

(a) el vector $\vec{b}$, que volem projectar sobre $F = \operatorname{Col}\mathsf{A}$,

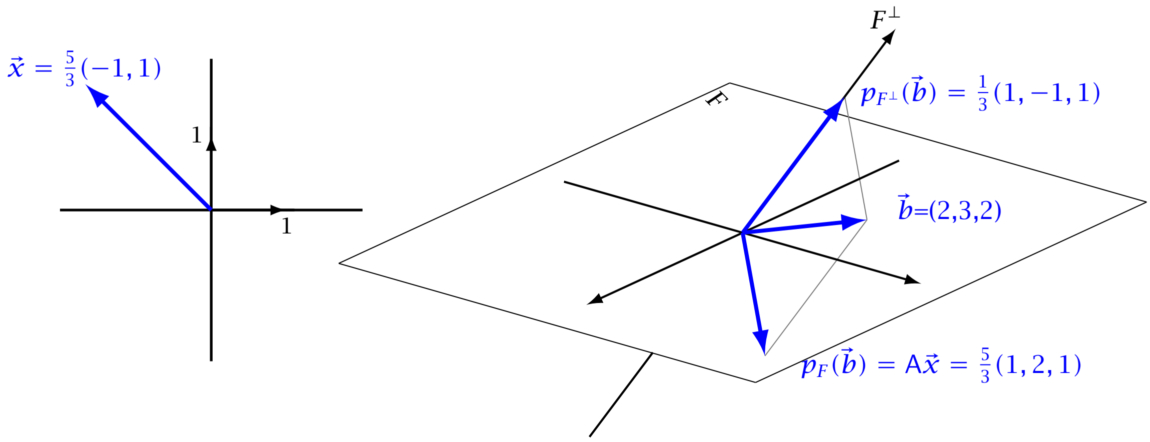
(b) el vector $p_F(\vec{b})$, projecció ortogonal de $\vec{b}$ sobre F,

(c) el vector $p_{F^\perp}(\vec{b}) = \vec{b} - p_F(\vec{b})$, projecció ortogonal de $\vec{b}$ sobre $F^\perp$ i

(d) el vector $\vec{x}$, solució del sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$, per al qual $p_F(\vec{b}) = \mathsf{A}\vec{x}$.

Tots aquests vectors són importants; molt particularment, el vector $\vec{x}$; en els problemes d'aproximació òptima el que ens interessarà, en realitat, és $\vec{x}$, en comptes de $p_F(\vec{b})$.

La figura següent mostra aquests quatre vectors en el cas de l'exemple (15.2.): la projecció ortogonal del vector $\vec{b} = (2,3,2)$ sobre el subespai $F = \langle(1,1,0),(2,3,1)\rangle$ és $p_F(\vec{b}) = (5/3)(1,2,1)$; la projecció sobre $F^\perp$ és $p_{F^\perp}(\vec{b}) = (1/3)(1,-1,1)$ i el vector $\vec{x}$ és $(5/3)(-1,1)$.[1]



Propietats 16.1.

Suposem que A és una matriu $m \times n$.

1. *$\mathsf{A}^t\mathsf{A}$ és una matriu simètrica $n \times n$.*
2. *L'espai nul de $\mathsf{A}^t\mathsf{A}$ és el mateix que el de A.*
3. *L'espai fila de $\mathsf{A}^t\mathsf{A}$ és el mateix que el de A.*
4. *El rang de $\mathsf{A}^t\mathsf{A}$ és el mateix que el de A.*
5. *$\mathsf{A}^t\mathsf{A}$ és una matriu invertible si i només si $\operatorname{rang}\mathsf{A} = n$ (és a dir, si les columnes de A són linealment independents).*

Demostració:

[1] Aquest exemple ens dóna una nova visió de les matrius: la matriu A *transforma* l'espai $\mathbb{R}^2$ en un subespai de $\mathbb{R}^3$ (l'espai columna de A). En la unitat 20. desenvoluparem aquesta idea, estudiant les matrius com a transformacions lineals.

1. $(\mathsf{A}^t\mathsf{A})^t = \mathsf{A}^t\,(\mathsf{A}^t)^t = \mathsf{A}^t\mathsf{A}$.
2. Provarem que $\operatorname{Nul}\mathsf{A} \subset \operatorname{Nul}\mathsf{A}^t\mathsf{A}$ i $\operatorname{Nul}\mathsf{A}^t\mathsf{A} \subset \operatorname{Nul}\mathsf{A}$:
   (a) Si $\vec{b} \in \operatorname{Nul}\mathsf{A}$, llavors $\mathsf{A}\vec{b} = \vec{0}$, així que
   $$\mathsf{A}^t\mathsf{A}\vec{b} = \mathsf{A}^t(\mathsf{A}\vec{b}) = \mathsf{A}^t\vec{0} = \vec{0}$$
   Per tant, $\operatorname{Nul}\mathsf{A} \subset \operatorname{Nul}\mathsf{A}^t\mathsf{A}$.
   (b) Si $\vec{b} \in \operatorname{Nul}\mathsf{A}^t\mathsf{A}$, llavors $\mathsf{A}\vec{b}$ es troba en $\operatorname{Nul}\mathsf{A}^t$ i també en $\operatorname{Col}\mathsf{A}$.
   Però $\operatorname{Col}\mathsf{A} \cap \operatorname{Nul}\mathsf{A}^t = \{\vec{0}\}$, així que $\mathsf{A}\vec{b} = \vec{0}$ i $\vec{b} \in \operatorname{Nul}\mathsf{A}$.
   Per tant, $\operatorname{Nul}\mathsf{A}^t\mathsf{A} \subset \operatorname{Nul}\mathsf{A}$.
3. $\operatorname{Fil}\mathsf{A} = (\operatorname{Nul}\mathsf{A})^\perp = (\operatorname{Nul}(\mathsf{A}^t\mathsf{A}))^\perp = \operatorname{Fil}\mathsf{A}^t\mathsf{A}$
4. Això és obvi, perquè els rangs són la dimensions dels espais fila.
5. $\mathsf{A}^t\mathsf{A}$ és invertible quan té rang màxim, és a dir, quan $\operatorname{rang}\mathsf{A}^t\mathsf{A} = n$. Tenint en compte l'apartat anterior, això equival a què el rang de A siga n. $\square$

☞ El sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$ és compatible sempre, però només és determinat quan les columnes de la matriu A són independents (perquè acabem de veure que aquesta és la condició perquè $\mathsf{A}^t\mathsf{A}$ siga invertible).

☞ En el cas que la matriu $\mathsf{A}^t\mathsf{A}$ és invertible podem trobar una fórmula explícita per a la projecció: la solució del sistema
$$\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$$
és
$$\vec{x} = (\mathsf{A}^t\mathsf{A})^{-1}\,\mathsf{A}^t\vec{b}$$
de manera que la projecció ortogonal, $\vec{b}_F = \mathsf{A}\vec{x}$, serà
$$\boxed{\vec{b}_F = \mathsf{A}\,(\mathsf{A}^t\mathsf{A})^{-1}\,\mathsf{A}^t\vec{b}} \tag{16.1}$$

16.2. Teorema de l'aproximació òptima

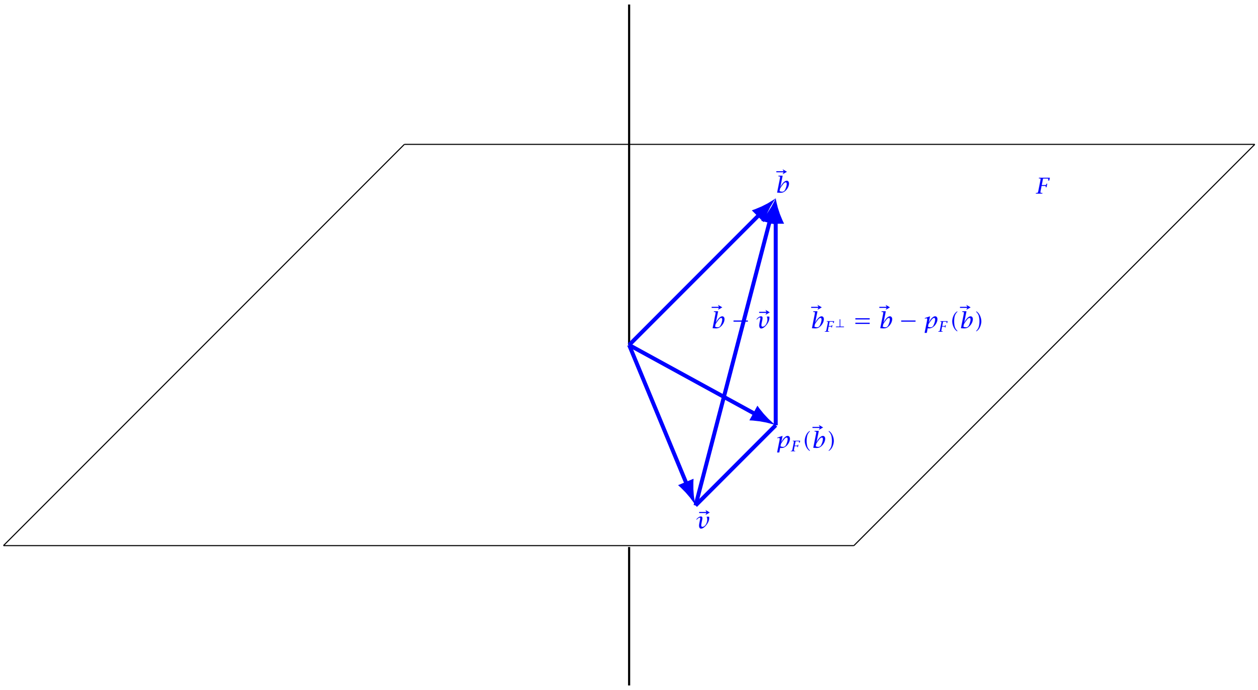
Aquest teorema té una gran importància, entre altres motius, perquè és la base teòrica del mètode d'aproximació per mínims quadrats.

Teorema 16.2. (Aproximació òptima)

El vector de F més pròxim al vector $\vec{b}$ és la projecció ortogonal de $\vec{b}$ sobre el subespai F.

Demostració: En farem dues versions, de la demostració, una de geomètrica i una altra d'algèbrica.

(a) Des del punt de vista geomètric: el vector $\vec{b}$, la projecció ortogonal $p_F(\vec{b})$ i qualsevol altre vector de F, $\vec{v}$, determinen un triangle rectangle, així que la distància $\left\|\vec{b} - \vec{v}\right\|$ (la hipotenusa del triangle) és més gran que $\left\|\vec{b} - p_F(\vec{b})\right\|$ (un catet).



(b) Prova algèbrica: Si $\vec{v}$ és un vector de F, el vector $\vec{b}-\vec{v}$ el podem escriure com

$$\vec{b}-\vec{v}=(\vec{b}-p_F(\vec{b}))+(p_F(\vec{b})-\vec{v})=p_{F^\perp}(\vec{b})+(p_F(\vec{b})-\vec{v})$$

on els vectors $p_{F^\perp}(\vec{b})$ i $p_F(\vec{b})-\vec{v}$ són ortogonals perquè el primer es troba en $F^\perp$ i el segon en F. Aleshores,

$$\begin{aligned}\left\|\vec{b}-\vec{v}\right\|^2 &= (\vec{b}-\vec{v})\cdot(\vec{b}-\vec{v})\\ &= \big((p_{F^\perp}(\vec{b})+(p_F(\vec{b})-\vec{v})\big)\cdot\big(p_{F^\perp}(\vec{b})+(p_F(\vec{b})-\vec{v})\big)\\ &= \underbrace{p_{F^\perp}(\vec{b})\cdot p_{F^\perp}(\vec{b})}_{\left\|p_{F^\perp}(\vec{b})\right\|^2}+2\underbrace{p_{F^\perp}(\vec{b})\cdot(p_F(\vec{b})-\vec{v})}_{0}+\underbrace{(p_F(\vec{b})-\vec{v})\cdot(p_F(\vec{b})-\vec{v})}_{\left\|p_F(\vec{b})-\vec{v}\right\|^2}\\ \left\|\vec{b}-\vec{v}\right\| &= \sqrt{\left\|p_{F^\perp}(\vec{b})\right\|^2+\left\|p_F(\vec{b})-\vec{v}\right\|^2}\end{aligned}$$

I aquesta norma és mínima quan $\left\|p_F(\vec{b})-\vec{v}\right\|=0$, és a dir, quan $\vec{v}=p_F(\vec{b})$. □

Aquest teorema justifica la importància de la projecció ortogonal de $\vec{b}$ sobre $F^\perp$: aquest vector mesura la distància entre el vector $\vec{b}$ i el subespai F:

$$d(\vec{b},F)=\min\{d(u,v)\,:\,v\in F\}=\left\|\vec{b}-p_F(\vec{b})\right\|=\left\|p_{F^\perp}(\vec{b})\right\|$$

Exemple 16.1.

Trobeu la projecció ortogonal del vector $\vec{b}=(3,2,3)$ sobre el subespai $F=\operatorname{Col}\mathsf{A}$, on

$$\mathsf{A}=\begin{bmatrix}1&1&0\\0&1&1\\1&2&1\end{bmatrix}$$

i la distància del vector $\vec{b}$ al subespai F.

Aquesta matriu té rang 2, perquè al tercera fila és la suma de les altres dues. En conseqüència, el sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x}=\mathsf{A}^t\vec{b}$ no serà determinat. Calculem la projecció seguint els passos habituals:

(a) Ja sabem que $F=\operatorname{Col}\mathsf{A}=\operatorname{Col}\begin{bmatrix}1&1&0\\0&1&1\\1&2&1\end{bmatrix}$

(b) Calculem els productes $\mathsf{A}^t\mathsf{A}$ i $\mathsf{A}^t\vec{b}$:

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 6 & 3 \\ 1 & 3 & 2 \end{bmatrix} \qquad \mathsf{A}^t\vec{b} = \begin{bmatrix} 6 \\ 11 \\ 5 \end{bmatrix}$$

(c) Resolem el sistema $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$ esglaonant la matriu ampliada

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 3 & 6 & 3 & 11 \\ 1 & 3 & 2 & 5 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 4/3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Com esperàvem, aquest sistema és indeterminat. La solució general és

$$\vec{x} = \left(1 + \alpha, \frac{4}{3} - \alpha, \alpha\right)$$

i la projecció de $\vec{b}$ sobre F serà

$$p_F(\vec{b}) = \mathsf{A}\vec{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ \frac{4}{3} \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right) = \frac{1}{3}\begin{bmatrix} 7 \\ 4 \\ 11 \end{bmatrix} + \alpha \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 7 \\ 4 \\ 11 \end{bmatrix}$$

La part paramètrica es fa zero, en multiplicar-la per A *perquè* $\operatorname{Nul}\mathsf{A}^t\mathsf{A} = \operatorname{Nul}\mathsf{A}$*!* i qualsevol solució del sistema ens dóna el mateix vector projecció.

La distància de $\vec{b}$ a F és

$$d(\vec{b}, F) = \left\|\vec{b} - p_F(\vec{b})\right\| = \left\|(3,2,3) - \frac{1}{3}(7,4,11)\right\| = \frac{2}{\sqrt{3}} \quad \square$$

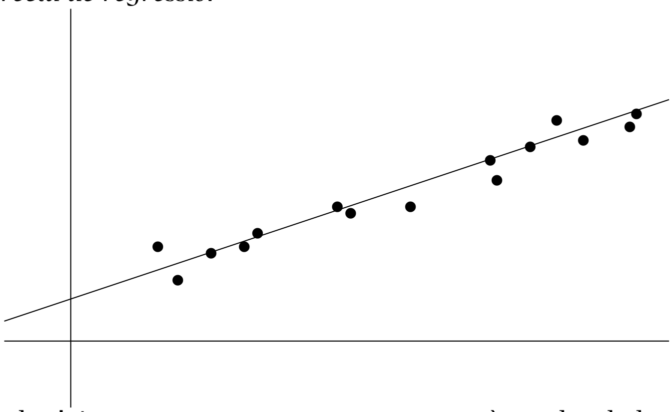
16.3. Sistemes incompatibles i mínims quadrats

En molts problemes pràctics es disposa de dades experimentals i es vol trobar una funció que aproxime aquestes dades. El cas més simple és el de la *recta de regressió*.

Suposem que estudiant un determinat fenòmen es prenen diverses medicions i les dades obtingudes es reflexen en aquesta taula

$$\{(x_1, y_1), (x_2, y_2), \ldots, (x_p, y_p)\} \quad (16.2)$$

Si en representar-les en un diagrama cartesià obtenim el resultat de la figura adjunta, és raonable deduir que aquestes dades es troben aproximadament sobre una recta $y = a + bx$.



Fins i tot, podriem pensar que les dades reals s'ajusten exactament a una recta, però que les dades experimentals contenen petits errors. En qualsevol cas, el problema seria el següent: quina és la recta que s'ajusta més a aquests valors?

Si les dades estigueren realment sobre la recta $y = a + bx$ tindríem

$$\begin{aligned} a + bx_1 &= y_1 \\ a + bx_2 &= y_2 \\ &\cdots \\ a + bx_p &= y_p \end{aligned} \tag{16.3}$$

però, excepte en el cas que només tinguem un parell de punts, allò normal és que aquest sistema siga incompatible.[2]

[2] En el tipus de problemes que estem estudiant, és normal que els sistemes lineals que resulten siguen incompatibles, perquè generalment es tracta de sistemes amb moltes equacions i poques incògnites, i com que es basen en valor experimentals, allò sorprenent seria que els sistemes foren compatibles.

Si escrivim el sistema (16.3) com

$$\mathsf{A}\vec{x} = \vec{b}$$

i es tracta d'un sistema incompatible, podem entendre que una solució *aproximada* d'aquest sistema és un vector $\vec{x}$ que minimitza la distància entre $\mathsf{A}\vec{x}$ i $\vec{b}$), ço és, la norma $\left\|\vec{b} - \mathsf{A}\vec{x}\right\|$. Si $\vec{x} = (a, b)$ és aquest vector, aleshores $y = a + bx$ és la *recta de regressió*.

DEFINICIÓ 16.1.

La recta de regressió corresponent a la taula de valors (16.2) és la recta d'equació $y = a + bx$ on $\vec{x} = (a, b)$ és el vector que minimitza la norma $e = \left\|\vec{b} - \mathsf{A}\vec{x}\right\|$.

Aquest problema es pot generalitzar substituint el sistema lineal (16.3) per qualsevol altre sistema d'equacions lineals.

DEFINICIÓ 16.2.

El problema de mínims quadrats aplicat al sistema lineal

$$\mathsf{A}\vec{x} = \vec{b} \tag{16.4}$$

consisteix a trobar el vector $\vec{x}$ que minimitza la norma $e = \left\|\vec{b} - \mathsf{A}\vec{x}\right\|$. Aquesta norma és l'error de mínims quadrats del problema.

Naturalment, si el sistema (16.4) és compatible, les solucions del problema de mínims quadrats són les solucions del sistema lineal. En cas contrari, la solució ens la proporciona el teorema de l'aproximació òptima: com que el vector de $F = \operatorname{Col}\mathsf{A}$ més pròxim a $\vec{b}$ és la projecció ortogonal de $\vec{b}$ sobre F, la solució del problema de mínims quadrats és un vector $\vec{x}$ tal que $\mathsf{A}\vec{x}$ és $p_F(\vec{b})$; és a dir, qualsevol solució del sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$.

PROPIETATS 16.3. (SOLUCIONS DEL PROBLEMA DE MÍNIMS QUADRATS)

(a) Les solucions del problema de mínims quadrats aplicat al sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ són les solucions del sistema lineal $\mathsf{A}\vec{x} = p_{\operatorname{Col}\mathsf{A}}(\vec{b})$.

(b) Les solucions del problema de mínims quadrats aplicat al sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ són les solucions del sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$.

(c) Si el sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és compatible, aleshores té les mateixes solucions que el sistema $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$, és a dir, que les solucions del problema de mínims quadrats són les solucions ordinàries.

EXEMPLE 16.2.

Trobeu la recta de regressió que ajusta la taula de valors.

x	3,9	4,5	5,2	5,4	6
y	2,1	2,8	2,4	2,8	3,1

i calculeu-ne l'error de mínims quadrats.

Les equacions (16.3), en el nostre cas són aquestes:

$$\begin{aligned} a + 3{,}9b &= 2{,}1 \\ a + 4{,}5b &= 2{,}8 \\ a + 5{,}2b &= 2{,}4 \\ a + 5{,}4b &= 2{,}8 \\ a + 6b &= 3{,}1 \end{aligned}$$

Per tant, es tracta de trobar els nombres a i b que minimitzen $\left\|\vec{b} - \mathsf{A}\vec{x}\right\|$ on

$$\mathsf{A} = \begin{bmatrix} 1 & 3{,}9 \\ 1 & 4{,}5 \\ 1 & 5{,}2 \\ 1 & 5{,}4 \\ 1 & 6 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} a \\ b \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 2{,}1 \\ 2{,}8 \\ 2{,}4 \\ 2{,}8 \\ 3{,}1 \end{bmatrix}$$

Ací, les equacions normals $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$, són

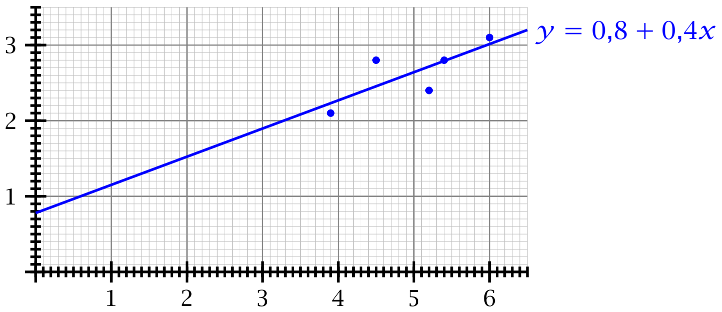
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3.9 & 4.5 & 5.2 & 5.4 & 6 \end{bmatrix} \begin{bmatrix} 1 & 3{,}9 \\ 1 & 4{,}5 \\ 1 & 5{,}2 \\ 1 & 5{,}4 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 3.9 & 4.5 & 5.2 & 5.4 & 6 \end{bmatrix} \begin{bmatrix} 2{,}1 \\ 2{,}8 \\ 2{,}4 \\ 2{,}8 \\ 3{,}1 \end{bmatrix}$$

o bé,

$$\begin{bmatrix} 5 & 25 \\ 25 & 127{,}66 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 13{,}2 \\ 66{,}99 \end{bmatrix}$$

que és un sistema compatible determinat, la solució del qual, arrodonida a una xifra decimal, és $a = 0{,}8$, $b = 0{,}4$, de manera que la recta de regressió és

$$y = 0{,}8 + 0{,}4x$$



L'error de mínims quadrats és

$$\left\|\vec{b} - \mathsf{A}\vec{x}\right\| = \left\| \begin{bmatrix} 2{,}1 \\ 2{,}8 \\ 2{,}4 \\ 2{,}8 \\ 3{,}1 \end{bmatrix} - \begin{bmatrix} 1 & 3{,}9 \\ 1 & 4{,}5 \\ 1 & 5{,}2 \\ 1 & 5{,}4 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} 0{,}8 \\ 0{,}4 \end{bmatrix} \right\| \approx 0{,}6$$

Exemple 16.3.

Comproveu que el sistema lineal

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 15 \\ 7 \\ 7 \end{bmatrix}$$

és incompatible i trobeu-ne la solució per mínims quadrats.

El sistema és incompatible, perquè la tercera fila de la matriu $\mathsf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ és combinació lineal de les dues primeres: $(1/4)(4,0) + (1/2)(0,2) = (1,1)$ i aquesta relació no es dóna en el vector $\vec{b} = (15,7,7)$: $(1/4)15 + (1/2)7 = 7{,}25 \neq 7$. Per a trobar la solució per mínims quadrats,

(a) Calculem els productes $\mathsf{A}^t\mathsf{A}$ i $\mathsf{A}\vec{b}$:

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \qquad \mathsf{A}^t\vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 15 \\ 7 \\ 7 \end{bmatrix} = \begin{bmatrix} 67 \\ 21 \end{bmatrix}$$

(b) Resolem el sistema lineal $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}\vec{b}$. La forma esglaonada reduïda de la matriu ampliada $\left[\begin{array}{cc|c} 17 & 1 & 67 \\ 1 & 5 & 21 \end{array}\right]$ és $\left[\begin{array}{cc|c} 1 & 0 & 3{,}74 \\ 0 & 1 & 3{,}45 \end{array}\right]$, així que la solució és $\vec{x} = \begin{bmatrix} 3{,}74 \\ 3{,}45 \end{bmatrix}$.

La projecció del vector $\vec{b}$ sobre l'espai columna de A és

$$p_{\mathrm{Col\,A}}(\vec{b}) = \mathsf{A}\vec{x} = \begin{bmatrix} 14{,}96 \\ 6{,}9 \\ 7{,}19 \end{bmatrix}$$

que sembla un vector bastant pròxim al desitjat $(15, 7, 7)$ El vector error és

$$\vec{b} - \mathsf{A}\vec{x} = \begin{bmatrix} 0{,}04 \\ 0{,}1 \\ -0{,}19 \end{bmatrix}$$

i l'error de mínims quadrats,

$$\left\|\vec{b} - \mathsf{A}\vec{x}\right\| = 0.048 \quad \square$$

16.4. Exercicis

Exercici 16.1. Trobeu el vector del subespai $F = \langle(1, 1, 0, 0), (0, 1, 1, 0)\rangle$ més pròxim al vector $\vec{v} = (2, 1, 2, 1)$.

Exercici 16.2. Calculeu la distància del vector $\vec{v} = (2, 1, -3)$ al subespai $F = \mathrm{Col}\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Exercici 16.3. Trobeu la solució per mínims quadrats del sistema lineal $\mathsf{A}\vec{x} = \vec{b}$, on

$$\mathsf{A} = \begin{bmatrix} 1 & 10 \\ 0 & -5 \\ -1 & 0 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix}$$

Quin és l'error de mínims quadrats?

Exercici 16.4. De la funció $y = f(x)$ coneixem la taula de valors següent:

x	0	1	2	3	4
y	−1	1	−1	−7	−17

Busqueu els polinomis $y = a_0 + a_1x + \cdots + a_kx^k$ de graus 1,2 i 3 que aproximen millor, en el sentit dels mínims quadrats, aquesta funció. Calculeu en cada cas els errors de mínims quadrats i analitzeu els resultats.

Exercici 16.5. Un conductor vol estudiar el consum de benzina en funció de la velocitat. Per això, observa els consums del seu trajecte habitual per l'autovia, recorrent-lo a diferents velocitats. Aquests són els resultats que obté:

velocitat mitjana (km/h)	90	105	117	120
consum (l)	6	6,5	7,5	8,2

Calculeu l'equació $y = a + bx$ de la recta que aproxima, en el sentit dels mínims quadrats, aquesta taula (x representa la velocitat mitjana i y el consum). Quin consum es pot esperar a una velocitat mitjana de 130 km/h?

Exercici 16.6. Un professor decideix que el mètode habitual de qualificació, la mitjana aritmètica de les notes dels exàmens, no és bastant justa. Ell pensa que cada alumne té la seua nota *natural*, la que mereix tenint en compte el seu esforç i els seus coneixements i que les notes de cada examen es desvien d'aquesta nota per diversos motius, de manera que la nota *natural* de l'alumne ha de ser el nombre més pròxim a totes les notes parcials, en el sentit dels mínims quadrats, així que aplica aquest mètode.

Un estudiant té aquestes notes en les cinc proves que s'han fet al llarg del curs: $7, 7, 3, 5, 9$. Quina és la mitjana aritmètica d'aquestes notes? Quina nota li correspon si fem servir el criteri dels mínims quadrats?

17. El mètode d'ortonormalització de Gram-Schmidt i la factorització QR

A la primera part d'aquest curs, per resoldre un sistema lineal, hem fet servir l'estratègia de l'esglaonament, els dos casos més interessants de la qual han estat els algorismes de Gauss i Gauss-Jordan, que sempre es poden interpretar com a factoritzacions de la matriu de coeficients. La factorització LU n'és un cas particular.
En aquesta unitat fem alguna cosa semblant per resoldre els problemes de mínims quadrats. Cercarem una base ortonormal de l'espai columna de la matriu implicada i observarem que això és equivalent a una factorització A = QR on les columnes de la matriu Q són ortonormals i R és una matriu triangular superior.

17.1. Operacions elementals per columnes

Les operacions elementals per columnes són equivalents a multiplicacions per matrius elementals (per la dreta).

Propietats 17.1. (Operacions elementals per columnes)

(a) *Multiplicar una matriu* A *per la matriu elemental* $\mathsf{E}_{i,j}$ *és equivalent a intercanviar les columnes* i *i* j *de* A.

(b) *Multiplicar una matriu* A *per la matriu elemental* $\mathsf{E}_i(\alpha)$ *és equivalent a multiplicar la columna* i *de* A *per* α.

(c) *Multiplicar una matriu* A *per la matriu elemental* $\mathsf{E}_{j,i}(\alpha)$ *és equivalent a sumar a la columna* i *de* A α *vegades la columna* j. □

☞ La matriu que suma a la columna i de A α vegades la columna j *no és* $\mathsf{E}_{i,j}(\alpha)$ *sinó* $\mathsf{E}_{j,i}(\alpha)$. Aquesta és l'única diferència amb el cas de les operacions elementals per files.

17.2. Obtenció de bases ortonormals. Mètode de Gram-Schmidt i factorització QR

L'algorisme de Gram-Schmidt transforma una base B_1 d'un subespai F de $\mathbb{K}^n$ en una base ortonormal B_2. El procès consisteix a construir iterativament una nova base, substituint cada vector per la seua projecció sobre l'ortogonal a l'embolcall lineal dels vectors anteriors.

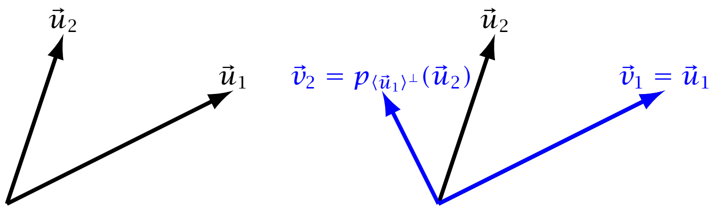
17.2.1. Algorisme de Gram-Schmidt amb dos i tres vectors vectors

Començarem amb un subespai F de dimensió dos. Si el conjunt $B_1 = \{\vec{u}_1, \vec{u}_2\}$ és una base del subespai F, com que el vector $\vec{u}_2$ és la suma de les seues projeccions ortogonals sobre la recta generada per $\vec{u}_1$ i l'ortogonal d'aquesta recta, $\vec{u}_2 = p_{\langle\vec{u}_1\rangle}(\vec{u}_2) + p_{\langle\vec{u}_1\rangle^\perp}(\vec{u}_2)$,

$$p_{\langle\vec{u}_1\rangle^\perp}(\vec{u}_2) = \vec{u}_2 - p_{\langle\vec{u}_1\rangle}(\vec{u}_2)$$

és un vector ortogonal a $\vec{u}_1$ i $\{\vec{u}_1, p_{\langle\vec{u}_1\rangle^\perp}(\vec{u}_2)\}$ és una base ortogonal de F. Anomenarem $\{\vec{v}_1, \vec{v}_2\}$ aquesta nova base,

$$\begin{aligned}\vec{v}_1 &= \vec{u}_1\\ \vec{v}_2 &= p_{\langle\vec{v}_1\rangle^\perp}(\vec{u}_2)\end{aligned}$$



Ara, com que $p_{\langle\vec{u}_1\rangle}(\vec{u}_2) = \dfrac{\vec{v}_1\cdot\vec{u}_2}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1$,

$$p_{\langle\vec{u}_1\rangle^\perp}(\vec{u}_2) = \vec{u}_2 - p_{\langle\vec{u}_1\rangle}(\vec{u}_2) = \vec{u}_2 - \frac{\vec{v}_1\cdot\vec{u}_2}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1$$

Per simplificar escriurem $\alpha_{1,2} = \dfrac{\vec{v}_1\cdot\vec{u}_2}{\vec{v}_1\cdot\vec{v}_1}$ i tindrem

$$\vec{v}_2 = \vec{u}_2 - \alpha_{1,2}\vec{v}_1$$

La base $\{\vec{v}_1, \vec{v}_2\}$ és ortogonal, així que, perquè siga ortonormal hem de dividir cada vector per la seua norma: si $\vec{q}_1 = \dfrac{1}{\|\vec{v}_1\|}\vec{v}_1$ i $\vec{q}_2 = \dfrac{1}{\|\vec{v}_2\|}\vec{v}_2$, llavors $B_2 = \{\vec{q}_1, \vec{q}_2\}$ és una base ortonormal del subespai F.

En resum, per transformar la base de F $B_1 = \{\vec{u}_1, \vec{u}_2\}$ en una base ortonormal, $B_2 = \{\vec{q}_1, \vec{q}_2\}$, apliquem l'algorisme següent:

Mètode d'ortonormalització de Gram-Schmidt (amb dos vectors)

Si $B_1 = \{\vec{u}_1, \vec{u}_2\}$ és una base del subespai F,
aquest algorisme construeix una base ortonormal de F, $B_2 = \{\vec{q}_1, \vec{q}_2\}$.

1. Obtenció d'una base ortogonal (projecteu $\vec{u}_2$ sobre l'ortogonal de $\vec{u}_1$):

$$\vec{v}_1 = \vec{u}_1$$

$$\vec{v}_2 = \vec{u}_2 - \alpha_{1,2}\vec{v}_1 \qquad \left(\alpha_{1,2} = \frac{\vec{v}_1\cdot\vec{u}_2}{\vec{v}_1\cdot\vec{v}_1}\right)$$

2. Normalització (dividiu cada vector per la seua norma):

$$\vec{q}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1, \qquad \vec{q}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2, \qquad \dots \qquad \vec{q}_p = \frac{1}{\left\|\vec{v}_p\right\|}\vec{v}_p$$

EXEMPLE 17.1.

Obteniu una base ortonormal del subespai $F = \langle\vec{u}_1, \vec{u}_2\rangle$, on $\vec{u}_1 = (-2,-2,-1)$ i $\vec{u}_2 = (1,4,-1)$.

Obtenim una base ortogonal:

- El primer vector no l'hem de modificar:

$$\boxed{\vec{v}_1 = \vec{u}_1 = (-2,-2,-1)}$$

- Calculem $\alpha_{1,2}$ i la norma de $\vec{v}_1$:

$$\vec{v}_1\cdot\vec{v}_1 = (-2,-2,-1)\cdot(-2,-2,-1) = 9 \qquad \|\vec{v}_1\| = \sqrt{\vec{v}_1\cdot\vec{v}_1} = 3$$

$$\vec{v}_1\cdot\vec{u}_2 = (-2,-2,-1)\cdot(1,4,-1) = -9 \qquad \alpha_{1,2} = \frac{\vec{v}_1\cdot\vec{u}_2}{\vec{v}_1\cdot\vec{v}_1} = -1$$

- Càlcul del vector $\vec{v}_2$:

$$\boxed{\vec{v}_2 = \vec{u}_2 - \alpha_{1,2}\vec{v}_1 = (-1,2,-2)}$$

- Calculem la norma de $\vec{v}_2$:

$$\vec{v}_2\cdot\vec{v}_2 = (-1,2,-2)\cdot(-1,2,-2) = 9 \qquad \|\vec{v}_2\| = \sqrt{\vec{v}_2\cdot\vec{v}_2} = 3$$

$\{\vec{v}_1, \vec{v}_2\}$ és una base ortogonal de F.

Normalització: Per convertir els vectors en unitaris hem de dividir-los per les seues normes:

$$\boxed{\vec{q}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}(-2,-2,-1)}$$

$$\boxed{\vec{q}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{3}(-1,2,-2)}$$

$B_2 = \{\vec{q}_1, \vec{q}_2\}$ és una base ortonormal de F. □

Suposem ara que el subespai F és tridimensional; partirem d'una base $B_1 = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ i, com en el cas anterior, construirem en primer lloc una base ortogonal $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ i, després, una altra, $B_2 = \{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$, ortonormal.

Per trobar la base ortogonal, els dos vectors $\vec{v}_1$ i $\vec{v}_2$ els construïm de la mateixa manera:

1. El vector $\vec{v}_1$ va a ser el mateix $\vec{u}_1$:
$$\vec{v}_1 = \vec{u}_1$$

2. Com a $\vec{v}_2$ prendrem la projecció de $\vec{u}_2$ sobre l'ortogonal de la recta generada per $\vec{v}_1$:
$$\vec{v}_2 = p_{\langle\vec{v}_1\rangle^\perp}(\vec{u}_2) = \vec{u}_2 - \alpha_{1,2}\vec{v}_1, \qquad \alpha_{1,2} = \frac{\vec{v}_1 \cdot \vec{u}_2}{\vec{v}_1 \cdot \vec{v}_1}$$

3. Per acabar d'obtenir una base de F, hi haurem d'afegir un tercer vector $\vec{v}_3$, que es trobe en F i siga ortogonal a $\vec{u}_1$ i $\vec{u}_2$. Això ho podem aconseguir projectant ortogonalment el vector $\vec{u}_3$ sobre l'ortogonal de $\langle\vec{u}_1, \vec{u}_2\rangle$.
$$\vec{v}_3 = p_{\langle\vec{v}_1,\vec{v}_2\rangle^\perp}(\vec{u}_2) = \vec{u}_3 - \alpha_{1,3}\vec{v}_1 - \alpha_{2,3}\vec{v}_2, \qquad \alpha_{1,3} = \frac{\vec{v}_1 \cdot \vec{u}_3}{\vec{v}_1 \cdot \vec{v}_1}, \quad \alpha_{2,3} = \frac{\vec{v}_2 \cdot \vec{u}_3}{\vec{v}_2 \cdot \vec{v}_2}$$

Llavors, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ és una base ortogonal de F, així que, per convertir-la en ortonormal, dividim cada vector per la seua norma. :

$$\vec{q}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 \qquad \vec{q}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 \qquad \vec{q}_3 = \frac{1}{\|\vec{v}_3\|}\vec{v}_3$$

El conjunt $B_2 = \{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ és una base ortonormal de F.

Exemple 17.2.

Obteniu una base ortonormal del subespai $F = \langle\vec{u}_1, \vec{u}_2, \vec{u}_3\rangle$, on $\vec{u}_1 = (1,2,0,2)$, $\vec{u}_2 = (0,9,6,0)$ i $\vec{u}_3 = (0,0,6,9)$.

Obtenim una base ortogonal:

– El primer vector no l'hem de modificar:

$$\boxed{\vec{v}_1 = \vec{u}_1 = (1,2,0,2)}$$

– Calculem $\alpha_{1,2}$, $\alpha_{1,3}$ i la norma de $\vec{v}_1$:

$$\vec{v}_1 \cdot \vec{v}_1 = (1,2,0,2) \cdot (1,2,0,2) = 9 \qquad \|\vec{v}_1\| = \sqrt{\vec{v}_1 \cdot \vec{v}_1} = 3$$
$$\vec{v}_1 \cdot \vec{u}_2 = (1,2,0,2) \cdot (0,9,6,0) = 18 \qquad \alpha_{1,2} = \frac{\vec{v}_1 \cdot \vec{u}_2}{\vec{v}_1 \cdot \vec{v}_1} = 2$$
$$\vec{v}_1 \cdot \vec{u}_3 = (1,2,0,2) \cdot (0,0,6,9) = 18 \qquad \alpha_{1,3} = \frac{\vec{v}_1 \cdot \vec{u}_3}{\vec{v}_1 \cdot \vec{v}_1} = 2$$

– Càlcul del vector $\vec{v}_2$:
$$\boxed{\vec{v}_2 = \vec{u}_2 - \alpha_{1,2}\vec{v}_1 = (-2,5,6,-4)}$$

- Calculem $\alpha_{2,3}$ i la norma de $\vec{v}_2$:

$$\vec{v}_2 \cdot \vec{v}_2 = (-2,5,6,-4)\cdot(-2,5,6,-4) = 81 \qquad \|\vec{v}_2\| = \sqrt{\vec{v}_2\cdot\vec{v}_2} = 9$$

$$\vec{v}_2 \cdot \vec{u}_3 = (-2,5,6,-4)\cdot(0,0,6,9) = 0 \qquad \alpha_{2,3} = \frac{\vec{v}_2\cdot\vec{u}_3}{\vec{v}_2\cdot\vec{v}_2} = 0$$

- Càlcul del vector $\vec{v}_3$:

$$\boxed{\vec{v}_3 = \vec{u}_3 - \alpha_{1,3}\vec{v}_1 - \alpha_{2,3}\vec{v}_2 = (-2,-4,6,5)}$$

- Calculem la norma de $\vec{v}_3$:

$$\vec{v}_3 \cdot \vec{v}_3 = (-2,-4,6,5)\cdot(-2,-4,6,5) = 81 \qquad \|\vec{v}_3\| = \sqrt{\vec{v}_3\cdot\vec{v}_3} = 9$$

$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ és una base ortogonal de F.

Normalització: Per convertir els vectors en unitaris hem de dividir-los per les seues normes:

$$\boxed{\vec{q}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \frac{1}{3}(1,2,0,2)}$$

$$\boxed{\vec{q}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \frac{1}{9}(-2,5,6,-4)}$$

$$\boxed{\vec{q}_3 = \frac{1}{\|\vec{v}_3\|}\vec{v}_3 = \frac{1}{9}(-2,-4,6,5)}$$

$B_2 = \{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ és una base ortonormal de F. □

17.2.2. La factorització QR amb tres vectors

Les transformacions que hem fet al darrer exemple es poden interpretar com una sèrie d'operacions elementals:

- Al vector $\vec{u}_2$ li hem restat un múltiple de $\vec{v}_1$
- Al vector $\vec{u}_3$ li hem restat un múltiple de $\vec{v}_1$ i, tot seguit, un múltiple de $\vec{v}_2$
- Hem dividit cada vector per la seua norma.

Si A és la matriu associada a la base B_1, és a dir, $\mathsf{A} = \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \vec{u}_3\end{bmatrix}$, llavors l'algorisme de Gram-Schmidth consisteix a fer unes quantes operacions elelemntals sobre les columnes d'aquesta matriu. En aquest sentit, aquest algorisme és anàleg al de Gauss (però, en comptes de fer operacions elementals amb les files, les fem amb les columnes i, en comptes de cercar zeros i uns, cerquem vectors ortogonals i unitaris).

En l'exemple 17.2., la matriu seria $\mathsf{A} = \begin{bmatrix}1&0&0\\2&9&0\\0&6&6\\2&0&9\end{bmatrix}$ i el procès, esquemàticament, és el següent (les operacions indicades es fan per columnes):

Obtenció d'una base ortogonal:

$$\mathsf{A} = \begin{bmatrix}1&0&0\\2&9&0\\0&6&6\\2&0&9\end{bmatrix} \xrightarrow[\mathsf{E}_{1,2}(-2)]{} \begin{bmatrix}1&-2&0\\2&5&0\\0&6&6\\2&-4&9\end{bmatrix} \xrightarrow[\mathsf{E}_{1,3}(-2)]{} \begin{bmatrix}1&-2&-2\\2&5&-4\\0&6&6\\2&-4&5\end{bmatrix} \xrightarrow[\mathsf{E}_{2,3}(0)]{} \begin{bmatrix}1&-2&-2\\2&5&-4\\0&6&6\\2&-4&5\end{bmatrix}$$

Normalització:

$$\begin{bmatrix}1&-2&-2\\2&5&-4\\0&6&6\\2&-4&5\end{bmatrix} \xrightarrow[\mathsf{E}_1\left(\frac{1}{3}\right)\mathsf{E}_2\left(\frac{1}{9}\right)\mathsf{E}_3\left(\frac{1}{9}\right)]{} \begin{bmatrix}1/3&-2/9&-2/9\\2/3&5/9&-4/9\\0&6/9&6/9\\2/3&-4/9&5/9\end{bmatrix} = \mathsf{Q}$$

Ací, la matriu Q té les columnes ortonormals. I, com que aquesta matriu ha esta obtinguda fent operacions elementals sobre les columnes de A, resulta que

$$\mathsf{A} = \mathsf{QR}$$

on R és el producte de les operacions elementals inverses de les que hem aplicat per trobar Q.

Aquesta matriu R és invertible (perquè és un producte de matrius elementals i triangular superior (perquè totes les matrius elementals que hem aplicat ho són). Per a calcular-la, com que

$$\mathsf{Q} = \mathsf{A}\,\underbrace{\mathsf{E}_{1,2}(-2)\mathsf{E}_{1,3}(-2)\mathsf{E}_{2,3}(0)\mathsf{E}_1\left(\frac{1}{3}\right)\mathsf{E}_2\left(\frac{1}{9}\right)\mathsf{E}_3\left(\frac{1}{9}\right)}_{\mathsf{R}^{-1}}$$

tindrem

$$\mathsf{R} = \mathsf{E}_3\,(9)\,\mathsf{E}_2\,(9)\,\mathsf{E}_1\,(3)\,\mathsf{E}_{2,3}(0)\mathsf{E}_{1,3}(2)\mathsf{E}_{1,2}(2)$$

Així que, aplicant a les files de la identitat les operacions elementals inverses de les que apliquem a A, obtenim

$$\mathsf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{1,2}(2)} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{1,3}(2)} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{\mathsf{E}_{2,3}(0)} \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\xrightarrow{\mathsf{E}_3(9)\mathsf{E}_2(9)\mathsf{E}_1(3)} \begin{bmatrix} 3 & 6 & 6 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \mathsf{R}$$

☞ En general, si els vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$ són linealment independents llavors, la matriu $\mathsf{A} = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$ es pot factoritzar d'aquesta manera:

$$\mathsf{A} = \underbrace{\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \vec{q}_3 \end{bmatrix}}_{\mathsf{Q}} \underbrace{\begin{bmatrix} \|v_1\| & \alpha_{1,2}\,\|v_1\| & \alpha_{1,3}\,\|v_1\| \\ 0 & \|v_2\| & \alpha_{2,3}\,\|v_1\| \\ 0 & & \|v_3\| \end{bmatrix}}_{\mathsf{R}}$$

Aquest o qualsevol altre producte $\mathsf{A} = \mathsf{QR}$ on Q té les columnes ortonormals i R és triangular superior i invertible és una *factorització QR de la matriu* A.

17.2.3. El cas general

Definició 17.1.

Qualsevol factorització de la matriu A *com a producte d'una matriu amb les columnes ortonormals i una altra de triangular superior invertible és una descomposició o factorització QR de* A.

Els dos exemples anteriors són suficients per a convèncer-nos que, si B_1 és una base d'un subespai F de $\mathbb{K}^n$, l'algorisme següent ens proporciona una base ortonormal de F, i una factorització QR de la matriu $\mathsf{A} = \mathsf{M}_{B_1}$.

Mètode d'ortonormalització de Gram-Schmidt

Si $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ és una base del subespai F llavors,
aquest algorisme construeix una base ortonormal de F, $B_2 = \{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_p\}$.

1. Obtenció d'una base ortogonal (projecteu cada vector sobre l'ortogonal dels anteriors):

$$\begin{aligned}
\vec{v}_1 &= \vec{u}_1 \\
\vec{v}_2 &= \vec{u}_2 - \alpha_{1,2}\vec{v}_1 && \left(\alpha_{1,2} = \frac{\vec{v}_1 \cdot \vec{u}_2}{\vec{v}_1 \cdot \vec{v}_1}\right) \\
\vec{v}_3 &= \vec{u}_3 - \alpha_{1,3}\vec{v}_1 - \alpha_{2,3}\vec{v}_2 && \left(\alpha_{i,3} = \frac{\vec{v}_i \cdot \vec{u}_3}{\vec{v}_i \cdot \vec{v}_i}, 1 \le i \le 2\right) \\
&\dots \\
\vec{v}_p &= \vec{u}_p - \alpha_{1,p}\vec{v}_1 - \alpha_{2,p}\vec{v}_2 - \cdots - \alpha_{p-1,p}\vec{v}_{p-1} && \left(\alpha_{i,p} = \frac{\vec{v}_i \cdot \vec{u}_p}{\vec{v}_i \cdot \vec{v}_i}, 1 \le i \le p-1\right)
\end{aligned}$$

2. Normalització (dividiu cada vector per la seua norma):

$$\vec{q}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1, \qquad \vec{q}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2, \qquad \dots \qquad \vec{q}_p = \frac{1}{\left\|\vec{v}_p\right\|}\vec{v}_p$$

Factorització QR deduïda de l'algorisme d'ortonormalització de Gram-Schmidt

Si les columnes de la matriu $\mathsf{A} = \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_p\end{bmatrix}$ són linealment independents, aleshores

$$\mathsf{A} = \mathsf{QR}$$

on

$$\mathsf{Q} = \begin{bmatrix}\vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_p\end{bmatrix}$$

és una matriu que té les columnes ortonormals i

$$\mathsf{R} = \begin{bmatrix}
\|\vec{v}_1\| & \alpha_{1,2}\|\vec{v}_1\| & \alpha_{1,3}\|\vec{v}_1\| & \cdots & \alpha_{1,p}\|\vec{v}_1\| \\
0 & \|\vec{v}_2\| & \alpha_{2,3}\|\vec{v}_2\| & \cdots & \alpha_{2,p}\|\vec{v}_2\| \\
0 & 0 & \|\vec{v}_3\| & \cdots & \alpha_{3,p}\|\vec{v}_1\| \\
\cdots & \cdots & \cdots & & \cdots \\
0 & 0 & 0 & \cdots & \left\|\vec{v}_p\right\|
\end{bmatrix}$$

és una matriu triangular superior invertible.

☞ El procés anterior funciona perquè les columnes de la matriu A són la base d'un subespai, es a dir, que les columnes de la matriu A són linealment independents.

☞ Si A és una matriu $m \times p$, llavors Q és també $m \times p$ i té les columnes ortonormals i la matriu R és triangular superior, $p \times p$, i invertible.

☞ El mètode de Gram-Schmidt és anàleg a l'algorisme de Gauss, en el sentit que es pot interpretar com una sèrie d'operacions elementals.

I, de la mateixa manera que l'algorisme de Gauss ens proporciona la factorització LU, el de Gram-Schmidt ens dóna una factorització QR.

Cal tenir en compte que la descomposició QR no és única, i que el mètode de Gram-Schmidt no és l'única estratègia que ens permet trobar una factorització QR (de fet, en la pràctica, els sistemes de càlcul científic no fan servir aquest algorisme sinó d'altres que són numèricament més estables).

17.3. Aplicacions de la descomposició QR

Acabarem la lliçó veient com se simplifica la resolució dels problemes bàsics quan coneixem una factorització QR de la matriu implicada. Suposem que la matriu $\mathsf{A} = \mathsf{M}_B$ és la matriu associada a una base B del subespai F i que $\mathsf{A} = \mathsf{QR}$ és una factorització QR de la matriu A.

17.3.1. Projecció ortogonal

La projecció ortogonal del vector $\vec{b}$ sobre el subespai F depén únicament de F; per tant, aquesta projecció no depén de la base de F que escollim. Com que les columnes de Q són una base ortonormal de F, podem aplicar la fórmula (15.3) (pàgina 131):

$$P_F(\vec{b}) = (\vec{q}_1 \cdot \vec{b})\vec{q}_1 + (\vec{q}_2 \cdot \vec{b})\vec{q}_2 + \cdots + (\vec{q}_p \cdot \vec{b})\vec{q}_p$$

17.3.2. Mínims quadrats

Per trobar la solució per mínims quadrats del sistema $\mathsf{A}\vec{x} = \vec{b}$ hem de resoldre el sistema $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$. Ara bé,

$$\mathsf{A}^t\mathsf{A} = \mathsf{R}^t\mathsf{Q}^t\mathsf{QR} = \mathsf{R}^t\mathsf{R}$$

(perquè $\mathsf{Q}^t\mathsf{Q} = \mathrm{I}$) de manera que $\mathsf{A}^t\mathsf{A}\vec{x} = \mathsf{A}^t\vec{b}$ és el mateix que

$$\mathsf{R}^t\mathsf{R}\vec{x} = \mathsf{R}^t\mathsf{Q}^t\vec{b}$$

O bé, com que R^t és invertible,

$$\mathsf{R}\vec{x} = \mathsf{Q}^t\vec{b}$$

i l'únic que hem de fer és resoldre un sistema triangular. D'altra banda, l'error de mínims quadrats és $e = \left\|\vec{b} - \mathsf{A}\vec{x}\right\|$, és a dir,

$$e = \|\vec{b} - \underbrace{\mathsf{QR}}_{\mathsf{A}}\underbrace{\mathsf{R}^{-1}\mathsf{Q}^t\vec{b}}_{\vec{x}}\| = \left\|\vec{b} - \mathsf{QQ}^t\vec{b}\right\|$$

17.4. Exercicis

17.4.1. Operacions elementals per files i columnes

Exercici 17.1. La matriu N l'hem obtinguda fent les operacions elementals següents sobre la matriu $\mathsf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -2 & -3 \end{bmatrix}$:
(a) Hem sumat la columna primera a la tercera; (b) Hem restat la columna segona a la tercera; (c) Hem intercanviat les dues files; (d) A la primera fila li hem sumat el doble de la segona. Calculeu la matriu N i trobeu dues matrius invertibles B i C de manera que $\mathsf{N} = \mathsf{BAC}$.

17.4.2. La factorització QR

Exercici 17.2. Trobeu una factorització QR de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$.

Exercici 17.3. Trobeu la projecció del vector $\vec{b} = (3, 5, 7, -3)$ sobre l'espai columna de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$.

Exercici 17.4. Resoleu el problema de mínims quadrats $\mathsf{A}\vec{x} = \vec{b}$ on $\mathsf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$, fent servir la descomposició QR de l'exercici anterior. Determineu també l'error de mínims quadrats en aquest problema.

Capítol 5

Espais vectorials i aplicacions lineals

Continguts

18. Espais vectorials

En aquesta lliçó introduïm la teoria dels espais vectorials. Fins ara només coneixem els espais $\mathbb{R}^n$; ara generalitzarem les idees que hem estat estudiant i trobarem que molts conjunts que ja coneixem tenen l'estructura d'espais vectorials.

18.1. Espais vectorials

L'eina fonamental en tot el treball amb sistemes d'equacions, matrius i subespais de $\mathbb{R}^n$ han estat les combinacions lineals. Per això, ara que pretenem generalitzar les idees a altres conjunts, treballarem amb objectes que ens permeten fer *combinacions lineals*; en altres paraules, hem de fer servir conjunts en els quals puguem fer sumes i multiplicacions per escalars. Entre els conjunts que ens ho permeten hi ha les matrius, els polinomis, les successions, les funcions...

Les dues definicions importants són aquestes:

Definició 18.1.

Un espai vectorial és un conjunt E, els elements del qual s'anomenen vectors, en el qual hi ha definides dues operacions: la suma de vectors i el producte d'un escalar per un vector, que tenen les propietats següents:

Propietats de la suma

Commutativa $\vec{u}_1 + \vec{u}_2 = \vec{u}_2 + \vec{u}_1$

Associativa $(\vec{u}_1 + \vec{u}_2) + \vec{u}_3 = \vec{u}_1 + (\vec{u}_2 + \vec{u}_3)$

Element neutre *Existeix un vector zero, $\vec{0}$ amb la propietat que $\vec{u} + \vec{0} = \vec{u} \quad \forall \vec{u} \in E$*

Elements oposats (o simètrics) *Cada vector $\vec{u}$ té un vector oposat, $-\vec{u}$, de manera que $\vec{u} + (-\vec{u}) = \vec{0}$*

Propietats del producte escalar-vector

Associativa $(\alpha_1\alpha_2)\vec{u} = \alpha_1(\alpha_2\vec{u})$

Distributiva respecte a la suma d'escalars $(\alpha_1 + \alpha_2)\vec{u} = \alpha_1\vec{u} + \alpha_2\vec{u}$

Distributiva respecte a la suma de vectors $\alpha(\vec{u}_1 + \vec{u}_2) = \alpha\vec{u}_1 + \alpha\vec{u}_2$

Element neutre $1\vec{u} = \vec{u}$

Definició 18.2.

Un vector $\vec{u}$ és combinació lineal dels vectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$ si existeixen escalars $\alpha_1, \alpha_2, \ldots, \alpha_p$ de manera que

$$\vec{u} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \ldots \alpha_p\vec{u}_p$$

Segons que treballem amb escalars reals o complexos parlarem d'espais vectorials reals o d'espais vectorials complexos i, quan no siga necessari especificar quin és el conjunt d'escalars, parlarem d'espais vectorials *sobre* $\mathbb{K}$.

A partir de la definició es poden deduir totes les propietats del càlcul amb vectors. La més interessant fa referència al vector zero i a l'escalar zero: el producte d'un escalar per un vector és zero, només, quan un dels dos és zero.

Propietats 18.1.

Suposem, que $\vec{u}$ és un vector i que α és un escalar. llavors,

(a) $0\vec{u} = \vec{0}$

(b) $\alpha\vec{0} = \vec{0}$

(c) Si $\alpha\vec{u} = \vec{0}$, llavors $\alpha = 0$ o $\vec{u} = \vec{0}$ $\square$

18.1.1. Exemples

Evidentment, $\mathbb{K}^n$ és un espai vectorial. Però hi ha molts altres exemples interessants. Vegem-ne alguns.

L'espai nul El conjunt $O = \{\vec{0}\}$ (format per un sol vector) és un espai vectorial. Evidentment, l'única manera de definir-hi les operacions és aquesta:

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

i totes les propietats de la definició es compleixen, perquè el resultat de qualsevol operació és $\vec{0}$.

Les matrius $m \times n$ En el conjunt $\mathcal{M}_{m\times n}$ hi ha definides les dues operacions, suma de matrius i producte d'un escalar per una matriu, i es compleixen totes les propietats de la definició. En conseqüència, aquest conjunt és un espai vectorial.[1]

El polinomis de la indeterminada x Representem com $\mathbb{K}[x]$ el conjunt de tots els polinomis de la forma

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

(on els coeficients són elements de $\mathbb{K}$). En aquest conjunt també tenim les operacions convenients:

$$(a_n x^n + \cdots + a_1 x + a_0) + (b_n x^n + \cdots + b_1 x + b_0) = (a_n + b_n)\, x_n + \cdots + (a_1 + b_1)\, x + (a_0 + b_0)$$
$$\alpha\,(a_n x^n + \cdots + a_1 x + a_0) = \alpha a_n x^n + \cdots + \alpha a_1 x + \alpha a_0$$

i es poden comprovar fàcilment les vuit propietats de la definició. Per tant, $\mathbb{K}[x]$ és un espai vectorial.

El conjunt de totes les successions En aquest espai, les operacions són

$$\{a_n : n = 1, 2, \ldots\} + \{b_n : n = 1, 2, \ldots\} = \{a_n + b_n : n = 1, 2, \ldots\}$$
$$\alpha\,\{a_n : n = 1, 2, \ldots\} = \{\alpha a_n : n = 1, 2, \ldots\}$$

Les funcions contínues en un interval El conjunt $C_0(I)$ format per totes les funcions contínues en l'interval I també és un espai vectorial (recordem que la suma de dues funcions contínues i el producte d'un nombre per una funció contínua també són funcions contínues.

18.2. Subespais d'un espai vectorial

Els subespais d'un espai vectorial E es defineixen de la mateixa manera que els subespais de $\mathbb{R}^n$.

Definició 18.3.

Un subespai de l'espai vectorial E és un subconjunt $F \subset E$ que té aquestes dues propietats:

(a) El vector $\vec{0}$ és un element de F:

$$\vec{0} \in F$$

(b) Si fem qualsevol combinació lineal amb dos elements de F el vector que en resulta també és un element de F:

$$\text{Si } u_1, u_2 \in F \text{ llavors } \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 \in F$$

(per a dos escalars α_1 i α_2 qualssevol).

De manera equivalent, es pot definir un subespai com un subconjunt de E que, amb les mateixes operacions definides en E, és també un espai vectorial.

[1] El conjunt de totes les matrius *no és* un espai vectorial, perquè les matrius de dimensions diferents no poden sumar-se. El que tenim ací són diversos espais: el de les matrius 2×2, el de les matrius 4×5, etc.

18.2.1. Exemples

L'espai E i el conjunt $\{\vec{0}\}$ són, evidentment, subespais de E. D'aquests subespais en direm *impropis*.

Ací veurem alguns exempes interessants de subespais dels espais que hem definit a la secció anterior.

Subespais de matrius En l'espai de les matrius quadrades $n \times n$, els conjunts de les matrius diagonals, de les triangulars superiors, de les triangulars inferiors, el de les matrius simètriques i el de les antisimètriques són subespais.

Exemple 18.1.

> Proveu que el conjunt $\mathscr{S}_n$ de les matrius simètriques $n \times n$ és un subespai de l'espai $\mathscr{M}_{n\times n}$

És clar que $\mathscr{S}_n \subset \mathscr{M}_{n\times n}$. A més a més,

(a) La matriu O és un element de $\mathscr{S}_n$, perquè és una matriu simètrica.

(b) Si A_1 i A_2 són matrius simètriques $n \times n$, llavors

$$(\alpha_1\mathsf{A}_1 + \alpha_2\mathsf{A}_2)^t = \alpha_1\mathsf{A}_1^t + \alpha_2\mathsf{A}_2^t = \alpha_1\mathsf{A}_1 + \alpha_2\mathsf{A}_2$$

així que $\alpha_1\mathsf{A}_1 + \alpha_2\mathsf{A}_2$ també és una matriu simètrica. □

Subespais de polinomis Representem com $\mathbb{K}_n[x]$ el conjunt dels polinomis de grau, com a molt, n. Aquest conjunt és un subespai de $\mathbb{K}[x]$, perquè la suma de dos polinomis i el producte d'un escalar per un polinomi no augmenten el grau dels polinomis.

Un polinomi és *parell* si només conté potències parelles de la indeterminada (per exemple, el polinomi $p(x) = 1 - x^2 + x^4$ és parell); de manera anàloga, un polinomi és senar si només conté potències senars de la indeterminada (per exemple, el polinomi $p(x) = x + 2x^3 + x^7$ és senar). Els conjunt de tots els polinomis parells és un subespai de $\mathbb{K}[x]$.

Subespais de successions El conjunt de les successions convergents és un subespai del de totes les successions (recordem que la suma de dues successions convergents és convergent, i que també ho és el producte d'un nombre per una successió convergent).

També són subespais els conjunts de les successions que convergeixen a zero i el de les succesions $\{a_n\}$ per a les quals la sèrie $\sum a_n$ és convergent.

Subespais de funcions Representem com $C_1(I)$ el conjunt de les funcions que són derivables en l'interval I amb derivada contínua; $C_2(I)$ és el conjunt de les funcions derivables dues vegades en I amb derivada segona contínua (i anàlogament es defineixen els conjunts $C_3(I)$, $C_4(I)$...). Tots aquests conjunts són subespais de $C_0(I)$.

18.3. Dependència lineal

Una *relació de dependència* entre els vectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p$ és una igualtat de la forma

$$\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_p\vec{u}_p = \vec{0}$$

on $\alpha_1, \alpha_2, \ldots, \alpha_p$ són escalars. És clar que, elegint tots els escalars iguals a zero obtenim la relació de dependència

$$0\vec{u}_1 + 0\vec{u}_2 + \cdots + 0\vec{u}_p = \vec{0}$$

Aquesta és la relació de dependència *trivial*.

Definició 18.4.

> *Un conjunt de vectors S és linealment dependent si existeix alguna relació de dependència no trivial entre vectors de S. En cas contrari, diem que S és linealment independent.*

☞ Un conjunt que continga el vector zero sempre serà linealment dependent.

☞ Si S és un conjunt que només conté un element, i aquest element no és el zero, llavors S és linealment independent.

EXEMPLE 18.2.

Estudieu si el conjunt de polinomis $S = \{1, x-1, (x-1)^2\}$ és linealment dependent o independent.

Suposem que

$$\alpha_1 + \alpha_2(x-1) + \alpha_3(x-1)^2 = 0$$

Derivant aquesta igualtat obtenim

$$\alpha_2 + 2\alpha_3(x-1) = 0$$

i, tornant a derivar,

$$2\alpha_3 = 0$$

D'aquestes tres igualtats es dedueix que $\alpha_1 = \alpha_2 = \alpha_3 = 0$, de manera que S és linealment independent.

EXEMPLE 18.3.

Estudieu si el conjunt de matrius 2×2

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \right\}$$

és linealment dependent o independent.

Com que

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

el conjunt S és linealment dependent.

PROPIETAT 18.2.

Suposem que S és un conjunt que conté, almenys, dos elements. Aleshores,
S és linealment dependent si i només si existeix algun vector en S que és combinació lineal dels altres. □

En l'exemple anterior, la tercera matriu és la suma de les altres dues; per això S és linealment dependent.

18.4. EXERCICIS

EXERCICI 18.1. Estudieu si els conjunts F següents són o no subespais dels espais vectorials E que s'indiquen en cada cas.

(a) F és el conjunt de les matrius quadrades 2×2 A per a les quals $a_{11} = a_{22}$ i $a_{12} = 2a_{21}$. $E = \mathcal{M}_{2\times 2}$.

(b) F és el conjunt de les matrius quadrades 2×2 A per a les quals $a_{11} = a_{22} = 0$ i $a_{12} = a_{21} = 2$. $E = \mathcal{M}_{2\times 2}$.

(c) F és el conjunt dels polinomis de grau parell. $E = \mathbb{R}[x]$.

(d) F és el conjunt dels polinomis de la forma $p(x) = ax + bx^2$ on a i b són constants qualssevol. $E = \mathbb{R}[x]$.

(e) F és el conjunt dels polinomis de la forma $p(x) = 2+2ax$ on a és una constant qualsevol. $E = \mathbb{R}_1[x]$.

EXERCICI 18.2. Estudieu si les successions

$$\{a_n\} = \{1, 1, 1, \dots, 1, \dots\}$$
$$\{b_n\} = \{1, 2, 4, \dots, 2^n \dots\}$$
$$\{c_n\} = \{1, 3, 9, \dots, 3^n, \dots\}$$

són linealment independents.

Exercici 18.3. Estudieu si els conjunts de polinomis

$$\mathscr{P}_1 = \{x, 1 + 2x, 1 - 3x, 1 - 4x^2\}, \qquad \mathscr{P}_2 = \{1 + x, 1 + 2x, x - x^2\}$$

són linealment independents. En cas que no ho siguen, trobeu alguna relació de dependència no trivial entre els seus elements.

Exercici 18.4. Estudieu si el conjunt de matrius 2×2,

$$\mathscr{A} = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

és linealment independent. És possible *ampliar* aquest conjunt, afegint una matriu A, de manera que el nou conjunt continue essent linealment independent?

Exercici 18.5. Estudieu si el conjunt de funcions $\{\cos x, \sin x, \cos 2x, \sin 2x\}$ és linealment independent.

19. Base d'un espai vectorial

Ara traslladarem el concepte de base als espais vectorials arbitraris.

19.1. Embolcall lineal d'un conjunt de vectors i conjunts generadors

Si S és un conjunt de vectors, l'*embolcall lineal* de S és el conjunt $\langle S\rangle$ de totes les combinacions lineals dels elements de S.

Per exemple, l'embolcall lineal de

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

és el conjunt de les matrius triangulars superiors 2×2, perquè

$$\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} = a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

☞ L'embolcall lineal $\langle S\rangle$ és el subespai més petit que conté a S.

Definició 19.1.

Un conjunt S és generador de l'espai E si $\langle S\rangle = E$.

Exemple 19.1.

Estudieu si el conjunt $S = \{1, 1+x, 1-x, x-x^2\}$ és generador de $\mathbb{K}_2[x]$.

Es tracta de veure si qualsevol polinomi de grau, com a molt, 2 és combinació lineal dels vectors de S, és a dir, si, per a qualsevol polinomi $p(x) = a_0 + a_1x + a_2x^2$, existeixen escalars α_1, α_2, α_3, α_4, de manera que

$$a_0 + a_1x + a_2x^2 = \alpha_1 + \alpha_2(1+x) + \alpha_3(1-x) + \alpha_4(x-x^2)$$

o, equivalentment,

$$a_0 + a_1x + a_2x^2 = (\alpha_1 + \alpha_2 + \alpha_3) + (\alpha_2 - \alpha_3 + \alpha_4)x - \alpha_4x^2$$

Igualant els coeficients, tindrem

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 \qquad &= a_0 \\ \alpha_2 - \alpha_3 + \alpha_4 &= a_1 \\ -\alpha_4 &= a_2 \end{aligned}$$

Sistema d'equacions lineals que podem escriure en forma matricial com

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

Aquest sistema és compatible i, en conseqüència, S és generador de $\mathbb{K}_2[x]$.

Definició 19.2.

Un espai vectorial E és de dimensió finita (o finitament generat) si hi ha un conjunt finit S de manera que $\langle S\rangle = E$.

EXEMPLES 19.2.

Els espais $\mathbb{R}^n$ són de dimensió finita, perquè ja en coneixem diversos conjunts generadors finits (per exemple, la base canònica, o qualsevol altra base).
L'espai $\mathcal{M}_{2\times 3}$ també és finitament generat, perquè

$$\left\langle \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\} \right\rangle = \mathcal{M}_{2\times 3}$$

L'espai $\mathbb{R}_3[x]$ és finitament generat, perquè

$$\left\langle \left\{ 1, x, x^2, x^3 \right\} \right\rangle = \mathbb{R}_3[x].$$

En canvi, no és finitament generat l'espai $\mathbb{R}[x]$, perquè si ho fora, tindríem un conjunt finit de polinomis

$$S = \{p_1(x), p_2(x), \dots, p_r(x)\}$$

que generaria l'espai. Ara bé, si n és el grau màxim dels polinomis de S, llavors és evident que cap polinomi de grau $n+1$ en serà combinació lineal. En conseqüència, no és possible que S siga generador.

19.2. BASES

DEFINICIÓ 19.3.

Una base de l'espai vectorial E és un conjunt de vectors B que és linealment independent i generador de E.

EXEMPLE 19.3.

Estudieu si el conjunt $S = \{1, 1+x, 1-x, x-x^2\}$ és base de $\mathbb{K}_2[x]$.

En un exemple anterior hem provat que aquest conjunt és generador de $\mathbb{K}_2[x]$. Però no és linealment independent, perquè

$$-1 + \frac{1}{2}(1+x) + \frac{1}{2}(1-x) = 0$$

Per tant, no és una base.

EXEMPLE 19.4.

Estudieu si el conjunt $S = \{1+x, 1-x, x-x^2\}$ és base de $\mathbb{K}_2[x]$.

Es tracta d'estudiar si aquest conjunt és independent i generador. Perquè siga generador, qualsevol vector n'ha de ser combinació lineal:

$$a_0 + a_1 x + a_2 x^2 = \alpha_1(1+x) + \alpha_2(1-x) + \alpha_3(x-x^2)$$

Reordenant segons les potències de x,

$$a_0 + a_1 x + a_2 x^2 = (\alpha_1 + \alpha_2) + (\alpha_1 - \alpha_2 + \alpha_3)x - \alpha_3 x^2$$

Igualem coeficents,

$$\begin{aligned} \alpha_1 + \alpha_2 \qquad &= a_0 \\ \alpha_1 - \alpha_2 + \alpha_3 &= a_1 \\ -\alpha_3 &= a_2 \end{aligned}$$

En forma matricial,

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

El rang de la matriu d'aquest sistema lineal és 3. Per tant, el sistema és compatible i el conjunt S és generador.

Per decidir si es tracta d'un conjunt linealment independent hem d'estudiar l'equació

$$\alpha_1(1+x)+\alpha_2(1-x)+\alpha_3(x-x^2)=0$$

Que és equivalent al sistema lineal

$$\begin{bmatrix}1 & 1 & 0\\ 1 & -1 & 1\\ 0 & 0 & -1\end{bmatrix}\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

I, com que aquest sistema és determinat, l'única solució que té és $\alpha_1=\alpha_2=\alpha_3=0$, de manera que S és independent. En definitiva, S és una base de $\mathbb{K}_2[x]$.

19.2.1. Coordenades d'un vector respecte a una base

Com en el cas de $\mathbb{R}^n$, la diferència entre un conjunt generador qualsevol i una base és que només hi ha una manera d'expressar un vector com a combinació lineal dels vectors de la base.

Si el conjunt B és una base de l'espai E, llavors qualsevol vector de E s'expressa com a combinació lineal dels elements de B *de forma única*.

Aquesta propietat ens permet definir les coordenades d'un vector respecte a una base.

Definició 19.4.

Si $B=\{\vec{u}_1,\vec{u}_2,\dots,\vec{u}_n\}$ és una base de l'espai E i el vector $\vec{u}$ s'expressa com

$$\vec{u}=\alpha_1\vec{u}_1+\alpha_2\vec{u}_2+\cdots+\alpha_n\vec{u}_n$$

llavors els escalars d'aquesta combinació lineal són les coordenades de $\vec{u}$ respecte a la base B. Simbòlicament, escriurem

$$\vec{u}_B=(\alpha_1,\alpha_2,\dots,\alpha_n)$$

Exemple 19.5.

Trobeu les coordenades del polinomi $p(x)=3+x-x^2$ respecte a la base (de $\mathbb{K}_2[x]$) $B=\{1+x,1-x,x-x^2\}$.

En l'exemple 19.4. ja hem vist que el conjunt B és una base de $\mathbb{K}_2[x]$. Ara hem de cercar els escalars α_1, α_2 i α_3 perquè

$$\alpha_1(1+x)+\alpha_2(1-x)+\alpha_3(x-x^2)=3+x-x^2$$

Reordenem el polinomi del costat esquerre,

$$(\alpha_1+\alpha_2)+(\alpha_1-\alpha_2+\alpha_3)x-\alpha_3x^2=3+x-x^2$$

i igualem els coeficients:

$$\begin{aligned}\alpha_1+\alpha_2\qquad &= 3\\ \alpha_1-\alpha_2+\alpha_3 &= 1\\ -\alpha_3 &= -1\end{aligned}$$

Aquest sistema lineal és determinat i la solució és aquesta: $\alpha_1=3/2$, $\alpha_2=3/2$, $\alpha_3=1$. Per tant,

$$p(x)_B=(3/2,3/2,1)$$

19.2.2. Bases canòniques

En les unitats anteriors hem estudiat a bastament les bases dels espais $\mathbb{R}^n$: qualsevol conjunt format per n vectors independents és una base; la base canònica de $\mathbb{R}^n$ és

$$\mathscr{C}=\{(1,0,\dots,0).(0,1,\dots,0)\dots,(0,0,\dots,1)\}$$

La base canònica de l'espai $\mathcal{M}_{m\times n}$ és aquesta:

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 0 & \dots & 0 \\ \dots & & & \\ \dots & & & \\ 0 & 0 & \dots & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & \dots & 0 \\ \dots & & & \\ \dots & & & \\ 0 & 0 & \dots & 0 \end{bmatrix} \dots, \begin{bmatrix} 0 & 0 & \dots & 0 \\ \dots & & & \\ \dots & & & \\ 0 & 0 & \dots & 1 \end{bmatrix} \right\}$$

(notem que aquesta base té mn elements).

Per al conjunt dels polinomis, $\mathbb{K}[x]$, la base canònica és

$$\mathcal{C} = \{1, x, x^2, \dots, x^n, \dots\}$$

(infinits elements).

19.2.3. El teorema de la dimensió. Obtenció de bases

En aquest apartat treballem només amb espais de dimensió finita. Les propietats fonamentals de les bases són aquestes:

Propietats 19.1.

Suposem que E és un espai vectorial de dimensió finita.

(a) Si el conjunt finit S genera tot l'espai E, llavors S conté una base de E.

(b) Si el conjunt S és linealment independent, llavors S està contingut en una base de E.

*(c) **Teorema de la dimensió**: Totes les bases de E tenen el mateix nombre d'elements.*

El teorema de la dimensió assegura que totes les bases tenen el mateix nombre d'elements. Això justifica la definició següent:

Definició 19.5. (Dimensió d'un espai vectorial)

Siga E un espai vectorial qualsevol.

Si E és de dimensió finita, llavors la dimensió de E és el nombre d'elements d'una base qualsevol.

Si E és l'espai nul, la dimensió de E és 0.

En altre cas, la dimensió de E és ∞.

Per exemple, $\dim \mathbb{R}^n = n$, $\dim \mathbb{R}[x] = \infty$, $\dim \mathcal{M}_{m\times n} = mn$.

Exemple 19.6.

Trobeu una base i la dimensió del subespai de $\mathcal{M}_{2\times 2}$

$$F = \left\langle \left\{ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 5 \\ 1 & 0 \end{bmatrix} \right\} \right\rangle$$

El conjunt

$$S = \left\{ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 5 \\ 1 & 0 \end{bmatrix} \right\}$$

genera F. Però, si no és linealment independent, no és una base. És evident que les relacions de dependència entre aquestes matrius són les mateixes que les que hi ha entre els vectors

$$(1, 1, -1, 0), (-2, 1, 2, 0), (0, 3, 0, 0), (-1, 5, 1, 0)$$

i aqueste relacions les podem obtenir esglaonant la matriu

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

La forma esglaonada reduïda és

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Això vol dir que les matrius tercera i quarta són combinacions lineals de les dues primeres:

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

En conseqüència,

$$B = \left\{ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 1 \\ 2 & 0 \end{bmatrix} \right\}$$

és una base de F i $\dim F = 2$.
☞ Aquest exemple confirma la primera propietat: un conjunt generador en un espai de dimensió finita sempre conté una base.

19.3. Exercicis

Exercici 19.1. Trobeu una base de cadascun dels espais següents.

(a) F és el conjunt de les matrius quadrades 2×2 A per a les quals $a_{11} = a_{22}$ i $a_{12} = 2a_{21}$.

(b) F és el conjunt de les matrius simètriques 2×2.

(c) F és el conjunt de les matrius antisimètriques 2×2.

(d) F és el conjunt dels polinomis de grau parell.

(e) F és el conjunt dels polinomis de la forma $p(x) = a + 2ax$ on a és una constant qualsevol.

Exercici 19.2. Trobeu una base de l'espai de polinomis $\left\langle 1 + x, 1 - x^2, x + x^2, 2 + x - x^2 \right\rangle$.

Exercici 19.3. Trobeu les coordenades respecte a la base

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

de la matriu $\mathsf{A} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$.

Exercici 19.4. Proveu que el conjunt

$$B = \{1, x - 1, (x - 1)^2, (x - 1)^3\}$$

és una base de $\mathbb{R}_3[x]$ i calculeu les coordenades respecte aquesta base del polinomi $p(x) = x^3 - x + 1$.

Exercici 19.5. Proveu que el conjunt

$$B = \{1, x - 1, (x - 1)^2, (x - 1)^3, \dots, (x - 1)^n, \dots\}$$

és una base de $\mathbb{R}[x]$ i calculeu les coordenades respecte aquesta base d'un polinomi arbitrari $p(x)$.

Exercici 19.6. (a) Proveu que el conjunt S de les successions que són solució de la recurrència

$$a_{n+2} - 4a_n = 0$$

és un subespai vectorial de l'espai de totes les successions.

(b) Proveu que aquest subespai és de dimensió 2.

(c) Comproveu que les successions $a_n = 2^n$ i $b_n = (-2)^n$ són solucions de la recurrència linealment independents i deduïu que qualsevol solució de la recurrència n'és combinació lineal.

(d) Resoleu la recurrència pels mètodes típics de l'anàlisi matemàtica i compareu els resultats.

20. Aplicacions lineals

Per acabar d'entendre la importància de les matrius en els problemes de caràcter lineal ens falta encara fer un últim pas: interpretar la multiplicació matriu-vector com un *operador* que *transforma* l'espai. Si A és una matriu $m \times n$ i $\vec{x}$ un vector $n \times 1$, aleshores $\mathsf{A}\vec{x}$ és un vector $m \times 1$, de manera que podem entendre que la multiplicació *transforma* els vectors de $\mathbb{R}^n$ en vectors de $\mathbb{R}^m$.
Aquest és l'exemple més simple d'aplicació lineal.

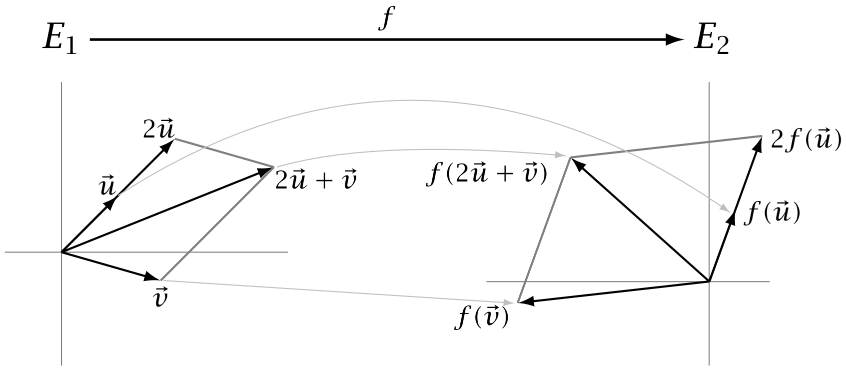
20.1. Aplicacions lineals

Un aplicació f entre dos espais vectorials és lineal si conserva les combinacions lineals.

Definició 20.1.

Donats dos espais vectorials sobre $\mathbb{K}$, E_1 i E_2, una aplicació $f: E_1 \longrightarrow E_2$, és lineal si

$$f(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) = \alpha_1 f(\vec{x}_1) + \alpha_2 f(\vec{x}_2), \qquad \forall \vec{x}_1, \vec{x}_2 \in E, \quad \forall \alpha_1, \alpha_2 \in \mathbb{K}$$



Exemple 20.1.

Si A és una matriu $m \times n$, llavors l'aplicació

$$\begin{aligned} f\colon \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \vec{x} &\rightsquigarrow f(\vec{x}) = \mathsf{A}\vec{x} \end{aligned}$$

és lineal.

Aquesta aplicació és lineal perquè

$$f(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) = \mathsf{A}(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2) = \alpha_1\mathsf{A}\vec{x}_1 + \alpha_2\mathsf{A}\vec{x}_2 = \alpha_1 f(\vec{x}_1) + \alpha_2 f(\vec{x}_2) \quad \square$$

Exemple 20.2.

La transposició de matrius,

$$\begin{aligned} f\colon \mathcal{M}_{m\times n} &\longrightarrow \mathcal{M}_{n\times m} \\ \mathsf{A} &\rightsquigarrow f(\mathsf{A}) = \mathsf{A}^t \end{aligned}$$

és lineal.

La prova és immediata:

$$f(\alpha_1\mathsf{A}_1 + \alpha_2\mathsf{A}_2) = (\alpha_1\mathsf{A}_1 + \alpha_2\mathsf{A}_2)^t = \alpha_1\mathsf{A}_1^t + \alpha_2\mathsf{A}_2^t = \alpha_1 f(\mathsf{A}_1) + \alpha_2 f(\mathsf{A}_2) \quad \square$$

La derivada també és lineal:

EXEMPLE 20.3.

L'aplicació

$$\begin{aligned} D\colon \mathbb{R}[x] &\longrightarrow \mathbb{R}[x] \\ p(x) &\rightsquigarrow D(p(x)) = p'(x) \end{aligned}$$

és lineal.

$$D(\alpha_1 p_1(x) + \alpha_2 p_2(x)) = (\alpha_1 p_1(x) + \alpha_2 p_2(x))' = \alpha_1 p_1'(x) + \alpha_2 p_2'(x) = \alpha_1 D(p(x)) + \alpha_2 D(p(x)) \quad \square$$

I la integral:

EXEMPLE 20.4.

L'aplicació

$$\begin{aligned} F\colon C_0([a,b]) &\longrightarrow \mathbb{R} \\ f &\rightsquigarrow F(f) = \int_a^b f(t)dt \end{aligned}$$

és lineal

$$\int_a^b \Big(\alpha_1 f_1(t) + \alpha_2 f_2(t)\Big)dt = \alpha_1 \int_a^b f_1(t)dt + \alpha_2 \int_a^b f_2(t)dt \quad \square$$

20.1.1. NUCLI I IMATGE

DEFINICIÓ 20.2.

Si $f : E_1 \longrightarrow E_2$ és una aplicació lineal, el nucli de f ($\operatorname{Nuc} f$ o $\ker f$) és el conjunt de totes les antiimatges del vector $\vec{0}$.
La imatge de f ($\operatorname{Im} f$ o $f(E_1)$) és el conjunt de les imatges de tots els vectors de E_1.

EXEMPLE 20.5.

L'aplicació

$$\begin{aligned} f\colon \mathbb{R}^3 &\longrightarrow \mathbb{R}[x] \\ (u_1,u_2,u_3) &\rightsquigarrow f((u_1,u_2,u_3)) = u_2 + (u_1+u_3)x \end{aligned}$$

és lineal. Trobeu-ne el nucli i la imatge.

La imatge és $\mathbb{R}_1[x]$, perquè és clar que totes les imatges són polinomis de grau 0 o 1 i que qualsevol polinomi de la forma $a + bx$ té l'antiimatge $(b, a, 0)$.

Un vector (u_1, u_2, u_3) és al nucli si i només si

$$f(u_1,u_2,u_3) = 0 \Leftrightarrow u_2 + (u_1+u_3)x = 0 \Leftrightarrow \begin{bmatrix} u_2 = 0 \\ u_1 + u_3 = 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} u_2 = 0 \\ u_1 = -u_3 \end{bmatrix}$$

així que el nucli és $\operatorname{Nuc} f = \langle(-1,0,1)\rangle$.

EXEMPLE 20.6.

L'aplicació

$$\begin{aligned} f\colon \mathcal{M}_{2\times 2} &\longrightarrow \mathcal{M}_{2\times 2} \\ \mathsf{A} &\rightsquigarrow f(\mathsf{A}) = \mathsf{A} - \mathsf{A}^t \end{aligned}$$

és lineal. Proveu que el nucli i la imatge són, respectivament, el conjunt de les matrius 2×2 simètriques i el de les antisimètriques.

Una matriu A és al nucli si $\mathsf{A} - \mathsf{A}^t = \mathsf{O}$, és a dir, si $\mathsf{A} = \mathsf{A}^t$. Per tant, el nucli de f és el conjunt de les matrius simètriques 2×2.

D'altra banda, la matriu B és a la imatge si hi ha una matriu A per a la qual $\mathsf{A} - \mathsf{A}^t = \mathsf{B}$. En eixe cas, $\mathsf{B}^t = (\mathsf{A} - \mathsf{A}^t)^t = \mathsf{A}^t - \mathsf{A} = -\mathsf{B}$. Això vol dir que totes les imatges són matrius antisimètriques. Per altra banda, si B és una matriu antisimètrica, llavors $f(\frac{1}{2}\mathsf{B}) = \frac{1}{2}(\mathsf{B} - \mathsf{B}^t) = \frac{1}{2}(\mathsf{B} + \mathsf{B}) = \mathsf{B}$, així que totes matrius antisimètriques són a la imatge. $\square$

Propietats 20.1.

Siga f una aplicació lineal entre els espais E_1 i E_2.

1. *Dos vectors $\vec{u}$ i $\vec{v}$ tenen la mateixa imatge si i només si $\vec{u} - \vec{v} \in \operatorname{Nuc} f$.*
2. *El nucli de f és un subespai de E_1.*
3. *L'aplicació f és injectiva si i només si $\operatorname{Nuc} f = \{\vec{0}\}$.*
4. *La imatge $\operatorname{Im} f$ és un subespai de E_2.*
5. *Fórmula de les dimensions. Si E_1 és de dimensió finita, llavors*

$$\dim \operatorname{Nuc} f + \dim \operatorname{Im} f = \dim E_1 \quad \square$$

20.2. Transformacions lineals de $\mathbb{R}^n$ a $\mathbb{R}^m$

Com hem vist adés, si A és una matriu $m \times n$, llavors $f(\vec{x}) = \mathsf{A}\vec{x}$ és una aplicació lineal que va de $\mathbb{R}^n$ a $\mathbb{R}^m$. Per exemple, si $\mathsf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix}$ i $\vec{x} = (x_1, x_2, x_3)$ és un vector tridimensional qualsevol, el producte $\mathsf{A}\vec{x}$ és el vector bidimensional

$$\mathsf{A}\vec{x} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -2 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -2x_2 + x_3 \end{bmatrix}$$

☞ Notem que les tres columnes de la matriu A són les imatges dels vectors de la base canònica $\mathcal{E}_3 = \{(1,0,0), (0,1,0), (0,0,1)\}$:

$$f(1,0,0) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$f(0,1,0) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

$$f(0,0,1) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

així que

$$\boxed{\mathsf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \begin{bmatrix} f(\vec{e}_1) & f(\vec{e}_2) & f(\vec{e}_3) \end{bmatrix}}$$

Es pot provar fàcilment que qualsevol transformació lineal entre els espais $\mathbb{R}^n$ i $\mathbb{R}^m$ és d'aquest tipus, és a dir que

$$\boxed{\text{Si } f : \mathbb{R}^n \longrightarrow \mathbb{R}^m \text{ és lineal llavors } f(\vec{x}) = \mathsf{A}\vec{x} \text{ on } \mathsf{A} = \begin{bmatrix} f(\vec{e}_1) & f(\vec{e}_2) & \cdots & f(\vec{e}_n) \end{bmatrix}}$$

20.3. Interpretació geomètrica

Una manera interessant d'entendre les transformacions lineals és interpretar-les com a transformacions geomètriques.

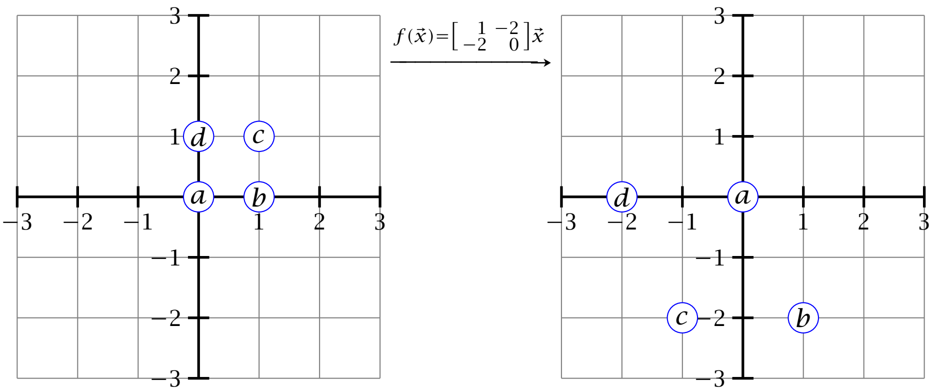
Considerem, per exemple, la transformació lineal

$$f(\vec{x}) = \begin{bmatrix} 1 & -2 \\ -2 & 0 \end{bmatrix} \vec{x}$$

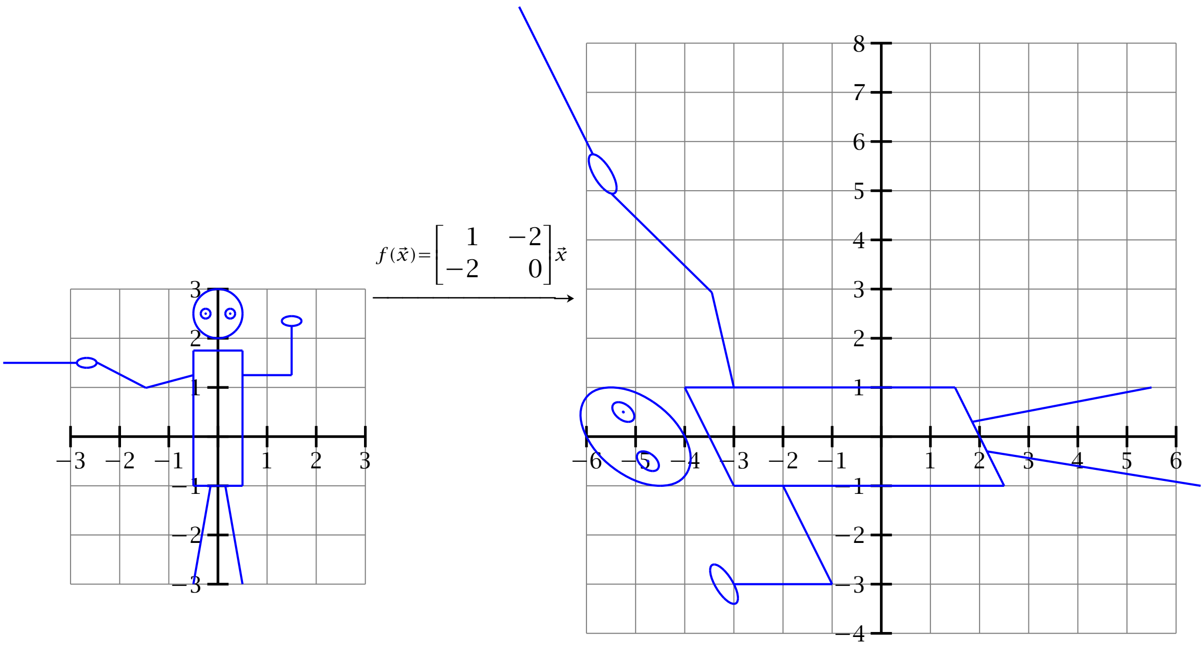
Podem calcular les imatges d'uns quants vectors,

$$f(0,0) = (0,0) \quad f(1,0) = (1,-2) \quad f(1,1) = (-1,-2) \quad f(0,1) = (-2,0)$$

i representar-les gràficament.



Però l'efecte geomètric es farà més evident si, en comptes de representar uns quants punts aïllats, mirem com es transforma tot un gràfic.



Observem que la figura es deforma, es fa més gran i gira. Fins i tot, una observació més acurada ens fa notar que aquesta imatge s'ha *reorientat*. Tot i això, la imatge *s'assembla molt* al dibuix original, perquè els segments de recta es transformen en segments de recta i les el·lipses en el·lipses; això no seria així si la transformació no fos lineal.

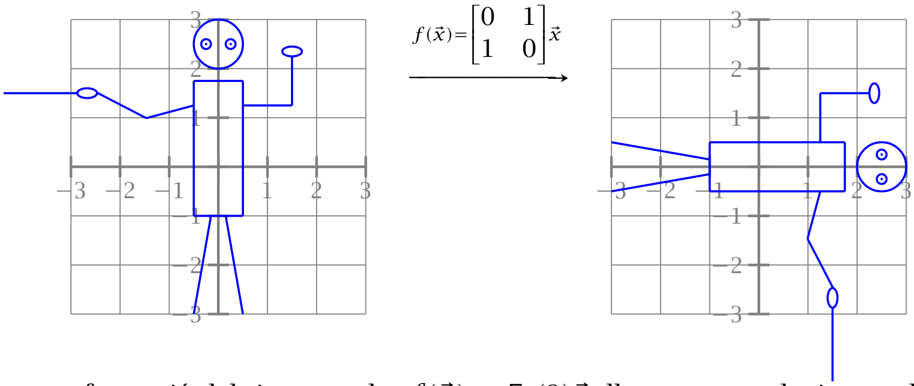
20.3.1. Transformacions lineals especials

En aquest apartat estudiarem l'efecte geomètric de les transformacions lineals associades a alguns tipus especials de matrius.

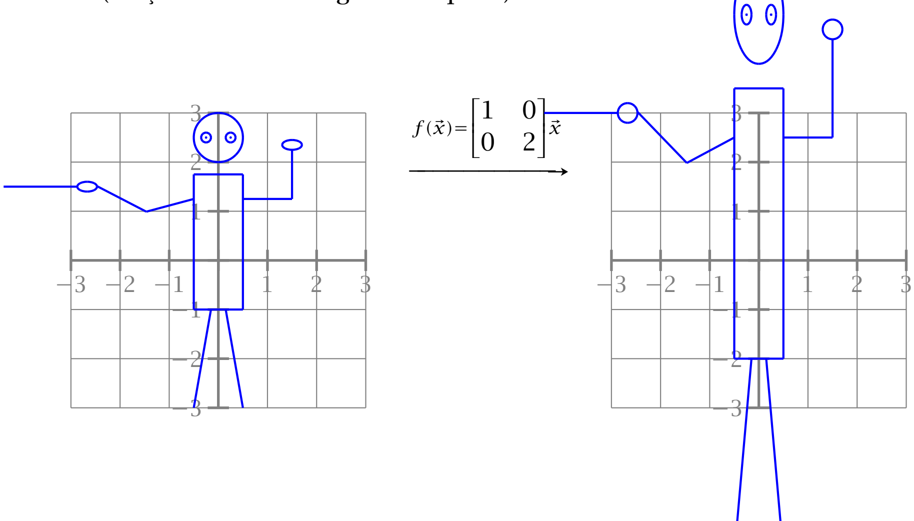
Matrius elementals En cada un dels gràfics que segueixen es visualitza l'efecte de la transformació lineal associada a una matriu elemental.

(a) Començarem amb la transformació lineal $f(\vec{x}) = \mathsf{E}_{1,2}\vec{x}$, és a dir, amb una matriu elemental del tipus permutació. Aquesta aplicació f transforma l'eix horitzontal en el vertical, i el vertical en l'horitzontal.

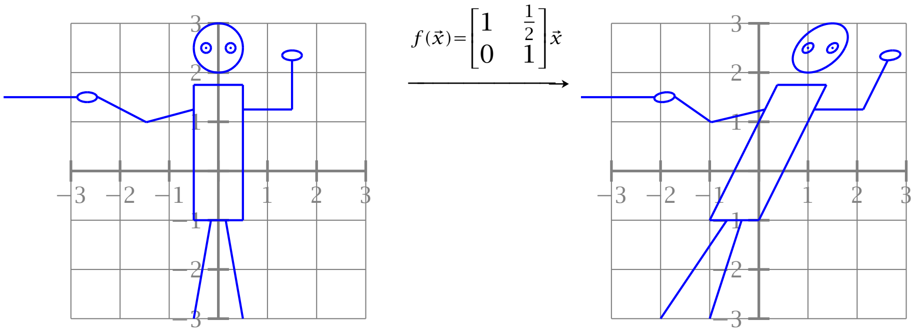
La imatge és el resultat d'aplicar la simetria respecte a la recta $x_1 = x_2$.



(b) Si fem la transformació del tipus escalat $f(\vec{x}) = \mathsf{E}_2(2)\vec{x}$, llavors es produeix una dilatació en la direcció vertical (l'alçada de les imatges es duplica).

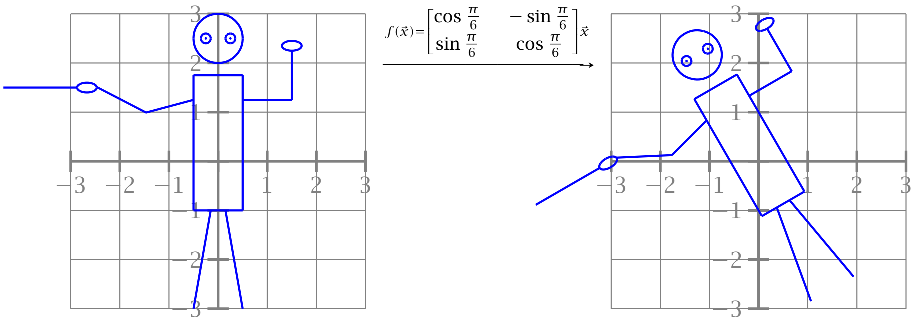


(c) Finalment, la transformació $f(\vec{x}) = \mathsf{E}_{1,2}(1/2)\vec{x}$ transforma el vector (x_1, x_2) en $(x_1 + (1/2)x_2, x_2)$, així que la figura *es doblega* cap a la dreta.

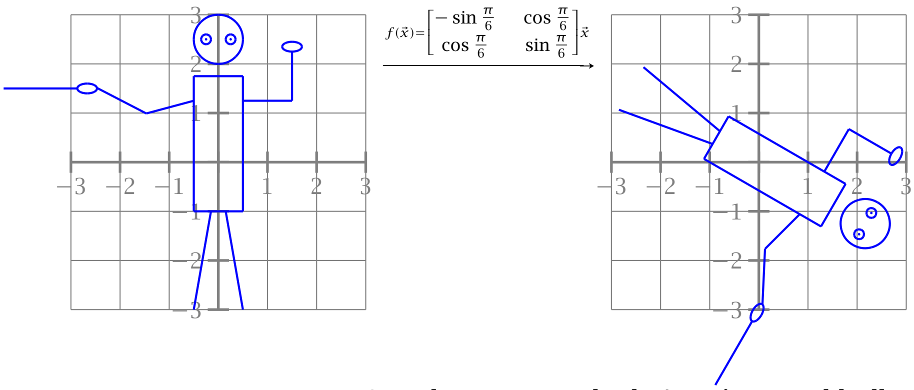


MATRIUS ORTOGONALS Quan la matriu canònica de la transformació lineal f és ortogonal es diu que f és una *isometria*, perquè la transformació lineal f conserva els angles i les distàncies, de manera que *no deformen les figures*. En el cas de les matrius 2×2, només hi ha dos tipus d'isometries: les rotacions i les simetries, així que en veurem un exemple de cada tipus.

(a) La transformació $f(\vec{x}) = \begin{bmatrix}\cos\frac{\pi}{6} & -\sin\frac{\pi}{6}\\ \sin\frac{\pi}{6} & \cos\frac{\pi}{6}\end{bmatrix}\vec{x}$ *gira* la imatge 30^o en el sentit contrari al de les agulles del rellotge.

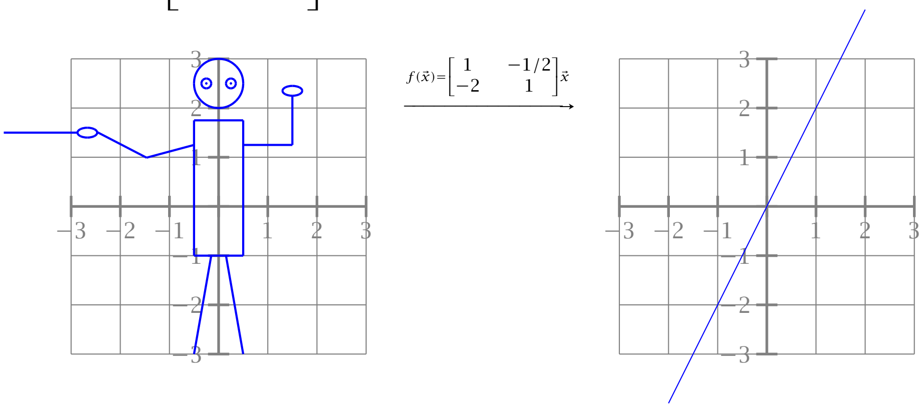


(b) La transformació $f(\vec{x}) = \begin{bmatrix} -\sin\frac{\pi}{6} & \cos\frac{\pi}{6} \\ \cos\frac{\pi}{6} & \sin\frac{\pi}{6} \end{bmatrix}\vec{x}$ produeix una simetria respecte a la recta que passa per l'origen i té un pendent de 15°.



Matrius quadrades no invertibles Quan la matriu quadrada A no és invertible, llavors l'aplicació f no és injectiva (perquè el sistema $A\vec{x} = \vec{0}$ és indeterminat, de manera que el vector zero té infinites antiimatges).

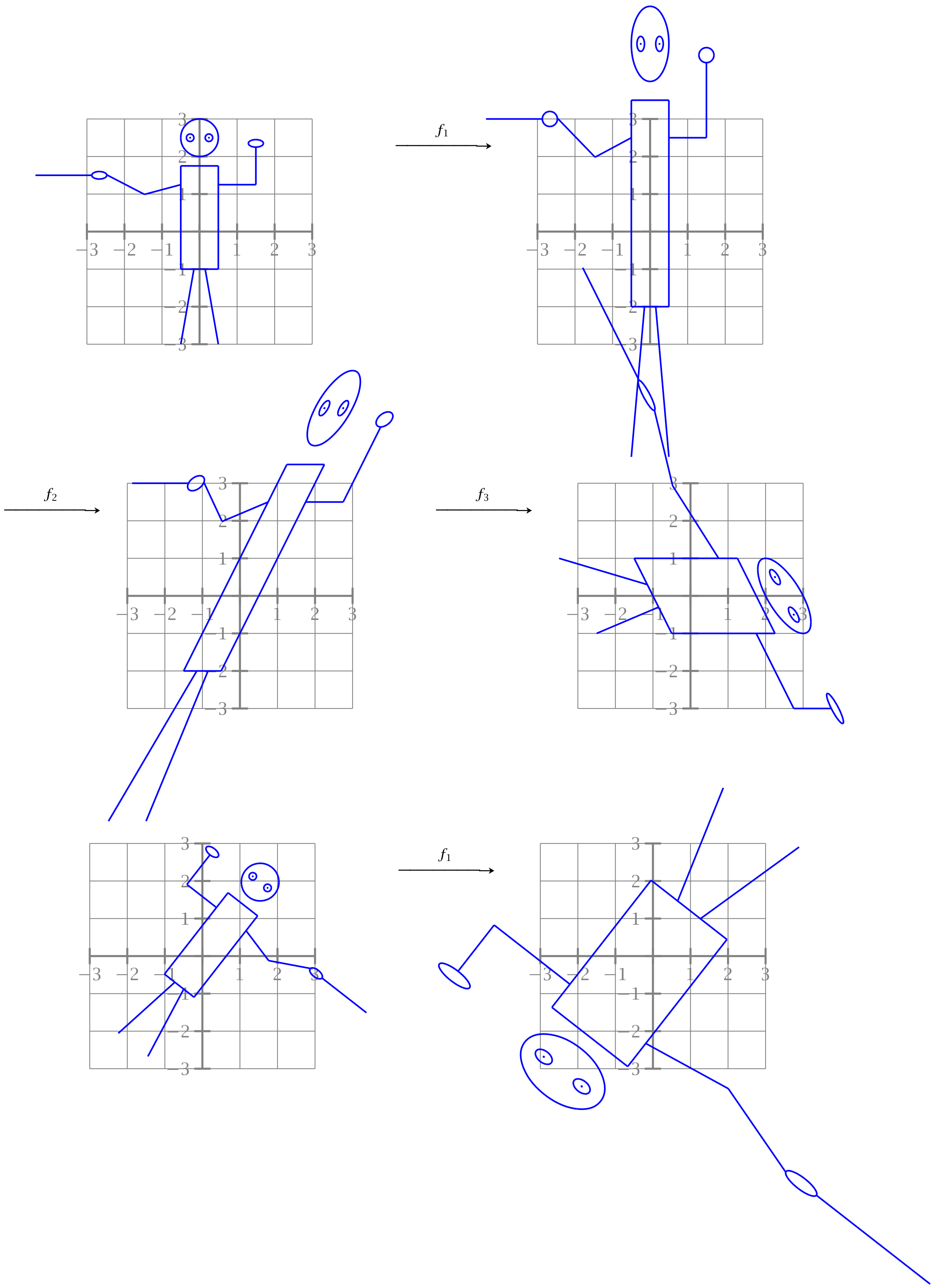
Per exemple, si $A = \begin{bmatrix} 1 & -1/2 \\ -2 & 1 \end{bmatrix}$, llavors totes les imatges es troben sobre una recta.



Producte de transformacions elementals Com que qualsevol matriu invertible és un producte de matrius elementals, la transformació lineal associada a una matriu regular és la composició de diverses transformacions elementals. Per exemple, la matriu $A = \begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix}$ es pot escriure com

$$A = E_{2,1}(-2)E_{1,2}(1/2)E_2(2) = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Si f és la transformació lineal associada a A i f_3, f_2 i f_1 són, respectivament les transformacions corresponents a $E_{2,1}(-2)$, $E_{1,2}(1/2)$ i $E_2(2)$, llavors $f = f_3 \circ f_2 \circ f_1$, de manera que l'efecte geomètric de f es pot reconstruir aplicant successivament tres transformacions elementals.



f_1
f_2
f_3
f_1

20.3.2. Propietats algebràiques i geomètriques

El nucli i la imatge de l'aplicació lineal $f(\vec{x}) = \mathsf{A}\vec{x}$ són dos dels subespais deduïts de la matriu A. Com que $f(\vec{x}) = \mathsf{A}\vec{x}$, $f(\vec{x}) = \vec{0}$ és equivalent a $\mathsf{A}\vec{x} = \vec{0}$ i resulta que

☞ el nucli de f coincideix amb l'espai nul de la matriu A.

Exemple 20.7.

Calculeu el nucli de l'aplicació lineal $f(\vec{x}) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \vec{x}$

$$\operatorname{Nuc} f = \operatorname{Nul} \mathsf{A} = \operatorname{Nul} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \operatorname{Nul} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1/2 \end{bmatrix} = \left\langle \left(-1, \frac{1}{2}, 1\right) \right\rangle \quad \square$$

D'altra banda, per a calcular la imatge de la transformació f haurem de trobar els vectors $\vec{b}$ per als quals hi ha algun $\vec{x}$ de manera que $f(\vec{x}) = \vec{b}$, és a dir, que $\mathsf{A}\vec{x} = \vec{b}$. En altres paraules, el vector $\vec{b}$ és a la imatge si el sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és compatible. Per tant,

☞ la imatge $\operatorname{Im} f$ coincideix amb l'espai columna $\operatorname{Col} \mathsf{A}$.

Exemple 20.8.

Calculeu el conjunt imatge de l'aplicació lineal $f(\vec{x}) = \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} \vec{x}$

$$\operatorname{Im} f = \operatorname{Col} \mathsf{A} = \operatorname{Col} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \mathbb{R}^2 \quad \square$$

El que sabem a prop de les dimensions dels espais nul i columna, $\dim \operatorname{Nul} \mathsf{A} + \dim \operatorname{Col} \mathsf{A} = n$, ens proporciona una nova justificació, per a les transformacions lineals entre $\mathbb{R}^n$ i $\mathbb{R}^m$, de la fórmula de les dimensions:

$$\dim \operatorname{Nuc} f + \dim \operatorname{Im} f = n \quad \square$$

Ara, si el nucli de f és l'espai nul de A i la imatge de f n'és l'espai columna, *quin paper hi juguen l'espai fila i l'espai nul esquerre?* Recordem que $\mathbb{R}^n$ és la suma ortogonal de l'espai fila i l'espai nul,

$$\mathbb{R}^n = \operatorname{Fil} \mathsf{A} \oplus \operatorname{Nul} \mathsf{A}$$

i $\mathbb{R}^m$ és la suma ortogonal de l'espai columna i l'espai nul esquerre,

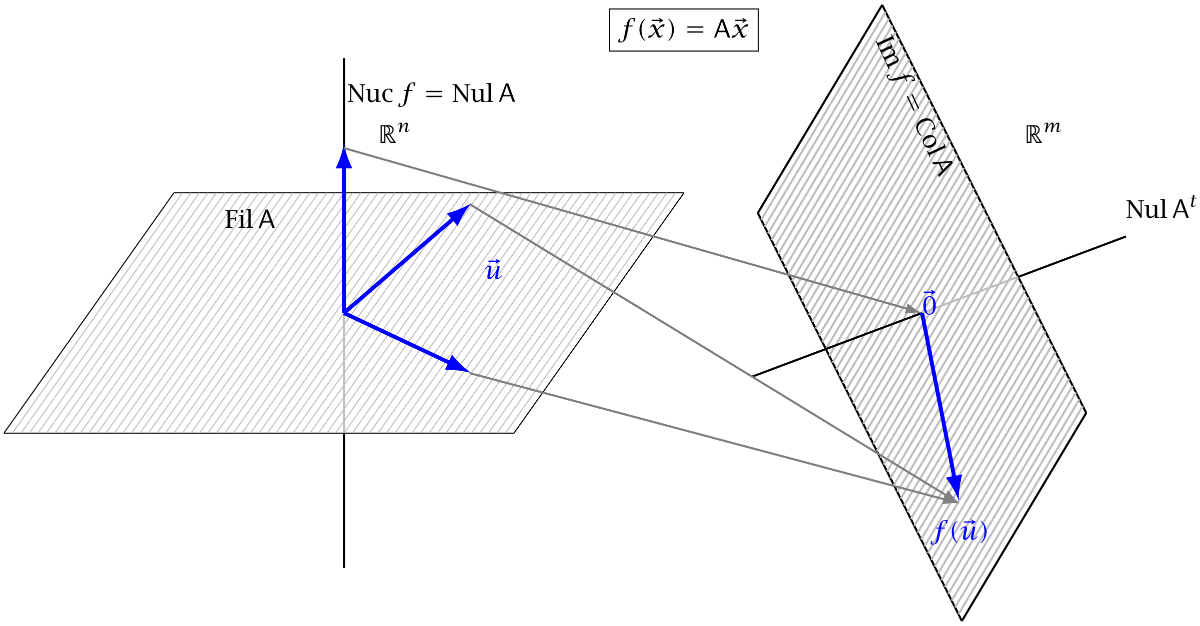
$$\mathbb{R}^m = \operatorname{Col} \mathsf{A} \oplus \operatorname{Nul} \mathsf{A}^t$$

Així doncs, si $\vec{u}$ és un vector de $\mathbb{R}^n$, llavors el podem descompondre com a suma de les projeccions ortogonals sobre l'espai fila i l'espai nul: $\vec{u} = p_{\operatorname{Fil} \mathsf{A}}(\vec{u}) + p_{\operatorname{Nul} \mathsf{A}}(\vec{u})$. Aleshores,

$$f(\vec{u}) = f(p_{\operatorname{Fil} \mathsf{A}}(\vec{u})) + \underbrace{f(p_{\operatorname{Nul} \mathsf{A}}(\vec{u}))}_{0} = f(p_{\operatorname{Fil} \mathsf{A}}(\vec{u}))$$

Això vol dir que la imatge de qualsevol vector coincideix amb la de la projecció d'aquest vector sobre l'espai fila de A. En conseqüència, *la imatge de l'espai fila és l'espai columna.*

- L'espai fila es transforma en l'espai columna (és a dir, en el conjunt de les imatges).
- L'espai nul va al vector zero.
- L'espai nul esquerre és l'ortogonal del conjunt de les imatges.



Els quatre subespais de la matriu A i la transformació lineal $f(\vec{x}) = \mathsf{A}\vec{x}$

Propietat 20.2.

L'aplicació f, restrigida als espais fila i columna,

$$\begin{aligned} f\colon \text{Fil}\,\mathsf{A} &\longrightarrow \text{Col}\,\mathsf{A} \\ \vec{x} &\rightsquigarrow f(\vec{x}) = \mathsf{A}\vec{x} \end{aligned}$$

és bijectiva.

Demostració: Si $\vec{x}$ i $\vec{y}$ són dos vectors de l'espai fila i $f(\vec{x}) = f(\vec{y})$, llavors $f(\vec{x} - \vec{y}) = \vec{0}$, de manera que $\vec{x} - \vec{y}$ es troba en Fil A i en Nul A; com que la intersecció d'aquests dos subespais és nul·la, $\vec{x} = \vec{y}$. Això demostra que l'aplicació és injectiva. Com que ja sabíem que la imatge de l'espai fila és l'espai columna, aquesta aplicació també és suprajectiva. $\square$

Això ens proporciona una nova justificació del fet que els espais fila i columna tenen la mateixa dimensió (o del fet que els rangs per files i columnes són iguals).

20.4. Matrius associades a una aplicació lineal

El fet que les aplicacions lineals entre espais $\mathbb{R}^n$ tenen la forma $f(x) = \mathsf{A}\vec{x}$ es pot generalitzar fàcilment a qualsevol aplicació lineal entre espais vectorials de dimensió finita, la qual cosa ens permetrà convertir el càlcul de les imatges a un producte matriu-vector.

Si definim la matriu de f respecte a un parell de bases d'aquesta manera,

Suposem que $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ i $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ són bases dels espais de dimensió finita E_1 i E_2 i que $f\colon E_1 \longrightarrow E_2$ és una aplicació lineal. La matriu de f respecte a les bases B_1 i B_2 és aquesta:

$$\mathsf{M}(f, B_1, B_2) = \begin{bmatrix} f(\vec{u}_1)_{B_2} & f(\vec{u}_2)_{B_2} & \dots & f(\vec{u}_n)_{B_2} \end{bmatrix}$$

llavors, fent servir aquesta matriu, les imatges a traves de les aplicacions lineals entre espais de dimensió finita es poden calcular com

$$f(\vec{u})_{B_2} = \mathsf{M}(f, B_1, B_2)\vec{u}_{B_1}$$

EXEMPLE 20.9.

L'aplicació

$$\begin{aligned} f\colon \mathbb{R}_2[x] &\longrightarrow \mathbb{R}_3[x] \\ a_0 + a_1x + a_2x^2 &\rightsquigarrow f(a_0 + a_1x + a_2x^2) = a_2x + (a_1 - a_0)x^2 + 2(a_0 - a_1)x^3 \end{aligned}$$

és lineal. Trobeu la matriu de f respecte a les bases $B_1 = \{1, x, x^2\}$ i $B_2 = \{1, x, x^2, x^3\}$ i feu-la servir per calcular la imatge del polinom $p(x) = 1 + x + x^2$.
Calculeu bases del nucli i la imatge de f.

Calculem les imatges de cada vector de B_1 i les escrivim com a combinació lineal dels vectors de la base B_2:

$$\begin{aligned} f(1) &= -x^2 + 2x_3 & \qquad f(1)_{B_2} &= (0, 0, -1, 2) \\ f(x) &= x^2 - 2x^3 & f(x)_{B_2} &= (0, 0, 1, -2) \\ f(x^2) &= x & f(x^2)_{B_2} &= (0, 1, 0, 0) \end{aligned}$$

La matriu és aquesta:

$$\mathsf{M}(f, B_1, B_2) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{bmatrix}$$

El vector de coordenades de $p(x) = 1 + x + x^2$ respecte a B_1 és $p(x)_{B_1} = (1, 1, 1)$. Per tant,

$$f(p(x))_{B_2} = \mathsf{M}(f, B_1, B_2)p(x)_{B_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Per tant,

$$f(p(x)) = x$$

La matriu $\mathsf{M}(f, B_1, B_2)$ té rang 2 (la segona columna és proporcional a la primera). Per tant, els polinomis corresponents a les columnes primera i tercera són una base de la imatge:

$$B_{\mathrm{Im}\, f} = \{-x^2 + 2x^3, x\}$$

El nucli el podem calcular resolent el sistema lineal $\mathsf{M}(f, B_1, B_2)\vec{x} = \vec{0}$. Com que la forma esglaonada reduïda de la matriu és

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

la solució és $a_0 = a_1$, $a_2 = 0$. Per tant, el nucli és el conjunt dels polinomis de la forma $a_0 + a_0x$, i com a base podem elegir

$$B_{\mathrm{Nuc}\, f} = \{1 + x\} \quad \square$$

20.4.1. Matrius associades en diverses bases

Suposem que $\mathsf{A} = \mathsf{M}(f, B_1, B_2)$ és la matriu associada a l'aplicació f entre els espais E_1 i E_2 i respecte a les bases B_1 i B_2. Si canviem les bases B_1 i B_2 per unes altres, B'_1 i B'_2 llavors, obtindrem una altra matriu associada a l'aplicació f, $A' = \mathsf{M}(f, B'_1, B'_2)$.

Per tal de trobar la relació que hi ha entre les matrius A i A', notem que, si $\vec{u}$ és un vector qualsevol de l'espai E_1,

$$f(\vec{u})_{B_2} = \mathsf{A}\vec{u}_{B_1} \tag{20.1}$$

$$f(\vec{u})_{B'_2} = \mathsf{A}'\vec{u}_{B'_1} \tag{20.2}$$

Ara bé, tenint en comte la fórmula del canvi de base, si $\mathsf{P} = \mathsf{M}_{B'_1 B_1}$ (la matriu del canvi de base de B'_1 a B_1) i $Q = \mathsf{M}_{B'_2 B_2}$ llavors

$$\begin{aligned} f(\vec{u})_{B_2} &= \mathsf{Q} f(\vec{u})_{B'_2} \\ \vec{u}_{B_1} &= \mathsf{P}\vec{u}_{B'_1} \end{aligned}$$

podem substituir les coordenades dels vectors en (20.1),

$$\mathsf{Q} f(\vec{u})_{B'_2} = \mathsf{AP}\vec{u}_{B'_1}$$

$$f(\vec{u})_{B'_2} = \mathsf{Q}^{-1}\mathsf{AP}\vec{u}_{B'_1} \tag{20.3}$$

El següent esquema justifica l'expressió (20.3): Calcular la imatge $f(\vec{u})$ fent servir les bases B_1 i B_2 és equivalent al procés següent: en primer lloc canviem de B_1 a B'_1, tot seguit calculem la imatge fent servir la matriu de f respecte a les bases B'_1 i B'_2 i, finalment, canviem de B'_2 a B_2.

$$\begin{tikzcd} \vec{u}_{B'_1} \arrow[r, "M_{B'_1 B_1}"] \arrow[rrr, bend right, "M(f, B'_1, B'_2)"'] & \vec{u}_{B_1} \arrow[r, "M(f, B_1, B_2)"] & f(\vec{u}_1)_{B_2} \arrow[r, "M_{B_2 B'_2}"] & f(\vec{u}_1)_{B'_2} \end{tikzcd}$$

Com que aquestes expressions són vàlides per a qualsevol vector $\vec{u}$, de (20.3) i (20.2) deduïm que

$$\mathsf{A}' = \mathsf{Q}^{-1}\mathsf{AP}$$

o bé,

$$\mathsf{A} = \mathsf{QA}'\mathsf{P}^{-1}$$

En resum,

Fórmula del canvi de bases per a aplicacions lineals
Si A és la matriu de f respecte a les bases B_1 i B_2, A' és la matriu de f respecte a les bases B'_1 i B'_2, P és la matriu de canvi de la base B'_1 a B_1, Q és la matriu de canvi de la base B'_2 a B_2 llavors,

$$\boxed{\mathsf{A} = \mathsf{QA}'\mathsf{P}^{-1}} \tag{20.4}$$

És interessant observar que aquesta fórmula significa que la matriu A es pot obtenir fent sobre A' operacions elementals per files i per columnes.

Exemple 20.10.

Trobeu la matriu de l'aplicació lineal

$$\begin{aligned} f\colon \mathbb{R}_2[x] &\longrightarrow \mathbb{R}_3[x] \\ a_0 + a_1 x + a_2 x^2 &\rightsquigarrow f(a_0 + a_1 x + a_2 x^2) = a_2 x + (a_1 - a_0)x^2 + 2(a_0 - a_1)x^3 \end{aligned}$$

respecte a les bases $B'_1 = \{1, x^2, 1 + x\}$ i $B'_2 = \{x^2 - 2x^3, x, x^2, 1\}$ i compareu els resultats amb els de l'exemple 20.9. per comprovar que s'hi verifica la relació (20.4).

Si expressem les imatges dels vectors de la base B'_1 respecte a la base B'_2,

$$\begin{aligned} f(1) &= -x^2 + 2x_3 & \qquad f(1)_{B'_2} &= (-1, 0, 0, 0) \\ f(x^2) &= x & f(x)_{B'_2} &= (0, 1, 0, 0) \\ f(1 + x) &= 0 & f(x^2)_{B'_2} &= (0, 0, 0, 0) \end{aligned}$$

trobem

$$\mathsf{A}' = \mathsf{M}(f, B_1', B_2') = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Les bases de l'exemple 20.9. són les canòniques $B_1 = \{1, x, x^2\}$ i $B_2 = \{1, x, x^2, x^3\}$, així que les matrius de canvi de base $M_{B_1'B_1}$ i $M_{B_2'B_2}$ són

$$\mathsf{P} = \mathsf{M}_{B_1'B_1} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathsf{Q} = \mathsf{M}_{B_2'B_2} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 \end{bmatrix}$$

En aquell exemple vam trobar que la matriu de f respecte a les bases B_1 i B_2 és

$$\mathsf{A} = \mathsf{M}(f, B_1, B_2) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{bmatrix}$$

Llavors,

$$\begin{aligned} \mathsf{QA'P}^{-1} &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & -2 & 0 \end{bmatrix} \\ &= \mathsf{A} \quad \square \end{aligned}$$

20.5. Exercicis

Exercici 20.1. Estudieu si les aplicacions següents són lineals

(a) $f\colon \mathbb{R}[x] \longrightarrow \mathbb{R}$, $p(x) \rightsquigarrow f(p(x)) = p(0)$

(b) $f\colon \mathcal{M}_{2\times 2} \longrightarrow \mathbb{R}$, $\mathsf{A} \rightsquigarrow f(\mathsf{A}) = a_{11} + a_{22}$

(c) $f\colon \mathbb{R} \longrightarrow \mathbb{R}_2[x]$, $a \rightsquigarrow f(a) = a + 2ax - ax^2$

(d) $f\colon \mathbb{R} \longrightarrow \mathbb{R}_2[x]$, $a \rightsquigarrow f(a) = a + 2ax - x^2$

Exercici 20.2. Trobeu bases del nucli i la imatge de l'aplicació lineal $f\colon \mathcal{M}_{2\times 2} \longrightarrow \mathcal{M}_{2\times 3}$ definida per

$$f\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} a_{11} + a_{22} & a_{12} - a_{21} & 0 \\ 0 & a_{11} - a_{12} + a_{21} + a_{22} & a_{22} \end{bmatrix}$$

Exercici 20.3. Proveu que l'aplicació de $\mathbb{R}^3$ en $\mathbb{R}^2$

$$f(x_1, x_2, x_3) = (x_1 + x_2 - 2x_3, x_1 + x_2 + x_3)$$

és una transformació lineal i calculeu una base del nucli i una de la imatge de f.

Exercici 20.4. De la transformació lineal f sabem que la matriu $\mathsf{M}(f, B, B)$, on B és la base $B = \{(1,1), (1,-1)\}$ és

$$\mathsf{M}(f, B, B) = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$$

Quina és la matriu canònica de f?

Exercici 20.5. Trobeu la matriu $\mathsf{M}(f, B, B)$ de l'aplicació lineal

$$\begin{aligned} f\colon \mathbb{R}_2[x] &\longrightarrow \mathbb{R}_2[x] \\ p(x) &\rightsquigarrow f(p(x)) = p'(x) - p(x) \end{aligned}$$

respecte a la base $B = \{1, x, x^2\}$ i proveu que f és bijectiva.

Un *endomorfisme* és una aplicació lineal de $\mathbb{R}^n$ en el mateix espai $\mathbb{R}^n$. Normalment, si treballem amb un endomorfisme ens interessarà treballar amb la mateixa base per als originals i les imatges.

Exercici 20.6. Trobeu la matriu $\mathsf{M}(f, B, B)$ de l'endomorfisme de $\mathbb{R}^2$ $f(\vec{u}) = (3u_1 + u_2, u_1 + 3u_2)$ respecte a la base $B = \{(1, -1), (1, 1)\}$.

Exercici 20.7. (a) Trobeu la matriu $\mathsf{M}(f, B_1, B_2)$ de l'aplicació lineal

$$\begin{aligned} f\colon \mathbb{R}_3[x] &\longrightarrow \mathbb{R}_2[x] \\ a_0 + a_1x + a_2x^2 + a_3x^3 &\rightsquigarrow f(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0 + a_1) + (a_1 + a_2)x + (a_2 + a_3)x^2 \end{aligned}$$

respecte a les bases $B_1 = \{1, x, x + x^2, 1 + x^3\}$ i $B_2 = \{1, 1 + x, x + x^2\}$

(b) Trobeu les coordenades respecte la base B_2 de la imatge $f(2 + 2x + x^2 + x^3)$.

(c) Trobeu el nucli de f.

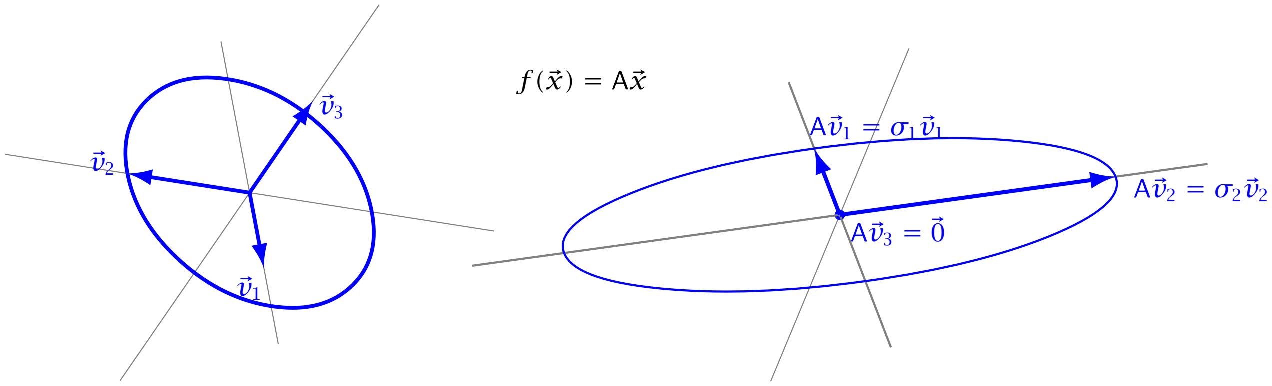
Exercici 20.8. 1. Determineu les dimensions del nucli i la imatge de la transformació lineal $f(\vec{x}) = \mathsf{A}\vec{x}$, si

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 2 & 4 \end{bmatrix}$$

2. Trobeu una base, B_1, de $\mathbb{R}^3$ que siga la unió d'una base de l'espai fila i una de l'espai nul de A.

3. Construïu també una base de $\mathbb{R}^4$, B_2, de manera que la matriu de f respecte a aquestes bases siga

$$M(f, B_1, B_2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



Tercera part
Diagonalització

CAPÍTOL 6

DETERMINANTS

Continguts

21. Determinants

En aquesta unitat introduïm els determinants. Probablement ja n'estàs familiaritzat i coneixes la seua utilitat per a resoldre els problemes que hem anat estudiant al llarg del curs i també alguna tècnica per a calcular-los (si més no, quan treballem amb matrius 2×2 o 3×3). Si és així, a hores d'ara hauríes d'estar convençut que l'ús de les operacions elementals és molt més pràctic que no el dels determinants.
En aquest curs anem a fer-ne un estudi mínim. En veurem les propietats realment bàsiques i aprendrem a calcular-los, per tal de poder fer-los servir en el càlcul dels valors propis.

Tradicionalment s'ha donat molta importància als determinants en l'Àlgebra Lineal, perquè ens proporcionen diversos criteris interessants (per a determinar el rang d'una matriu, per a decidir si una matriu és invertible) i, encara millor, fórmules explícites (per a calcular, per exemple, les solucions d'un sistema d'equacions o la inversa d'una matriu invertible). De fet, els determinants són anteriors a les matrius (que varen ser introduïdes precisament per a *generar* determinants).[1]

En canvi, actualment els determinants tenen molts detractors, perquè el càlcul d'un determinant requereix una gran quantitat d'operacions (fins i tot quan treballem amb matrius de poques dimensions) i l'ús d'altres tècniques (com ara, les operacions elementals i els algorismes del tipus Gauss) permet resoldre els mateixos problemes amb molt menys esforç. El que passa realment és que el determinant ha perdut la seua utilitat com a eina de càlcul, front a aquestes tècniques força més eficients, però encara conserva un gran interès teòric, pel fet que proporciona fórmules explícites per expressar les solucions dels problemes; a més a més, fora de l'àlgebra lineal, hi té força aplicacions.

21.1. El determinant d'una matriu

El determinant d'una matriu quadrada és un nombre real o complex (segons que treballem amb matrius reals o complexes). Representem com $\mathcal{M}(\mathbb{K})$ el conjunt de totes les matrius quadrades (de qualsevol ordre, incloses les matrius quadrades 1×1) amb entrades en $\mathbb{K}$.

Definició 21.1. (L'aplicació determinant)

El determinant és una aplicació $\det : \mathcal{M}(\mathbb{K}) \longrightarrow \mathbb{K}$ *(és a dir, que la imatge* $\det A$ *de la matriu quadrada* A *és un nombre) que verifica les següents propietats:*

1. *Si es permuten dues files de la matriu, aleshores el determinant canvia de signe:*

$$\det \begin{bmatrix} A_1 \\ \dots \\ A_i \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix} = -\det \begin{bmatrix} A_1 \\ \dots \\ A_j \\ \dots \\ A_i \\ \dots \\ A_n \end{bmatrix}$$

2. *Si una fila de la matriu* A *és combinació lineal de dos vectors fila, és a dir, si* $A_i = \alpha B_i + \beta C_i$*,*

$$\det A = \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ \alpha B_i + \beta C_i \\ \dots \\ A_n \end{bmatrix} = \alpha \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ B_i \\ \dots \\ A_n \end{bmatrix} + \beta \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ C_i \\ \dots \\ A_n \end{bmatrix}$$

3. *El determinant d'una matriu triangular superior és el producte de les entrades diagonals:*

$$\det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} = a_{11}a_{22}\cdots a_{nn}$$

[1] D'ací el nom de *matriu*.

☞ El determinant *no és una aplicació lineal*: en general, $\det(\alpha_1 A + \alpha_2 B) \neq \alpha_1 \det A + \alpha_2 \det B$.

☞ El determinant de la matriu A es pot representar com $\det A$ o bé com $|A|$; per exemple, en comptes de $\det \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$, podem escriure $\begin{vmatrix} 1 & -1 \\ 2 & 3 \end{vmatrix}$.

Aquesta definició és una mica estranya. Perquè siga correcta, caldria provar que efectivament existeix una aplicació amb totes aquestes propietats i que, a més, no n'hi ha cap altra aplicació que també les complisca. Nosaltres ens conformarem amb comprovar que quan la matriu A és 2×2 aquesta aplicació és el determinant que ja coneixem del batxillerat; tot seguit calcularem el determinant d'una matriu qualsevol i després n'estudiarem les propietats més interessants. Abans de tot estudiarem els efectes de les operacions elementals sobre els determinants.

21.2. Determinants i operacions elementals

La primera condició de la definició fa referència directament a una operació elemental del tipus permutació. De fet, podem reescriure la condició 1. d'aquesta manera: $\det(E_{i,j}A) = -\det A$.

De la segona condició podem deduir el que passa amb el determinant quan hi fem una operació elemental del tipus escalat: triant $\beta = 0$ tindrem

$$\det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ \alpha A_i \\ \dots \\ A_n \end{bmatrix} = \alpha \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_i \\ \dots \\ A_n \end{bmatrix}$$

és a dir, $\det(E_i(\alpha)A) = \alpha \det A$.

Per tal de determinar l'efecte d'una operació elemental del tipus reducció hem de provar abans una propietat senzilla.

Propietat 21.1.

Si una matriu té dues files iguals, aleshores el determinant d'aquesta matriu és nul.

Demostració: Si suposem que $A_i = A_j$,

$$\det A = \det \begin{bmatrix} A_1 \\ \dots \\ A_i \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix} = -\det \begin{bmatrix} A_1 \\ \dots \\ A_j \\ \dots \\ A_i \\ \dots \\ A_n \end{bmatrix} = -\det A$$

Però $\det A = -\det A$ només és possible si $\det A = 0$. □

Ara ja podem estudiar l'efecte de l'operació de reducció:

$$\det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_i + \alpha A_j \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_i \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix} + \alpha \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_j \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix} = \det \begin{bmatrix} A_1 \\ A_2 \\ \dots \\ A_i \\ \dots \\ A_j \\ \dots \\ A_n \end{bmatrix}$$

és a dir, $\det(E_{i,j}(\alpha)A) = \det A$.

En resum,

Propietats 21.2. (Efecte de les operacions elementals sobre els determinants)

1. *Una operació elemental del tipus permutació canvia el signe del determinant:*

$$\det \mathsf{A} = -\det(\mathsf{E}_{i,j}\mathsf{A})$$

2. *Una operació elemental del tipus escalat multiplica el determinant pel factor d'escala:*

$$\det \mathsf{A} = \frac{1}{\alpha}\det(\mathsf{E}_i(\alpha)\mathsf{A})$$

3. *Una operació elemental del tipus reducció no canvia el valor del determinant:*

$$\det \mathsf{A} = \det(\mathsf{E}_{i,j}(\alpha)\mathsf{A}) \quad \square$$

21.3. Càlcul de determinants

21.3.1. Determinants 2×2

Calculem el determinant de la matriu $\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

Si $a_{22} = 0$ llavors, la matriu A és triangular superior, així que $\det \mathsf{A} = a_{11}a_{22}$. En cas contrari, podem transformar la matriu A en una triangular superior fent una operació elemental:

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \xrightarrow{\mathsf{E}_{2,1}(-a_{21}/a_{11})} \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}}a_{12} \end{bmatrix} = \mathsf{U}$$

Per tant,

$$\det \mathsf{A} = \det(\mathsf{E}_{2,1}(-a_{21}/a_{11})\mathsf{U}) = \det \mathsf{U} = a_{11}\left(a_{22} - \frac{a_{21}}{a_{11}}a_{12}\right) = a_{11}a_{22} - a_{21}a_{12}$$

Fixem-nos que aquesta fórmula també és vàlida quan la matriu A és triangular superior, perquè en eixe cas,

$$\det \mathsf{A} = a_{11}a_{22} = a_{11}a_{22} - \underbrace{a_{21}}_{0}a_{12}$$

Si $\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ llavors, $\det \mathsf{A} = a_{11}a_{22} - a_{21}a_{12}$

Exemple 21.1.

Calculeu el determinant $\det \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix}$

$$\det \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix} = 1 \cdot 1 - 2 \cdot (-3) = 7 \quad \square$$

21.3.2. Aplicació de les operacions elementals al càlcul de determinants

En realitat ja sabem tot el necessari per a calcular el determinant de qualsevol matriu quadrada: mitjançant l'algorisme de Gauss (o qualsevol altra combinació d'operacions elementals) podem transformar fàcilment la matriu en una de triangular superior. Com que sabem com es modifica el valor del determinant cada vegada que hi fem una operació elemental, al final del procés obtindrem el resultat calculant el determinant d'una matriu triangular superior.

EXEMPLE 21.2.

Càlcul del determinant de la matriu $\mathsf{A} = \begin{bmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{bmatrix}$

Mirarem de reduir la matriu a la forma triangular. Primer de tot, eliminem els dos elements que hi ha davall a_{11}. Hauríem de sumar o restar a les files segona i tercera $3/2$ per la primera. Per a evitar l'ús de les fraccions, començarem multiplicant aquestes dues files per 2. Ara bé, cada vegada que multipliquem una fila per 2 el determinant es multiplicarà per 2, així que

$$\begin{vmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{vmatrix} = \frac{1}{2^2} \begin{vmatrix} 2 & 3 & 5 \\ -6 & 8 & 2 \\ 6 & 4 & -2 \end{vmatrix}$$

Ara ja podem eliminar més còmodament, sumant a la segona fila el triple de la primera i restant a la tercera el triple de la primera (això no canvia el determinant):

$$\begin{vmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{vmatrix} = \frac{1}{2^2} \begin{vmatrix} 2 & 3 & 5 \\ 0 & 17 & 17 \\ 0 & -5 & -17 \end{vmatrix}$$

Dividim la segona fila per 17:

$$\begin{vmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{vmatrix} = \frac{17}{2^2} \begin{vmatrix} 2 & 3 & 5 \\ 0 & 1 & 1 \\ 0 & -5 & -17 \end{vmatrix}$$

I tornem a eliminar, sumant 5 vegades la segona fila a la tercera:

$$\begin{vmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{vmatrix} = \frac{17}{2^2} \begin{vmatrix} 2 & 3 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & -12 \end{vmatrix}$$

La matriu de la dreta ja és triangular, així que el seu determinant és el producte de la diagonal:

$$\begin{vmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{vmatrix} = \frac{17}{2^2} 2(-12) = -102 \quad \square$$

EXEMPLE 21.3.

Calculeu el determinant de la matriu $\mathsf{A} = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 1 & 1 & 1 & 1 \\ 4 & 0 & 6 & 1 \\ 2 & 0 & 3 & 1 \end{bmatrix}$

$$\begin{vmatrix} 1 & -1 & 2 & -1 \\ 1 & 1 & 1 & 1 \\ 4 & 0 & 6 & 1 \\ 2 & 0 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 4 & -2 & 5 \\ 0 & 2 & -1 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 2 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 0 \quad \square$$

21.4. UN PARELL DE PROPIETATS FONAMENTALS

Els determinants més fàcils de calcular són els de les matrius elementals, si ens fixem que aquestes matrius les obtenim fent-hi només una operació elemental sobre la matriu identitat; i, com que el determinant de I és 1,

(a) La matriu E_{ij} s'obté intercanviant dues files de la matriu I, així que, $\boxed{\det \mathsf{E}_{i,j} = -\det \mathrm{I} = -1}$.

(b) La matriu $\mathsf{E}_i(\alpha)$ s'obté multiplicant la fila i de la identitat per α, així que $\boxed{\det \mathsf{E}_i(\alpha) = \alpha \det \mathrm{I} = \alpha}$.

(c) La matriu E_{ij} s'obté fent una operació que no canvia el valor del determinant. Per tant, $\boxed{\det E_{i,j}(\alpha) = \det I = 1}$.

En resum,

PROPIETAT 21.3.

Determinants de la matriu identitat i de les matrius elementals

(a) $\det I = 1$ (b) $\det E_{ij} = -1$ (c) $\det E_i(\alpha) = \alpha$ (d) $\det E_{ij}(\alpha) = 1$ □

I d'ací es dedueix que

Si E és una matriu elemental llavors,

$$\det(EA) = \det E \det A \quad \square$$

Això ens permet interpretar el procés de càlcul d'un determinant mitjançant operacions elementals com un producte de matrius:

Si la matriu B s'obté a partir de la matriu A fent-hi diverses operacions elementals, és a dir, si

$$B = E_1 E_2 \dots E_p A$$

llavors,

$$\det B = \det(E_1 E_2 \dots E_p A) = \det E_1 \det E_2 \dots \det E_p \det A$$

El que passa en realitat és que el determinant d'un producte de matrius és igual al producte dels determinants d'aquestes matrius (encara que no siguen elementals). Això ho veurem de seguida, però abans anem a provar la propietat més important d'aquesta unitat: una matriu és invertible quan el seu determinant és no nul.[2]

TEOREMA 21.4. (CARACTERITZACIÓ DE MATRIUS INVERTIBLES)

La matriu A és invertible si i només si $\det A \neq 0$

Demostració: Si A és invertible, aleshores sabem que A és un producte de matrius elementals. Per tant, el determinant de A serà el producte dels determinants d'aquestes matrius elementals. Però com el determinant d'una matriu elemental no és zero, tampoc no ho serà el de A.

En cas contrari, és a dir, si A no és invertible, aplicant-hi operacions elementals podem transformar-la en una matriu triangular que té una fila de zeros. Per tant, el determinant de A és zero. □

Ara ja podem provar que el determinant del producte és el producte dels determinants.

TEOREMA 21.5. (TEOREMA DE BINET)

Per a qualsevol parell de matrius d'ordre n, A i B, $\det(AB) = \det A \det B$

Demostració:

Si A no és invertible, aleshores AB tampoc no és invertible, de manera que els dos determinants, $\det(AB)$ i $\det A$ són iguals a zero i

$$\det A \det B = 0 \det B = 0 = \det(AB)$$

D'altra part, si A és invertible, podem escriure-la com un producte de matrius elementals, $A = E_1 E_2 \dots E_p$, de manera que

$$\det(AB) = \det(E_1 E_2 \dots E_p B) = \det E_1 \det E_) \dots \det E_p \det(B) = \det A \det B \quad \square$$

21.5. CÀLCUL DEL RANG D'UNA MATRIU

En aquest apartat, A és una matriu $m \times n$ (no necessàriament quadrada). Ja sabem que el rang de A és igual al nombre de files linealment independents i, al mateix temps, igual al nombre de columnes linealment independents. Per tant, si $\operatorname{rang} A = p$, hi ha exactament p files linealment independents; si B és la submatriu de A que formen aquestes p files llavors,

$$\operatorname{rang} A = \operatorname{rang} B = p$$

[2] Així que una de les coses que determina el determinant és la invertibilitat.

D'altra banda, B té exactament p columnes linealment independents, així que si C és la submatriu que queda quan ens quedem només amb aquestes columnes,

$$\operatorname{rang} \mathsf{A} = \operatorname{rang} \mathsf{B} = \operatorname{rang} \mathsf{C} = p$$

Aquesta matriu C és quadrada, $p \times p$, i de rang p, de manera que C és invertible i $\det \mathsf{C} \neq 0$.

El que hem provat és que si rang A és p llavors, A té una submatriu quadrada $p \times p$ el determinant de la qual és no nul. És fàcil veure que no hi pot haver cap altra matriu quadrada més gran que també tinga el determinant no nul.

Tot això vol dir que podem identificar el rang d'una matriu amb l'ordre de la matriu quadrada més gran que tinga el determinant no nul:

El rang d'una matriu $\mathsf{A} \in \mathcal{M}_{m\times n}$ coincideix amb l'ordre de la submatriu de A d'ordre màxim que té el determinant no nul.

Exemple 21.4.

Calculeu el rang de la matriu $\mathsf{A} = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 2 \end{bmatrix}$

El determinant d'aquesta matriu és

$$\det \mathsf{A} = \begin{vmatrix} 0 & 1 & 3 & 2 \\ 1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 2 & 1 & 0 \\ 0 & -1 & 2 & 2 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & -5 & -4 \\ 0 & 0 & 5 & 4 \end{vmatrix} = 0$$

De manera que el rang de A no pot ser 4. D'altra banda,

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

així que el rang és 3.

21.6. Determinants i operacions elementals per columnes

El determinant d'una matriu A i el de la seua matriu transposada A^t són iguals. Això és degut a que les transposades de les matrius elementals són matrius elementals del mateix tipus

$$\mathsf{E}_{i,j}^t = \mathsf{E}_{i,j} \qquad \mathsf{E}_i(\alpha)^t = \mathsf{E}_i(\alpha) \qquad \mathsf{E}_{i,j}(\alpha)^t = \mathsf{E}_{j,i}(\alpha)$$

D'ací es dedueix que tot el que hem fet fins ara treballant amb les files de la matriu A es pot fer igualment treballant amb les columnes (perquè el determinant no varia si es transposa la matriu), així que

(a) $\det \mathsf{A}^t = \det \mathsf{A}$.

(b) Si es permuten dues columnes de la matriu, aleshores el determinant canvia de signe.

(c) Si una columna de la matriu A és combinació lineal de dos vectors, és a dir, si $\vec{a}_i = \alpha\vec{b}_i + \beta\vec{c}_i$ llavors,

$$\begin{aligned} \det \mathsf{A} &= \det \begin{bmatrix} \vec{a}_1 & \dots & \alpha\vec{b}_i + \beta\vec{c}_i & \dots & \vec{a}_n \end{bmatrix} \\ &= \alpha \det \begin{bmatrix} \vec{a}_1 & \dots & \vec{b}_i & \dots & \vec{a}_n \end{bmatrix} + \beta \det \begin{bmatrix} \vec{a}_1 & \dots & \vec{c}_i & \dots & \vec{a}_n \end{bmatrix} \end{aligned}$$

(d) El determinant d'una matriu triangular inferior és el producte de les entrades diagonals de la matriu.

(e) Si una columna es multiplica per una constant llavors, el determinant també s'hi multiplica.

(f) Si a una columna se li suma un múltiple d'una altra, el determinant no varia.

21.7. Exercicis

Exercici 21.1. Calculeu els determinants de les matrius següents:

$$\text{(a)}\ \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} \qquad \text{(b)}\ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \qquad \text{(c)}\ \begin{bmatrix} x+1 & x-1 \\ x-1 & x+1 \end{bmatrix}$$

$$\text{(d)}\ \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 1 & 0 & -1 \end{bmatrix} \qquad \text{(e)}\ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 1 & 2 \end{bmatrix} \qquad \text{(f)}\ \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & 1 & 0 \end{bmatrix}$$

Exercici 21.2. Estudieu la invertibilitat de les matrius

$$\text{(a)}\ \mathsf{A} = \begin{bmatrix} 1 & 2 \\ -3 & 1 \end{bmatrix} \qquad \text{(b)}\ \mathsf{B} = \begin{bmatrix} 2 & 3 & 5 \\ -3 & 4 & 1 \\ 3 & 2 & -1 \end{bmatrix} \qquad \text{(c)}\ \mathsf{C} = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 1 & 1 & 1 & 1 \\ 4 & 0 & 6 & 1 \\ 2 & 0 & 3 & 1 \end{bmatrix}$$

Exercici 21.3. Calculeu els determinants següents, cercant en cada cas el mètode més eficient que pugueu trobar.

$$\text{(a)}\ \det \begin{bmatrix} x & x & x & x \\ x & y & y & y \\ x & y & z & z \\ x & y & z & t \end{bmatrix} \qquad \text{(b)}\ \det \begin{bmatrix} 1+x & 1 & 1 & \dots & 1 \\ 1 & 1+x & 1 & \dots & 1 \\ 1 & 1 & 1+x & \dots & 1 \\ \dots & & & & \\ 1 & 1 & 1 & \dots & 1+x \end{bmatrix} \ \text{(la matriu és d'ordre } n\text{)}$$

Exercici 21.4. Per a quins valors dels paràmetres són invertibles les matrius de l'exercici anterior?

Exercici 21.5. La matriu A l'hem obtinguda, a partir de la matriu identitat fent-hi les següents operacions elementals:

1. Hem permutat les files segona i quarta
2. A la fila tercera li hem sumat quatre vegades la primera
3. Hem multiplicat la columna segona per -2
4. Hem multiplicat la fila primera per 3.
5. Herm permutat les files primera i tercera.

Quin és el deteminant de A?

Exercici 21.6. Sabent que el determinant de la matriu 3×3 A és igual a 7, calculeu els determinants següents:

$$\text{(a)}\ \det \mathsf{A}^t \qquad \text{(b)}\ \det(2\mathsf{A}) \qquad \text{(c)}\ \det(\mathsf{E}_{1,3}\mathsf{A}) \qquad \text{(d)}\ \det(\mathsf{E}_{1,3}(-5)\mathsf{A})$$

$$\text{(e)}\ \det \mathsf{A}^2 \qquad \text{(f)}\ \det \mathsf{A}^{-1} \qquad \text{(g)}\ \det(\mathsf{E}_2(3)\mathsf{A}^{-1})$$

Exercici 21.7. Demostreu que la matriu

$$\begin{bmatrix} x+1 & x \\ x & x-1 \end{bmatrix}$$

és invertible (independentment del valor del paràmetre x) i calculeu-ne la inversa.

Exercici 21.8. (Determinant de Vandermonde) Calculeu els determinants

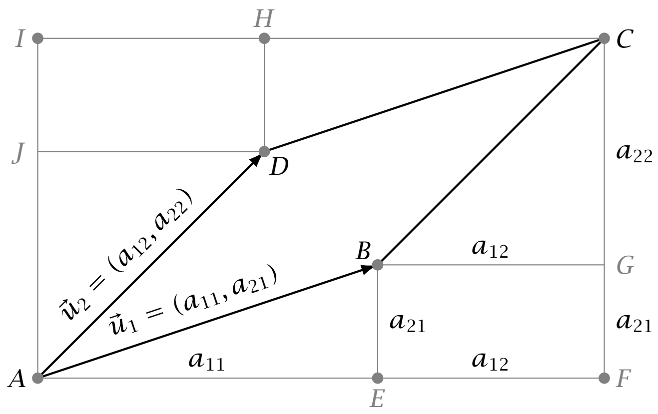
$$\text{(a)} \quad \begin{bmatrix} 1 & a \\ 1 & b \end{bmatrix} \qquad \text{(b)} \quad \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix} \qquad \text{(c)} \quad \begin{bmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{bmatrix}$$

Podeu trobar una fórmula (i provar-la) per al determinant següent?

$$\det \begin{bmatrix} 1 & a & a^2 & \dots & a^n \\ 1 & b & b^2 & \dots & b^n \\ \dots \\ 1 & c & c^2 & \dots & c^n \end{bmatrix}$$

Exercici 21.9. La figura representa el paral·lelogram $ABCD$ definit pels vectors $\vec{u}_1 = (a_{11}, a_{21})$ i $\vec{u}_2 = (a_{21}, a_{22})$. Calculeu la seua àrea (restant a l'àrea del rectangle *gran* $AFCI$ les dels polígons exteriors a $ABCD$).

Expressa el resultat que obtens com un determinant.



Exercici 21.10. Si $\mathsf{A} = \mathsf{LU}$ és una factorització LU de la matriu A, quina relació hi ha entre els determinants de les tres matrius A, L i U? (*No és suficient la resposta* $\det \mathsf{A} = \det \mathsf{L} \det \mathsf{U}$). Potser us convindrà distingir si la factorització és estricta o no ho és.

22. Altres propietats i aplicacions dels determinants

22.1. Desenvolupament per una columna o per una fila

A continuació deduirem una fórmula recursiva per al càlcul del determinant. Aquesta fórmula redueix el càlcul d'un determinant d'ordre n al càlcul d'uns quants (n) determinants d'ordre $n-1$.

Definició 22.1.

Si A *és una matriu quadrada, el menor complementari corresponent a l'element* a_{ij} *és el determinant de la submatriu* A_{ij}*, obtingut eliminant la fila* i *i la columna* j *de la matriu* A.
El nombre $c_{ij} = (-1)^{i+j} \det \mathsf{A}_{ij}$ *s'anomena cofactor corresponent a l'element* a_{ij}.

Suposem que volem calcular el determinant de la matriu

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{14} \\ a_{31} & a_{32} & a_{33} & a_{14} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

La primera columna podem escriure-la com

$$\vec{a}_1 = a_{11}(1,0,0,0) + a_{21}(0,1,0,0) + a_{31}(0,0,1,0) + a_{41}(0,0,0,1)$$

de manera que el determinant serà igual a

$$\begin{aligned}
\det \mathsf{A} &= a_{11} \begin{vmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{21} \begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 1 & a_{22} & a_{23} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{31} \begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 1 & a_{32} & a_{33} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{41} \begin{vmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \\ 1 & a_{42} & a_{43} & a_{44} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{21} \begin{vmatrix} 1 & a_{22} & a_{23} & a_{14} \\ 0 & a_{12} & a_{13} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} + a_{31} \begin{vmatrix} 1 & a_{32} & a_{33} & a_{14} \\ 0 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 0 & a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{41} \begin{vmatrix} 1 & a_{42} & a_{43} & a_{44} \\ 0 & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{14} \\ 0 & a_{32} & a_{33} & a_{14} \end{vmatrix}
\end{aligned}$$

(Els canvis de signe es deuen a que a la segona matriu hi fem una permutació de files, a la tercera dues i a la quarta tres.)

$$= \left|\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline 0 & & & \\ 0 & & \mathsf{A}_{11} & \\ 0 & & & \end{array}\right| - \left|\begin{array}{c|ccc} a_{21} & a_{12} & a_{13} & a_{14} \\ \hline 0 & & & \\ 0 & & \mathsf{A}_{21} & \\ 0 & & & \end{array}\right| + \left|\begin{array}{c|ccc} a_{31} & a_{12} & a_{13} & a_{14} \\ \hline 0 & & & \\ 0 & & \mathsf{A}_{31} & \\ 0 & & & \end{array}\right| - \left|\begin{array}{c|ccc} a_{41} & a_{12} & a_{13} & a_{14} \\ \hline 0 & & & \\ 0 & & \mathsf{A}_{41} & \\ 0 & & & \end{array}\right|$$

Fixem-nos que si apliquem l'algorisme de Gauss per transformar en triangular superior qualsevol d'aquestes matrius no canviarem en cap cas la primera fila de la matriu, de manera que

$$\begin{aligned}
\det \mathsf{A} &= a_{11} \det \mathsf{A}_{11} - a_{21} \det \mathsf{A}_{21} + a_{31} \det \mathsf{A}_{31} - a_{41} \det \mathsf{A}_{41} \\
&= a_{11} c_{11} + a_{21} c_{21} + a_{31} c_{31} + a_{41} c_{41}
\end{aligned}$$

Aquesta fórmula redueix el càlcul del determinant 4×4 a quatre determinants 3×3. Com que $\vec{a}_1 = (a_{11}, a_{21}, a_{31}, a_{41})$ és la primera columna de la matriu A en diem *el desenvolupament del determinant per la primera columna.* Evidentment, això mateix es pot fer amb una matriu de qualsevol altre ordre,

El determinant $\det \mathsf{A}$ es pot calcular sumant els productes de cada element de la primera columna de A pels seus respectius cofactors, és a dir,

$$\det \mathsf{A} = a_{11} c_{11} + a_{21} c_{21} + \cdots + a_{n1} c_{n1}$$

de manera que reduïm el càlcul d'un determinant $n \times n$ al de n determinants $(n-1) \times (n-1)$ (i $2n-1$ operacions més, entre sumes i productes). D'altra banda, es pot provar que la fórmula continua sent vàlida si en comptes de la primera columna n'escollim qualsevol altra, així que

Teorema 22.1. (Desenvolupament del determinant per una columna)

El determinant $\det \mathsf{A}$ *es pot calcular sumant els productes de cada element d'una columna (arbitrària) de* A *pels seus respectius cofactors, és a dir per a qualsevol* j,

$$\det \mathsf{A} = a_{1j} c_{1j} + a_{2j} c_{2j} + \cdots + a_{nj} c_{nj} \tag{22.1}$$

Finalment, com que el determinant de la matriu és el mateix que el de la transposada, si podem calcular-lo desenvolupant-lo per una columna, també ho podem fer amb una fila.

COROL·LARI 22.2. (DESENVOLUPAMENT DEL DETERMINANT PER UNA FILA)

El determinant $\det \mathsf{A}$ *es pot calcular sumant els productes de cada element d'una fila (arbitrària) de* A *pels seus respectius cofactors, és a dir per a qualsevol* i*,*

$$\det \mathsf{A} = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} \tag{22.2}$$

EXEMPLE 22.1.

Càlculeu el determinant de la matriu $\begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 1 \\ -2 & 1 & 2 \end{bmatrix}$

Podem desenvolupar-lo per qualsevol fila o columna. Desenvolupant per la tercera columna obtenim:

$$\begin{aligned} \begin{vmatrix} 2 & 1 & 4 \\ 1 & 0 & 1 \\ -2 & 1 & 2 \end{vmatrix} &= 4\begin{vmatrix} 1 & 0 \\ -2 & 1 \end{vmatrix} - 1\begin{vmatrix} 2 & 1 \\ -2 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 4(1\cdot 1 - 2\cdot 0) - 1(2\cdot 1 + 2\cdot 1) + 2(2\cdot 0 - 1\cdot 1) \\ &= -2 \end{aligned}$$

Si elegim la segona fila,

$$\begin{aligned} \begin{vmatrix} 2 & 1 & 4 \\ 1 & 0 & 1 \\ -2 & 1 & 2 \end{vmatrix} &= -1\begin{vmatrix} 1 & 4 \\ 1 & 2 \end{vmatrix} + 0\begin{vmatrix} 2 & 4 \\ -2 & 2 \end{vmatrix} - 1\begin{vmatrix} 2 & 1 \\ -2 & 1 \end{vmatrix} \\ &= -1(1\cdot 2 - 4\cdot 1) + 0 - 1(2\cdot 1 - 1\cdot(-2)) \\ &= -2 \end{aligned}$$

Observeu que l'elecció d'una fila que conté un zero ens ha estalviat el càlcul d'un determinant.

22.1.1. COMPARACIÓ ENTRE EL MÉTODE DE GAUSS I EL DE DESENVOLUPAMENT PER FILES O COLUMNES

Ara per ara tenim dos métodes per a calcular el determinant: el mètode de Gauss i l'aplicació recursiva d'alguna fórmula de desenvolupament per una fila o columna. En aquest apartat esbrinarem quina estratègia resultarà més eficient. La millor manera de fer aquesta comparació consistirà en comptar el nombre d'operacions que cal realitzar en cada cas.

En primer lloc, calcularem el nombre d'operacions necessàries per a calcular un determinant pel métode de Gauss: Observem que per a canviar per zeros tots els elements per sota a_{11} cal fer el següent càlcul: Canviar la fila A_i ($2 \le i \le n$) per $\mathsf{A}_i - (a_{i1}/a_{11})\mathsf{A}_1$. Així, per cada fila de la segona a la n, haurem de

1. Calcular a_{i1}/a_{11}: una operació.
2. Canviar a_{i1} per zero: cap operació.
3. Canviar a_{ij}, $2 \le j \le n$ per $a_{ij} - (a_{i1}/a_{11})a_{1j}$: una suma i un producte per cada j: $2(n-1)$ operacions.

Això ens dóna un total de $1 + 0 + 2(n-1)$ operacions per cada fila, és a dir,

$$(n-1)(1 + 2(n-1)) = (n-1) + 2(n-1)^2$$

operacions per a reduir la primera columna. És clar que per reduir la segona columna caldrà fer $(n-2) + 2(n-2)^2$ operacions i així successivament. En definitiva, per reduir A a la forma triangular les operacions que hem de fer són

$$\begin{aligned}
\Big[(n-1)+2(n-1)^2\Big]+\Big[(n-2)+2(n-2)^2\Big]+\cdots+\Big[1+2(1)^2\Big]\\
&=\Big[(n-1)+(n-2)+2+\cdots+1\Big]+2\Big[(n-1)^2+(n-2)^2+\cdots+1^2\Big]\\
&=\frac{(n-1)n}{2}+2\frac{n(2n-1)(n-1)}{6}\\
&=\frac{n(n-1)(4n+1)}{3}
\end{aligned}$$

Finalment, cal multiplicar els n elements diagonals: $n-1$ productes. Així doncs,

El nombre total d'operacions per calcular un determinant pel mètode de Gauss és

$$n-1+\frac{n(n-1)(4n+1)}{3}$$

Vegem ara quantes operacions caldria fer per tal de calcular un determinant d'ordre n aplicant successivament el mètode de desenvolupament per una fila fins a reduir-lo completament. A la fórmula

$$\det \mathsf{A} = a_{1j}c_{1j}+a_{2j}c_{2j}+\cdots+a_{nj}c_{nj}$$

s'hi sumen n termes, és a dir, cal fer-hi $n-1$ sumes. Cadascun dels termes que hi sumem és un producte. Per tant cal fer n productes on un dels factors és un determinant d'ordre $n-1$; en total, $n-1$ sumes, n productes i n determinants d'ordre $n-1$. Així que, anomenant a_n al nombre d'operacions corresponent a un determinant d'ordre n, el nombre d'operacions serà aquest:

$$a_n = (n-1)+n+na_{n-1}$$

Aquesta és una recurrència no lineal, i no és gens fàcil d'obtenir el valor exacte de a_n; però podem fer-ne una estimació:

$$a_n = (n-1)+n+na_{n-1} = 2n+na_{n-1}-1 > na_{n-1} > n(n-1)a_{n-2} > \cdots > n!$$

Per calcular el determinant desenvolupant per files cal fer més de $n!$ operacions.

El quadre 22.1 compara els dos mètodes per a matrius d'ordres $2, 3, \ldots, 25$ (s'ha calculat el nombre exacte d'operacions, amb l'ajut del programa `derive`) i mostra clarament que el mètode de desenvolupament per files o columnes no és gens recomanable (excepte potser per a matrius d'ordres molt petits).

☞ En qualsevol cas, en la pràctica pot ser convenient una combinació dels dos mètodes i, en general, de les propietats conegudes dels determinants per tal de simplificar al màxim els càlculs.

Exemple 22.2.

Calculeu el determinant de la matriu $\mathsf{A} = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 4 & 2 \end{bmatrix}$

Aquesta matriu té diverses files i columnes amb només dos entrades no nul·les. Farem una operació elemental per a conseguir un nou zero en la primera columna: li sumarem a la tercera fila la segona (això no canvia el valor del determinant):

$$|\mathsf{A}| = \begin{vmatrix} 0 & 1 & 3 & 2 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & -1 & 4 & 2 \end{vmatrix}$$

n	Gauss	Desenvolupant per files	
2	4	3	
3	15	14	
4	37	63	
5	74	324	
6	130	1 955	
7	209	13 698	
8	315	109 599	
9	452	986 408	(1 milió d'operacions)
10	624	9 864 099	(10 milions d'operacions)
11	835	108 505 110	
12	1 089	1 302 061 343	
13	1 390	16 926 797 484	
14	1 742	236 975 164 803	
15	2 149	3 554 627 472 074	
16	2 615	56 874 039 553 215	
17	3 144	966 858 672 404 688	
18	3 740	17 403 456 103 284 419	
19	4 407	330 665 665 962 403 998	
20	5 149	6 613 313 319 248 079 999	
21	5 970	138 879 579 704 209 680 020	
22	6 874	3 055 350 753 492 612 960 483	
23	7 865	70 273 067 330 330 098 091 154	
24	8 947	1 686 553 615 927 922 354 187 743	
25	10 124	42 163 840 398 198 058 854 693 624	(42 quatrilions d'operacions)

Quadre 22.1: Nombre d'operacions en el càlcul d'un determinant

Ara, com que només hi ha una entrada no nul·la a la primera columna, desenvolupem el determinant per aquesta columna:

$$|\mathsf{A}| = -\begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ -1 & 4 & 2 \end{vmatrix}$$

A la tercera fila li restem la primera,

$$|\mathsf{A}| = -\begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ -2 & 1 & 0 \end{vmatrix}$$

Desenvolupem per la tercera columna,

$$|\mathsf{A}| = -2\begin{vmatrix} 2 & 1 \\ -2 & 1 \end{vmatrix}$$

i a la segona fila li sumem la primera

$$|\mathsf{A}| = -2\begin{vmatrix} 2 & 1 \\ 0 & 2 \end{vmatrix} = -8$$

22.2. Aplicacions dels determinants

22.2.1. Resolució de sistemes lineals

Quan un sistema lineal de n equacions i n incògnites és determinat es pot resoldre fent servir la *regla de Cramer*, que consisteix en una fórmula per a cada una de les incògnites. Aquestes fórmules comporten el càlcul de $n + 1$ determinants d'ordre n, de manera que tenen poca utilitat pràctica (excepte en el cas de sistemes de dues o tres equacions, el nombre d'operacions que cal realitzar és grandíssim). Ara bé, les fórmules de Cramer són útils en algunes demostracions teòriques i, a més a més, permeten calcular alguna incògnita aïlladament (en algunes aplicacions, d'un determinat sistema només ens interessarà el valor d'alguna de les incògnites).

La regla de Cramer

Teorema 22.3. (Regla de Cramer) *Si* A *és una matriu invertible i* $A(i)$ *és la matriu que resulta de substituir la columna* i *de* A *pel vector* $\vec{b}$*, aleshores la solució del sistema* $A\vec{x} = \vec{b}$ *és*

$$x_i = \frac{|A(i)|}{|A|}, \qquad i = 1, 2, \dots, n \tag{22.3}$$

Demostració: Anomenem $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ a les columnes de A. Si $A\vec{x} = \vec{b}$ aleshores,

$$b = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \sum_{j=1}^{n} x_j\vec{a}_j$$

així que

$$\begin{aligned}
\det A(i) &= \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{b} & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&= \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \sum_{j=1}^{n} x_j\vec{a}_j & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&= \quad x_1 \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{a}_1 & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&\quad + x_2 \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{a}_2 & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&\quad + \dots\\
&\quad + x_i \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{a}_i & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&\quad + \dots\\
&\quad + x_n \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{a}_n & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix}\\
&= x_1 0 + x_2 0 + \dots x_i \det\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_{i-1} & \vec{a}_i & \vec{a}_{i+1} & \dots & \vec{a}_n\end{bmatrix} + \cdots + x_n 0\\
&= x_i \det A
\end{aligned}$$

De manera que

$$x_i = \frac{A(i)}{\det A} \quad \square$$

Exemple 22.3.

Solució del sistema $\begin{bmatrix}1 & 1\\ 2 & -1\end{bmatrix}\vec{x} = \begin{bmatrix}2\\ 1\end{bmatrix}$

$$x_1 = \frac{\begin{vmatrix}2 & 1\\ 1 & -1\end{vmatrix}}{\begin{vmatrix}1 & 1\\ 2 & -1\end{vmatrix}} = 1 \qquad x_2 = \frac{\begin{vmatrix}1 & 2\\ 2 & 1\end{vmatrix}}{\begin{vmatrix}1 & 1\\ 2 & -1\end{vmatrix}} = 1 \quad \square$$

Aplicació de la regla de Cramer a sistemes indeterminats Si el sistema $A\vec{x} = \vec{b}$ és indeterminat i el rang A és k, elegim una submatriu invertible de A d'ordre k, A_1. El sistema lineal que resulta de suprimir les files de A que no intervenen en A_1 té les mateixes solucions que $A\vec{x} = \vec{b}$ (perquè té el mateix nombre d'uns principals). En aquest subsistema es poden elegir com a principals les incògnites corresponents a les columnes de A_1 i, escrivint-lo en la forma

$$A_1\vec{x}_1 = b_1 - A_2\vec{x}_2$$

es pot resoldre mitjançant la regla de Cramer.

Exemple 22.4.

Solució del sistema $\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & -1\\ 3 & 3 & 1\end{bmatrix}\vec{x} = \begin{bmatrix}3\\ 5\\ 11\end{bmatrix}$

Com que els rangs de la matriu de coeficients i el de l'ampliada són els dos iguals a 2, el sistema és indeterminat. Elegim una submatriu de A que siga regular i d'ordre 2. Per exemple, $A_1 = \left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right]$, que correspon a les dues primeres files i a les columnes primera i tercera de A. Per tant, suprimim la tercera equació i elegim la segona variable com a lliure. El sistema que en resulta és:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 - x_2 \\ 5 - x_2 \end{bmatrix}$$

Resolent-lo per la regla de Cramer obtindrem:

$$x_1 = \frac{\begin{vmatrix} 3 - x_2 & 1 \\ 5 - x_2 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{-8 - 2x_2}{-2} = -4 - x_2$$

$$x_3 = \frac{\begin{vmatrix} 1 & 3 - x_2 \\ 1 & 5 - x_2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{2}{-2} = -1$$

així que la solució general és

$$\vec{x} = (-4, 0, -1) + \alpha(-1, 1, 0)$$

22.2.2. Càlcul de la inversa d'una matriu

Definició 22.2. (Matriu de cofactors) *Si* $\mathsf{A} \in \mathcal{M}_n$*, la matriu de cofactors de* A*, que representarem com* $\operatorname{cof} \mathsf{A}$*, es defineix com la matriu formada substituint cada entrada de* A *pel seu cofactor, és a dir,*

$$\operatorname{cof} \mathsf{A} = \begin{bmatrix} c_{11} & c_{12} & \dots c_{1n} \\ c_{21} & c_{22} & \dots c_{2n} \\ \dots & & \\ c_{n1} & c_{n2} & \dots c_{nn} \end{bmatrix}$$

Teorema 22.4.

Si A *és una matriu regular, llavors* $\mathsf{A}^{-1} = \frac{1}{|\mathsf{A}|} (\operatorname{cof} \mathsf{A})^t$.

Demostració: Si $\mathsf{A}^{-1} = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \dots & \vec{x}_n \end{bmatrix}$ aleshores $\mathsf{A}\mathsf{A}^{-1} = \mathrm{I}$, de manera que

$$\mathsf{A} \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \dots & \vec{x}_n \end{bmatrix} = \mathrm{I}$$

o, equivalentment,

$$\mathsf{A}\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \quad \mathsf{A}\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \quad \dots, \quad \mathsf{A}\vec{x}_n = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}$$

de manera que $\vec{x}_1$ és la solució del sistema $\mathsf{A}\mathsf{X} = \mathrm{I}^1$. Aplicant-hi la regla de Cramer,

$$\vec{x}_{11} = \frac{1}{\det \mathsf{A}} \begin{vmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & & & \\ 0 & a_{n2} & \dots & a_{nn} \end{vmatrix} = \frac{1}{\det \mathsf{A}} |\mathsf{A}_{11}|$$

$$\vec{x}_{21} = \frac{1}{\det \mathsf{A}} \begin{vmatrix} a_{11} & 0 & \dots & a_{1n} \\ a_{21} & 1 & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & 0 & \dots & a_{nn} \end{vmatrix} = -\frac{1}{\det \mathsf{A}} |\mathsf{A}_{12}|$$

i així successivament. □

22.3. Exercicis

Exercici 22.1. Calculeu el determinant de la matriu $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 5 & 0 & 0 \end{bmatrix}$.

Exercici 22.2. Calculeu el determinant de la matriu $\begin{bmatrix} -3 & 2 & 1 & 2 \\ 5 & 0 & 2 & 0 \\ 6 & -4 & 2 & 1 \\ 4 & 1 & -3 & -4 \end{bmatrix}$.

Exercici 22.3. Discutiu el sistema i, si és possible, apliqueu la Regla de Cramer per a resoldre'l:

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 10 \\ 3x_1 + 7x_2 + 4x_3 &= 3 \\ x_1 + 2x_2 + 2x_3 &= 3 \end{aligned}$$

Exercici 22.4. Discutiu el sistema i, en els casos determinats, apliqueu la Regla de Cramer per a resoldre'l:

$$\begin{aligned} ax_1 + bx_2 + x_3 &= 1 \\ x_1 + abx_2 + x_3 &= b \\ x_1 + bx_2 + ax_3 &= 1 \end{aligned}$$

Exercici 22.5. Fent servir els determinants, justifiqueu que la matriu següent, A, és invertible i calculeu la matriu inversa A^{-1}.

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}$$

Exercici 22.6. (Valors i vectors propis) Trobeu els valors del paràmetre λ per als quals la matriu

$$\mathsf{A}_\lambda = \begin{bmatrix} 1-\lambda & -2 & 0 \\ 0 & -1-\lambda & 0 \\ -1 & 1 & 2-\lambda \end{bmatrix}$$

no és invertible i calculeu l'espai nul de la matriu A_λ per a cadascun d'aquests valors.

Capítol 7

Vectors i valors propis. Diagonalització

Continguts

23. ENDOMORFISMES I MATRIUS DIAGONALITZABLES

En temes anteriors hem estudiat les factoritzacions $A = LU$ i $A = QR$ que estan relacionades amb la resolució de sistemes lineals i amb l'ortogonalitat i que es poden interpretar com a canvis de base en l'espai fila (en el cas de la factorització LU) o en l'espai columna (factorització QR).
Ara estudiarem els vectors propis de les matrius i els endomorfismes i mirarem d'obtenir una factorització $A = PDP^{-1}$, on la matriu D és diagonal. Això vol dir que farem un canvi de base tant en les files com en les columnes.

23.1. ENDOMORFISMES DIAGONALITZABLES I MATRIUS SEMBLANTS

Un *endomorfisme* de l'espai vectorial E és una aplicació lineal en la qual els espais inicial i final són el mateix E. Per exemple, una aplicació lineal $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ és un endomorfisme, però una aplicació lineal de $\mathbb{R}^3$ en $\mathbb{R}^2$ no ho és.

Si l'espai E és de dimensió finita, B és una base de E i $A_1 = M(f, B, B)$ és la matriu de f respecte aquesta base,[1] llavors la imatge d'un vector qualsevulla $\vec{x} \in E$ es pot calcular mitjançant la fórmula

$$f(\vec{x})_B = A_1 \vec{x}_B \tag{23.1}$$

Si elegim una altra base de E, diguem B', i si $A_2 = M(f, B', B')$, tindrem també

$$f(\vec{x})_{B'} = A_2 \vec{x}_{B'}$$

DEFINICIÓ 23.1.

Dues matrius quadrades són semblants si representen el mateix endomorfisme.

A l'apartat 20.4.1. ja vam trobar la relació que hi ha entre matrius que representen la mateixa aplicació lineal; en el nostre cas,

$$M(f, B', B') = M_{BB'} M(f, B, B) M_{B'B} = M_{B'B}^{-1} M(f, B, B) M_{B'B}$$

Això vol dir que *dues matrius* A_1 *i* A_2 *són semblants si existeix una matriu invertible* P *de manera que* $A_2 = P^{-1} A_1 P$ *(o* $A_1 = P A_2 P^{-1}$*).*

☞ L'objectiu d'aquest capítol és de trobar, si és possible, una matriu diagonal D que represente l'endomorfisme f, és a dir, una base B', de manera que la matriu $D = M(f, B', B')$ siga diagonal.

DEFINICIONS 23.2.

(a) *Una matriu quadrada* A *és diagonalitzable si és semblant a una matriu diagonal, és a dir, si existeixen dues matrius* D*, diagonal, i* P*, invertible, de manera que* $A = PDP^{-1}$.

(b) *Un endomorfisme* f *de l'espai de dimensió finita E és diagonalitzable si existeix una base de E de manera que la matriu de* f *respecte aquesta base és diagonal.*

Per entendre aquests conceptes, observem que, si la matriu A és diagonalitzable (és a dir, semblant a una matriu diagonal, D) i anomenem λ_1, λ_2, ... els elements diagonals de D, tindrem

$$\begin{aligned} A &= PDP^{-1} \\ AP &= PD \end{aligned}$$

$$A \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \cdots & \vec{p}_n \end{bmatrix} = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \cdots & \vec{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

$$\begin{bmatrix} A\vec{p}_1 & A\vec{p}_2 & \cdots & A\vec{p}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{p}_1 & \lambda_2\vec{p}_2 & \cdots & \lambda_n\vec{p}_n \end{bmatrix}$$

és a dir,

$$\boxed{A\vec{p}_1 = \lambda_1\vec{p}_1 \qquad A\vec{p}_2 = \lambda_2\vec{p}_2 \qquad \cdots \qquad A\vec{p}_n = \lambda_n\vec{p}_n}$$

[1] Com els espais de sortida i d'arribada són el mateix, hi podem triar la mateixa base.

Això vol dir que, quan multipliquem la matriu A pels vectors p_i, obtenim vectors proporcionals als originals (vectors que tenen la mateixa direcció); el factor de proporcionalitat λ_i és el que anomenarem un *valor propi* de la matriu A.

De manera anàloga, si l'endomorfisme f és diagonalitzable, i la matriu de f respecte a la base $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ és D llavors,

$$f(\vec{v}_1) = \lambda_1 \vec{v}_1 \qquad f(\vec{v}_2) = \lambda_2 \vec{v}_2 \qquad \cdots \qquad f(\vec{v}_n) = \lambda_n \vec{v}_n$$

i el càlcul de la imatge d'un vector, referit a la base B, és especialment senzill:

$$\begin{aligned} f(\alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n) &= \alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2) + \cdots + \alpha_n f(\vec{u}_n) \\ &= \alpha_1\lambda_1\vec{u}_1 + \alpha_2\lambda_2\vec{u}_2 + \cdots + \alpha_n\lambda_n\vec{u}_n \end{aligned}$$

Exemple 23.1.

L'endomorfisme

$$\begin{aligned} f\colon \mathbb{R}_1[x] &\longrightarrow \mathbb{R}_1[x] \\ a + bx &\rightsquigarrow f(a+bx) = (5a-6b) + (3a-4b)x \end{aligned}$$

és diagonalitzable, perquè si elegim com a base $B = \{1+x, 2+x\}$, tindrem

$$\begin{aligned} f(1+x) &= (5-6) + (3-4)x = -1 - x = -1(1+x) + 0(2+x) \\ f(2+x) &= (10-6) + (6-4)x = 4 + 2x = 0(1+x) + 2(2+x) \end{aligned}$$

així que la matriu de f respecte B és $D = \mathsf{M}(f, B, B) = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ i la imatge del polinomi $p(x) = a(1+x) + b(2+x)$ és

$$f(p(x)) = -1a(1+x) + 2b(2+x) \quad \square$$

23.2. Vectors propis i valors propis

Perquè l'endomorfisme f siga diagonalitzable, ha d'haver-hi una base $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ de E de manera que

$$f(\vec{u}_1) = \lambda_1\vec{u}_1,\ f(\vec{u}_2) = \lambda_2\vec{u}_2,\ \dots,\ f(\vec{u}_n) = \lambda_n\vec{u}_n$$

De manera anàloga, si la matriu A és diagonalitzable, tindrem $\mathsf{A} = \mathsf{PDP}^{-1}$, així que $\mathsf{AP} = \mathsf{PD}$. Columna a columna, això vol dir que

$$\mathsf{A}\begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

o bé

$$\mathsf{A}\vec{p}_1 = \lambda_1\vec{p}_1, \quad \mathsf{A}\vec{p}_2 = \lambda_2\vec{p}_2, \quad \dots, \quad \mathsf{A}\vec{p}_n = \lambda_n\vec{p}_n$$

Els vectors $\vec{x}$ per als quals $\mathsf{A}\vec{x} = \lambda\vec{x}$ els anomenem *vectors propis* de la matriu A.

Definicions 23.3. (Vectors propis i valors propis)

(a) Siga f un endomorfisme en l'espai vectorial E.

El vector $\vec{x} \neq \vec{0}$ és un vector propi de l'endomorfisme f, associat al valor propi λ, si $f(\vec{x}) = \lambda\vec{x}$.

(b) Siga A una matriu quadrada.

El vector $\vec{x} \neq \vec{0}$ és un vector propi de la matriu A, associat al valor propi λ, si $\mathsf{A}\vec{x} = \lambda\vec{x}$.

☞ Per diagonalitzar una matriu (o un endomorfisme) necessitem trobar una base de l'espai $\mathbb{K}^n$ formada per vectors propis.

☞ Tot i que $f(\vec{0}) = \vec{0}$, (o $A\vec{0} = \vec{0}$), el vector $\vec{0}$ no és un vector propi de f (ni de A), perquè en la definició s'exigeix que $\vec{x} \neq \vec{0}$.

☞ Si f és un endomorfisme de $\mathbb{K}^n$, llavors els valors i els vectors propis de f són els mateixos que els de la matriu canònica de f.

EXEMPLE 23.2.

Calculeu els valors propis i els vectors propis de la matriu $A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$ i comproveu que aquesta matriu és diagonalitzable.

Hem de trobar els valors de λ i $\vec{x}$ per als quals $A\vec{x} = \lambda\vec{x}$:

$$\begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad \text{és a dir,} \qquad \begin{aligned} -2x_1 + 2x_2 &= \lambda x_1 \\ 2x_1 + 1x_2 &= \lambda x_2 \end{aligned}$$

Això és un sistema lineal, però no està escrit en la forma normalitzada. Passant les incògnites al primer membre, el podem escriure com

$$\begin{aligned} (-2-\lambda)x_1 + 2x_2 &= 0 \\ 2x_1 + (1-\lambda)x_2 &= 0 \end{aligned}$$

i ara és clar que es tracta d'un sistema lineal homogeni. Com que els vectors propis no han de ser nuls, necessitem que aquest sistema siga indeterminat, és a dir, que

$$\operatorname{rang} \begin{bmatrix} -2-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} < 2$$

o, en termes de determinants, que

$$\begin{vmatrix} -2-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0$$

Calculant aquest determinant obtenim l'equació $-6 + \lambda + \lambda^2 = 0$, les arrels de la qual són $\lambda_1 = 2$ i $\lambda_2 = -3$.

☞ Els valors propis de la matriu A són $\lambda_1 = 2$ i $\lambda_2 = -3$.

Per a trobar els vectors propis hem de resoldre els dos sistemes lineals

$$\begin{bmatrix} -2-\lambda_1 & 2 \\ 2 & 1-\lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} -2-\lambda_2 & 2 \\ 2 & 1-\lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

és a dir,

$$\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

En altres paraules, hem de trobar els espais nuls de les matrius $\begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}$ i $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

(a) Vectors propis associats al valor propi $\lambda_1 = 2$:

$$\operatorname{Nul} \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} = \operatorname{Nul} \begin{bmatrix} 1 & -1/2 \\ 0 & 0 \end{bmatrix} = \langle (1,2) \rangle$$

(b) Vectors propis associats al valor propi $\lambda_2 = -3$:

$$\operatorname{Nul} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \operatorname{Nul} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \langle (-2,1) \rangle$$

☞ Els vectors propis associats al valor propi $\lambda_1 = 2$ són tots els vectors $\alpha(1,2)$, $\alpha \neq 0$.

☞ I els associats a $\lambda_2 = -3$, tots els de la forma $\alpha(-2,1)$, $\alpha \neq 0$.

Com que els vectors $(1,2)$ i $(-2,1)$ són linealment independents, la matriu A és diagonalitzable i

$$\boxed{\underbrace{\begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}^{-1}}_{P^{-1}}} \qquad \square$$

23.2.1. Càlcul dels vectors i els valors propis d'una matriu

L'exemple anterior, tot i ser molt simple, ens proporciona les claus per calcular els valors propis i els vectors propis d'una matriu i per decidir si aquesta matriu és diagonalitzable. Perquè $\vec{x}$ siga un vector propi de la matriu $n \times n$ A, associat al valor propi λ, han de passar dues coses:

1. $\mathsf{A}\vec{x} = \lambda\vec{x}$
2. $\vec{x} \neq 0$

La primera condició podem reescriure-la com $\mathsf{A}\vec{x} - \lambda\vec{x} = 0$, o bé, $(\mathsf{A} - \lambda\mathrm{I})\vec{x} = 0$, de manera que $\vec{x}$ és solució d'un sistema d'equacions lineals homogeni, és a dir, $\vec{x}$ ha de ser un element de l'espai nul de la matriu $\mathsf{A} - \lambda\mathrm{I}$.

Definició 23.4.

Si λ és un valor propi de la matriu A anomenem subespai propi associat a λ a l'espai nul de la matriu $\mathsf{A} - \lambda\mathrm{I}$. El representarem com $E_\lambda(\mathsf{A})$:

$$E_\lambda(\mathsf{A}) = \mathrm{Nul}\,(\mathsf{A} - \lambda\mathrm{I})$$

☞ Els vectors propis associats a λ són tots els vectors de $E_\lambda(\mathsf{A})$ tret del vector zero.

D'altra banda perquè es complisca la segona condició, ha d'haver algun vector distint de zero en $E_\lambda(\mathsf{A})$, és a dir, cal que la matriu $\mathsf{A} - \lambda\mathrm{I}$ *no siga* invertible, la qual cosa significa que $\mathrm{rang}(\mathsf{A} - \lambda\mathrm{I}) \leq n$ o que $\det(\mathsf{A} - \lambda\mathrm{I}) = 0$. El determinant $\det(\mathsf{A} - \lambda\mathrm{I})$ és un polinomi en λ de grau n (perquè restem λ a la diagonal de A). Per tant, els valors propis són les arrels d'una equació de grau n.

Definicions 23.5. (Polinomi característic i equació característica)

(a) El polinomi característic de la matriu A és $\det(\mathsf{A} - \lambda\mathrm{I})$.

(b) L'equació característica de la matriu A és $\det(\mathsf{A} - \lambda\mathrm{I}) = 0$.

☞ Els valors propis de la matriu A són les arrels de l'equació característica. I totes les matrius que representen el mateix endomorfisme tenen els mateixos valors propis, perquè

Propietat 23.1.

Si les matrius A_1 i A_2 són semblants, llavors les dues matrius tenen el mateix polinomi característic.

Demostració: Si les matrius són semblants, existeix una matriu invertible, P, tal que $\mathsf{A}_1 = \mathsf{P}\mathsf{A}_1\mathsf{P}^{-1}$. Llavors, el polinomi característic de A_1 és

$$\det(\mathsf{A}_1 - \lambda\mathrm{I}) = \det(\mathsf{P}\mathsf{A}_2\mathsf{P}^{-1} - \lambda\mathsf{P}\mathrm{I}\mathsf{P}^{-1}) = \det\left(\mathsf{P}(\mathsf{A}_2 - \lambda\mathrm{I})\,\mathsf{P}^{-1}\right) = \det\mathsf{P}\det(\mathsf{A}_2 - \lambda\mathrm{I})\det\mathsf{P}^{-1} = \det(\mathsf{A}_2 - \lambda\mathrm{I}) \qquad \square$$

Finalment, la matriu serà diagonalitzable si *hi trobem un nombre suficient de vectors propis*, és a dir, si podem construir una base de $\mathbb{K}^n$ formada per vectors propis.

Exemple 23.3.

Estudieu si la matriu $\mathsf{A} = \begin{bmatrix} -5 & 0 & -6 \\ 3 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}$ és diagonalitzable.

El polinomi característic de la matriu A és

$$\begin{vmatrix} -5-\lambda & 0 & -6 \\ 3 & 1-\lambda & 3 \\ 3 & 0 & 4-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -5-\lambda & -6 \\ 3 & 4-\lambda \end{vmatrix} = (1-\lambda)(-2+\lambda+\lambda^2) = (1-\lambda)^2(-2-\lambda)$$

Per tant, els valors propis són $\lambda_1 = 1$ i $\lambda_2 = -2$.

Els subespais propis són

$$E_{\lambda_1}(\mathsf{A}) = \mathrm{Nul}(\mathsf{A} - \mathrm{I}) = \mathrm{Nul}\begin{bmatrix} -6 & 0 & -6 \\ 3 & 0 & 3 \\ 3 & 0 & 3 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle (0,1,0), (-1,0,1) \rangle$$

$$E_{\lambda_2}(\mathsf{A}) = \mathrm{Nul}(\mathsf{A} + 2\mathrm{I}) = \mathrm{Nul}\begin{bmatrix} -3 & 0 & -6 \\ 3 & 3 & 3 \\ 3 & 0 & 6 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \langle (-2,1,1) \rangle$$

Com que els tres vectors propis $(0,1,0)$, $(-1,0,1)$ i $(-2,1,1)$ són linealment independents, la matriu A és diagonalitzable i

$$\boxed{\underbrace{\begin{bmatrix} -5 & 0 & -6 \\ 3 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{\mathsf{P}} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}}_{\mathsf{D}} \underbrace{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1}}_{\mathsf{P}^{-1}}}$$

Exemple 23.4.

Estudieu si la matriu $\mathsf{B} = \begin{bmatrix} 3 & 0 & -4 \\ 3 & -2 & -6 \\ 1 & 0 & -1 \end{bmatrix}$ és diagonalitzable.

El polinomi característic de la matriu B és

$$\begin{vmatrix} 3-\lambda & 0 & -4 \\ 3 & -2-\lambda & -6 \\ 1 & 0 & -1-\lambda \end{vmatrix} = (-2-\lambda)\begin{vmatrix} 3-\lambda & -4 \\ 1 & -1-\lambda \end{vmatrix} = (-2-\lambda)(1-2\lambda+\lambda^2) = (1-\lambda)^2(-2-\lambda)$$

Per tant, els valors propis són $\lambda_1 = 1$ i $\lambda_2 = -2$.

Els subespais propis són

$$E_{\lambda_1}(\mathsf{B}) = \mathrm{Nul}(\mathsf{B} - 1\mathrm{I}) = \mathrm{Nul}\begin{bmatrix} 2 & 0 & -4 \\ 3 & -3 & -6 \\ 1 & 0 & -2 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle (2,0,1) \rangle$$

$$E_{\lambda_2}(\mathsf{B}) = \mathrm{Nul}(\mathsf{B} + 2\mathrm{I}) = \mathrm{Nul}\begin{bmatrix} 5 & 0 & -4 \\ 3 & 0 & -6 \\ 1 & 0 & 1 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \langle (0,1,0) \rangle$$

Tots els valors propis d'aquesta matriu són múltiples de $(2,0,1)$ o de $(0,1,0)$, així que no és possible construir una base de $\mathbb{K}^3$ formada completament per vectors propis. La matriu no és diagonalitzable.

23.3. Matrius diagonalitzables

En els dos darrers exemples, el polinomi característic és el mateix, $(1-\lambda)^2(-2-\lambda)$, i, en conseqüència, els valors propis també són iguals; tot i això, una de les matrius és diagonalitzable i l'altra no ho és. El fet que el valor propi, $\lambda_1 = 1$, siga arrel doble del polinomi característic sembla exigir que el subespai propi E_{λ_1} tinga dimensió dos (com passa en l'exemple 23.3. però no en el 23.4.).

Definició 23.6. (Multiplicitat algèbrica i multiplicitat geomètrica)

Si λ és un valor propi de la matriu A*,*

- *La multiplicitat algèbrica de λ,* $\mathrm{malg}(\lambda, \mathsf{A})$*, és la multiplicitat de λ com a arrel del polinomi característic.*
- *La multiplicitat geometrica de λ,* $\mathrm{mgeo}(\lambda, \mathsf{A})$*, és la dimensió del subespai propi $E_\lambda(\mathsf{A})$.*

Exemple 23.5.

Determineu les multiplicitats algèbriques i geomètriques dels valors propis als exemples 23.3. i 23.4.

En tots dos casos, el polinomi característic és $(1-\lambda)^2(-2-\lambda)$. Per tant, les multiplicitats algèbriques són aquestes:

$$\begin{array}{ll} \operatorname{malg}(1, \mathsf{A}) = 2 & \operatorname{malg}(1, \mathsf{B}) = 2 \\ \operatorname{malg}(-2, \mathsf{A}) = 1 & \operatorname{malg}(-2, \mathsf{B}) = 1 \end{array}$$

Les multiplicitats geomètriques són

$$\begin{array}{ll} \operatorname{mgeo}(1, \mathsf{A}) = \dim E_1(\mathsf{A}) = 2 & \operatorname{mgeo}(1, \mathsf{B}) = \dim E_1(\mathsf{B}) = 1 \\ \operatorname{mgeo}(-2, \mathsf{A}) = \dim E_{-2}(\mathsf{A}) = 1 & \operatorname{mgeo}(-2, \mathsf{B}) = \dim E_{-2}(\mathsf{B}) = 1 \quad \square \end{array}$$

23.3.1. Suma directa dels subespais propis

Propietat 23.2.

La suma dels subespais propis de A *és directa.*

Demostració: Si λ_1, λ_2, ..., λ_r, són els valors propis distints de la matriu A, el que volem provar és que la suma

$$E_{\lambda_1}(\mathsf{A}) + E_{\lambda_2}(\mathsf{A}) + \cdots + E_{\lambda_r}(\mathsf{A})$$

és directa, és a dir, que

$$\boxed{\text{Si } \vec{x}_1 \in E_{\lambda_1}(\mathsf{A}),\ \vec{x}_2 \in E_{\lambda_2}(\mathsf{A}),\ \ldots,\ \vec{x}_r \in E_{\lambda_r}(\mathsf{A}), \text{ i } \vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_r = \vec{0}, \text{ llavors, } \vec{x}_1 = \vec{x}_2 = \cdots = \vec{x}_r = \vec{0}}$$

Per provar-ho, multipliquem l'expressió $\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_r = \vec{0}$ per A, A^2, ..., A^{r-1}:

$$\begin{aligned} \vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_r &= \vec{0} \\ \mathsf{A}\vec{x}_1 + \mathsf{A}\vec{x}_2 + \cdots + \mathsf{A}\vec{x}_r &= \vec{0} \\ \mathsf{A}^2\vec{x}_1 + \mathsf{A}^2\vec{x}_2 + \cdots + \mathsf{A}^2\vec{x}_r &= \vec{0} \\ &\cdots \\ \mathsf{A}^{r-1}\vec{x}_1 + \mathsf{A}^{r-1}\vec{x}_2 + \cdots + \mathsf{A}^{r-1}\vec{x}_r &= \vec{0} \end{aligned}$$

Però, com que els vectors són propis, això equival a

$$\begin{aligned} \vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_r &= \vec{0} \\ \lambda_1\vec{x}_1 + \lambda_2\vec{x}_2 + \cdots + \lambda_r\vec{x}_r &= \vec{0} \\ \lambda_1^2\vec{x}_1 + \lambda_2^2\vec{x}_2 + \cdots + \lambda_r^2\vec{x}_r &= \vec{0} \\ &\cdots \\ \lambda_1^{r-1}\vec{x}_1 + \lambda_2^{r-1}\vec{x}_2 + \cdots + \lambda_r^{r-1}\vec{x}_r &= \vec{0} \end{aligned}$$

Matricialment,

$$\begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \ldots & \vec{x}_r \end{bmatrix} \underbrace{\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \lambda_2^2 & \ldots & \lambda_2^{r-1} \\ \cdots \\ 1 & \lambda_r & \lambda_r^2 & \ldots & \lambda_r^{r-1} \end{bmatrix}}_{\Lambda} = \mathsf{O}$$

La matriu Λ és invertible, perquè és una matriu de Vandermonde amb paràmetres, $\lambda_1, \lambda_2, \ldots, \lambda_r$, distints. Així que

$$\begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \ldots & \vec{x}_r \end{bmatrix} = \mathsf{O} \quad \square$$

Com a conseqüència d'aquesta propietat,

$$\boxed{\dim\left(E_{\lambda_1}(\mathsf{A}) + E_{\lambda_2}(\mathsf{A}) + \cdots + E_{\lambda_r}(\mathsf{A})\right) = \dim E_{\lambda_1}(\mathsf{A}) + \dim E_{\lambda_2}(\mathsf{A}) + \cdots + \dim E_{\lambda_r}(\mathsf{A})}$$

I el que ha de passar, perquè la matriu A siga diagonalitzable, és que qualsevol vector siga combinació lineal dels vectors propis, és a dir, que $\mathbb{K}^n = E_{\lambda_1}(\mathsf{A}) \oplus E_{\lambda_2}(\mathsf{A}) \oplus \cdots \oplus E_{\lambda_r}(\mathsf{A})$, per tant,

Propietat 23.3. (Primera caracterització de la diagonabilitzabilitat)

Una condició necessària i suficient perquè la matriu $n \times n$ A siga diagonalitzable és que la suma de les multiplicitats geomètriques de tots els valors propis siga n. $\square$

La matriu A de l'exemple 23.3. és diagonalitzable perquè $\mathrm{mgeo}(1, A) + \mathrm{mgeo}(-2, A) = 3$. En canvi, en el cas de la matriu B de l'exemple 23.4., $\mathrm{mgeo}(1, B) + \mathrm{mgeo}(-2, B) = 2$ i per això, B no és diagonalitzable.

D'altra banda, també es podem provar que les multiplicitats geomètriques mai no són més grans que les multiplicitats algèbriques:

PROPIETAT 23.4.

SI λ_1 és un valor propi de la matriu A, llavors,

$$1 \leq \mathrm{mgeo}(\lambda_1) \leq \mathrm{malg}(\lambda_1)$$

Demostració: Que les multiplicitats són majors que zero és obvi, perquè hi ha algun vector propi. D'altra banda, com que $E_{\lambda_1}(A) = \mathrm{Nul}(A - \lambda_1 I)$ i

$$\mathbb{K}^n = \mathrm{Nul}(A - \lambda_1 I) \oplus \mathrm{Fil}(A - \lambda_1 I)$$

podem construir una base de $\mathbb{R}^n$ fent la unió d'una base de l'espai nul i una altra de l'espai fila. Si la multiplicitat geomètrica de λ_1 és m,

$$B = \{ \underbrace{\vec{u}_1, \ldots, \vec{u}_m}_{\text{Base de } \mathrm{Nul}(A-\lambda_1 I)}, \underbrace{\vec{v}_{m+1}, \ldots, \vec{v}_n}_{\text{Base de } \mathrm{Fil}(A-\lambda_1 I)} \}$$

Aleshores, la matriu A és semblant a

$$A' = \left[\begin{array}{ccc|ccc} \lambda_1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & & & \\ \cdots & \cdots & \cdots & & B & \\ 0 & \cdots & 0 & & & \end{array} \right]$$

Així que el polinomi característic és

$$\det(A' - \lambda I) = \det \left[\begin{array}{ccc|ccc} \lambda_1 - \lambda & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \lambda_1 - \lambda & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & & & \\ \cdots & \cdots & \cdots & & B - \lambda I & \\ 0 & \cdots & 0 & & & \end{array} \right] = (\lambda_1 - \lambda)^m \det(B - \lambda I)$$

i la multiplicitat algebraica de λ_1 és, almenys, igual a m. □

DEFINICIÓ 23.7.

Direm que un valor propi de la matriu A, λ, és geomètricament complet si $\mathrm{malg}(\lambda, A) = \mathrm{mgeo}(\lambda, A)$.

PROPIETAT 23.5.

Una condició necessària perquè la matriu $n \times n$ A siga diagonalitzable és que tots els valors propis de A siguen geomètricament complets. □

23.3.2. VALORS PROPIS COMPLEXOS

La qüestió de la diagonabilitzabilitat d'una matriu A és ben diferent si treballem amb matrius reals o si ho fem amb matrius complexes. Observem l'exemple següent.

EXEMPLE 23.6.

Estudieu la diagonabilitzabilitat de la matriu

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(a) considerant-la com una matriu real o (b) considerant-la com una matriu complexa.

El polinomi característic d'aquesta matriu és

$$\begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda - \mathrm{i})(\lambda + \mathrm{i})$$

de manera que, si la considerem com una matriu real, A no és diagonalitzable, perquè no té cap valor propi. En canvi, com a matriu complexa té dos valors propis (i, −i), els espais propis associats als quals són

$$E_{\mathrm{i}}(\mathsf{A}) = \mathrm{Nul}\begin{bmatrix} -\mathrm{i} & 1 \\ -1 & -\mathrm{i} \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & \mathrm{i} \\ 0 & 0 \end{bmatrix} = \langle(-\mathrm{i}, 1)\rangle$$
$$E_{-\mathrm{i}}(\mathsf{A}) = \mathrm{Nul}\begin{bmatrix} \mathrm{i} & 1 \\ -1 & \mathrm{i} \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & -\mathrm{i} \\ 0 & 0 \end{bmatrix} = \langle(\mathrm{i}, 1)\rangle$$

per tant la matriu A és semblant a la matriu diagonal $\mathsf{D} = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}$:

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -\mathrm{i} & \mathrm{i} \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}\begin{bmatrix} -\mathrm{i} & \mathrm{i} \\ 1 & 1 \end{bmatrix}^{-1} \qquad \square$$

Com que els valors propis són les arrels del polinomi característic, la diferència fonamental entre els casos real i complex és el fet que un polinomi (real) pot tenir arrels que no ho són, de reals, de manera que, en el cas complex, l'únic problema que pot haver-hi, perquè una matriu no siga diagonalitzable, és que no hi haja *prou* vectors propis, mentre que en el cas real també pot ser que el nombre de valors propis siga insuficient (perquè algunes arrels del polinomi característic no són reals). Aquestes observacions ens porten a caracteritzar les matrius diagonalitzables de manera diferent en els dos casos, real i complex.

Teorema 23.6. (Segona caracterització de les matrius diagonalitzables)

Cas complex: *La matriu quadrada* A *és diagonalitzable si i només si tots els valors propis de* A *són geomètricament complets.*

Cas real: *La matriu quadrada* A *és diagonalitzable si i només si tots els valors propis de* A *són reals i geomètricament complets.*

Segons aquesta propietat, per provar la diagonabilitzabilitat hem de comprovar que tots els valors propis compleixen la igualtat $\mathrm{mgeo}(\lambda, \mathsf{A}) = \mathrm{malg}(\lambda, \mathsf{A})$; tot i això, quan la multiplicitat algèbrica d'un valor propi és igual a 1 aquesta propietat es compleix automàticament.

El quadre següent mostra el procediment pràctic per estudiar la diagonalitzabilitat.

Estudi de la diagonalitzabilitat d'una matriu

1. Resoleu l'equació característica $\det(\mathsf{A} - \lambda \mathrm{I}) = 0$.

   Les arrels d'aquesta equació són els valors propis.
   Les multiplicitats d'aquestes arrels són les multiplicitats algèbriques corresponents.

2. Si treballeu amb matrius reals i alguna de les arrels no és real, llavors la matriu A no és diagonalitzable.

3. Calculeu les dimensions dels subespais propis $E_\lambda(\mathsf{A})$.

   Aquestes dimensions són les multiplicitats geomètriques.

4. Estudieu si cada valor propi amb multiplicitat algèbrica més gran que 1 és geomètricament complet, és a dir, si les multiplicitats algèbrica i geomètrica coincideixen. Si és així, la matriu és diagonalitzable; en cas contrari, no ho és.

Diagonalització de la matriu

En cas que la matriu siga diagonalitzable,

5. La matriu D és la diagonal dels valors propis, $\mathsf{D} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ (repetits tantes vegades com la seua multiplicitat).

6. Cerqueu una base de cada subespai propi.

7. Construïu la base B com a unió de les bases que heu obtingut, $B = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$, o la matriu $P = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix}$

   Pareu compte que l'ordenació dels vectors i els valors propis ha de ser coherent: cada $\vec{p}_i$ ha de ser un vector propi associat al valor propi λ_i.

23.4. Exercicis

Exercici 23.1. Estudieu si són diagonalitzables els següents endomorfismes de $\mathbb{R}^n$:

1. $f(x_1, x_2, x_3) = (3x_1 + 2x_2 + 4x_3, 2x_1 + 2x_3, 4x_1 + 2x_2 + 3x_3)$
2. $f(x_1, x_2, x_3) = (-x_1 - 3x_2 - 9x_3, 5x_2 + 18x_3, -2x_2 - 7x_3)$
3. $f(x_1, x_2, x_3) = (2x_1 + 5x_2 + 6x_3, -3x_2 + 2x_3, 5x_3)$
4. $f(x_1, x_2, x_3) = (x_1 + x_2, -x_1 + x_2, 5x_3)$
5. $f(x_1, x_2) = (2x_1 - x_2, -4x_1 + 2x_2)$
6. $f(x_1, x_2) = (x_1 + ax_2, ax_2)$

Exercici 23.2. Estudieu si les matrius

(a) $\mathsf{A} = \begin{bmatrix} 4 & 0 & -4 \\ -1 & 3 & 2 \\ 1 & 0 & 0 \end{bmatrix}$ (b) $\mathsf{B} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ (c) $\mathsf{C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ (d) $\mathsf{O} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

són diagonalitzables. En cas que ho siguen, diagonalitzeu-les.

Exercici 23.3. Determineu per a quins valors del paràmetre a és diagonalitzable la matriu real

$$\mathsf{A} = \begin{bmatrix} a & -a & 6 \\ 0 & 1 & 3 \\ 0 & 1 & -1 \end{bmatrix}$$

Exercici 23.4. Determineu per a quins valors del paràmetre a és diagonalitzable la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & a & 0 \\ a & 0 & a \end{bmatrix}$$

Diagonalitzeu-la en el cas $a = 0$.

Exercici 23.5. Proveu que els valors propis d'una matriu triangular són els elements de la diagonal principal d'aquesta matriu.

Exercici 23.6. Estudieu si els endomorfismes de $\mathbb{R}^3$ $f(x_1, x_2, x_3) = (-2x_1, x_1 + x_2, x_1 + x_2 + x_3)$ i $g(x_1, x_2, x_3) = (-2x_1, x_1 + x_2, x_1 + x_3)$ són diagonalitzables. En cas que ho siguen, trobeu una base de vectors propis.

Exercici 23.7. Proveu que si la matriu $n \times n$ A té n valors propis distints, llavors aquesta matriu és diagonalitzable.

Exercici 23.8. Proveu que la matriu $n \times n$

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & & & \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

és diagonalitzable i trobeu dues matrius P, invertible, i D, diagonal, de manera que $\mathsf{A} = \mathsf{PDP}^{-1}$ (podeu *inspirar-vos* en un exercici anterior).

24. Diagonalització ortogonal de les matrius simètriques (reals)

24.1. Valors propis complexos en matrius reals

El fet que els polinomis reals poden tenir arrels imaginàries és la causa que algunes matrius reals es poden diagonalitzar com a matrius complexes, però no com a matrius reals. Ara, quan passa això, si λ és un valor propi imaginari, amb vector propi associat $\vec{z}$, llavors $\overline{\lambda}$ també és un valor propi, amb $\overline{\vec{z}}$ com a vector propi, perquè, com que $\mathsf{A} = \overline{\mathsf{A}}$,

$$\mathsf{A}\overline{\vec{z}} = \overline{\mathsf{A}}\,\overline{\vec{z}} = \overline{\mathsf{A}\vec{z}} = \overline{\lambda\vec{z}} = \overline{\lambda}\,\overline{\vec{z}}$$

Exemple 24.1.

Els valors propis de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ -5 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

són $\lambda_1 = 2 + 2i$, $\lambda_2 = \overline{\lambda_1} = 2 - 2i$ i $\lambda_3 = 1$, i els subespais propis, $E_{\lambda_1} = \langle(1 + 2i, 5, 0)\rangle$, $E_{\lambda_2} = \overline{E_{\lambda_1}} = \langle(1 - 2i, 5, 0)\rangle$ i $E_{\lambda_3} = \langle(0, 0, 1)\rangle$, així que

$$\mathsf{A} = \begin{bmatrix} 1+2i & 1-2i & 0 \\ 5 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2+2i & 0 & 0 \\ 0 & 2-2i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1+2i & 1-2i & 0 \\ 5 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \qquad \square$$

És a dir, que els valors i els vectors propis imaginaris de les matrius reals sempre apareixen per parelles de nombres i vectors conjugats. En l'apartat següent aprofitarem això per a provar que les matrius reals simètriques només tenen valors propis reals.

24.2. Diagonalització de matrius simètriques

Les matrius simètriques reals són sempre diagonalitzables. Primer de tot, veurem que els valors propis de les matrius simètriques reals són tots reals.

Propietat 24.1.

Tots els valors propis de les matrius simètriques reals són reals.

Demostració: Si λ és un valor propi de la matriu simètrica A i $\vec{u}$ és un vector propi associat a λ, llavors, segons acabem de veure, $\overline{\lambda}$ també n'és valor propi, amb $\overline{\vec{u}}$ com a vector propi:

$$\mathsf{A}\vec{u} = \lambda\vec{u}$$
$$\mathsf{A}\overline{\vec{u}} = \overline{\lambda}\,\overline{\vec{u}}$$

Si multipliquem $\overline{\vec{u}}^t$ per la primera igualtat i $\vec{u}^t$ per la segona tindrem

$$\overline{\vec{u}}^t\mathsf{A}\vec{u} = \lambda\overline{\vec{u}}^t\vec{u}$$
$$\vec{u}^t\mathsf{A}\overline{\vec{u}} = \overline{\lambda}\vec{u}^t\overline{\vec{u}}$$

I si ara transposem una d'aquestes expressions tindrem

$$\left.\begin{array}{rr} & \overline{\vec{u}}^t\mathsf{A}\vec{u} = \lambda\overline{\vec{u}}^t\vec{u} \\ \left(\vec{u}^t\mathsf{A}\overline{\vec{u}}\right)^t = \left(\overline{\lambda}\vec{u}^t\overline{\vec{u}}\right)^t, & \overline{\vec{u}}^t\mathsf{A}\vec{u} = \overline{\lambda}\overline{\vec{u}}^t\vec{u} \end{array}\right\} \Rightarrow \lambda\overline{\vec{u}}^t\vec{u} = \overline{\lambda}\overline{\vec{u}}^t\vec{u} \Rightarrow (\lambda - \overline{\lambda})\overline{\vec{u}}^t\vec{u} = 0$$

de manera que, o bé $\lambda = \overline{\lambda}$, o $\overline{\vec{u}}^t\vec{u} = 0$. Ara bé, $\overline{\vec{u}}^t\vec{u}$, és el quadrat de la norma de $\vec{u}$, que només és zero quan $\vec{u} = \vec{0}$. Però això és impossible, perquè $\vec{u}$ és un vector propi. En conseqüència, $\lambda = \overline{\lambda}$ i λ és real. $\square$

A més a més, és fàcil provar que, per a cada valor propi real, hi ha algun vector propi real.

Teorema 24.2. (El teorema espectral)

Si A *és una matriu simètrica real aleshores,* A *és diagonalitzable ortogonalment, és a dir,*

$$\mathsf{A} = \mathsf{QDQ}^t$$

on D *és una matriu diagonal i* Q *és una matriu ortogonal.*

Demostració: Ho demostrarem per inducció sobre n.

1) El cas $n = 1$ és evident: $\begin{bmatrix} a \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} a \end{bmatrix}\begin{bmatrix} 1 \end{bmatrix}$.
2) Suposem-ho cert per a les matrius $(n-1)\times(n-1)$ i provem-ho per a una matriu A $n \times n$. Com que ja hem provat que els valors propis de A són reals, elegim-ne un qualsevol, λ, i un vector propi real i unitari associat a λ, $\vec{q}$. Si $\vec{v}$ és un vector (real) ortogonal a $\vec{q}$, llavors $\mathsf{A}\vec{v}$ també és ortogonal a $\vec{q}$, perquè

$$\mathsf{A}\vec{v}\cdot\vec{q} = (\mathsf{A}\vec{v})^t\vec{q} = \vec{v}^t\mathsf{A}\vec{q} = \vec{v}\cdot\mathsf{A}\vec{q} = \vec{v}\cdot\lambda\vec{q} = 0$$

Llavors, si $\{\vec{v}_2, \vec{v}_3, \dots, \vec{v}_n\}$ és una base ortonormal de $\langle\vec{q}\rangle^\perp$, el conjunt $\{\vec{q}, \vec{v}_2, \vec{v}_3, \dots, \vec{v}_n\}$ és una base ortonormal de $\mathbb{R}^n$ i

$$\mathsf{A} = \begin{bmatrix} \vec{q} & \vec{v}_2 & \vec{v}_3 & \dots & \vec{v}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & & & & \\ 0 & & & \mathsf{B} & \\ \dots & & & & \\ 0 & & & & \end{bmatrix}\begin{bmatrix} \vec{q}^t \\ \vec{v}_2^t \\ \vec{v}_3^t \\ \dots \\ \vec{v}_n^t \end{bmatrix} = \lambda_1\vec{q}\vec{q}^t + \begin{bmatrix} \vec{v}_2 & \vec{v}_3 & \dots & \vec{v}_n \end{bmatrix}\mathsf{B}\begin{bmatrix} \vec{v}_2^t \\ \vec{v}_3^t \\ \dots \\ \vec{v}_n^t \end{bmatrix}$$

on B és una matriu simètrica $(n-1)\times(n-1)$, així que B es pot diagonalitzar ortogonalment. D'ací es conclou fàcilment que la matriu A és diagonalitzable ortogonalment. ☐

☞ Si la matriu A és simètrica, podem reescriure $\mathsf{A} = \mathsf{QDQ}^t$ com

$$\mathsf{A} = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_n \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}\begin{bmatrix} \vec{q}_1^t \\ \vec{q}_2^t \\ \dots \\ \vec{q}_n^t \end{bmatrix}$$

$$A = \lambda_1\vec{q}_1\vec{q}_1^t + \lambda_2\vec{q}_2\vec{q}_2^t + \dots + \lambda_n\vec{q}_n\vec{q}_n^t$$

així que A és una combinació lineal de matrius de rang 1, les columnes de les quals són vectors propis; els pesos de la combinació són els valors propis:

> Si A és una matriu simètrica real, llavors
>
> $$A = \lambda_1\vec{q}_1\vec{q}_1^t + \lambda_2\vec{q}_2\vec{q}_2^t + \dots + \lambda_n\vec{q}_n\vec{q}_n^t$$
>
> on λ_1, ..., λ_n són els valors propis i $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$ una base ortonormal de vectors propis.

EXEMPLE 24.2.

> Diagonalització ortonormal de la matriu simètrica $\mathsf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ i expresseu-la com a combinació lineal de matrius de rang 1,
>
> $$A = \lambda_1\vec{q}_1\vec{q}_1^t + \lambda_2\vec{q}_2\vec{q}_2^t + \dots + \lambda_n\vec{q}_n\vec{q}_n^t$$
>
> on λ_1, ..., λ_n són els valors propis i $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$ una base ortonormal de vectors propis.

Els valors propis d'aquesta matriu són $\lambda_1 = 1$ i $\lambda_2 = 3$ i els subespais propis corresponents, $E_1 = \langle(1,-1)\rangle$ i $E_3 = \langle(1,1)\rangle$, de manera que el conjunt

$$B = \Big\{ \underbrace{\frac{1}{\sqrt{2}}(1,-1)}_{\vec{q}_1}, \underbrace{\frac{1}{\sqrt{2}}(1,1)}_{\vec{q}_2} \Big\}$$

és una base de vectors propis ortonormal. Llavors,

$$\mathsf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

A més,

$$
\begin{aligned}
\mathsf{A} &= 1\vec{q}_1\vec{q}_1^{\,t} + 3\vec{q}_2\vec{q}_2^{\,t} \\
&= 1\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} + 3\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \\
&= 1\begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix} + 3\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \quad \square
\end{aligned}
$$

Per tal de descriure el procés de diagonalització ortogonal, necessitem una última propietat.

PROPIETAT 24.3.

En una matriu simètrica real, els subespais propis corresponents a valors propis distints són ortogonals entre ells.

Demostració: Si $\vec{p}_1$ i $\vec{p}_2$ són dos vectors propis corresponents als valors propis λ_1 i λ_2, llavors

$$
\begin{aligned}
\mathsf{A}\vec{p}_1 \cdot \vec{p}_2 &= \vec{p}_1 \cdot \mathsf{A}\vec{p}_2 \\
\lambda_1\vec{p}_1 \cdot \vec{p}_2 &= \vec{p}_1 \cdot \lambda_2\vec{p}_2 \\
\lambda_1\vec{p}_1 \cdot \vec{p}_2 &= \lambda_2\vec{p}_1 \cdot \vec{p}_2 \\
(\lambda_1 - \lambda_2)\vec{p}_1 \cdot \vec{p}_2 &= 0
\end{aligned}
$$

i, per tant, $\vec{p}_1 \cdot \vec{p}_2 = 0$. $\square$

Diagonalització ortogonal d'una matriu simètrica real A

(a) Calculeu els valors propis $\lambda_1, \lambda_2, \ldots, \lambda_r$

(b) Calculeu els subespais propis $E_{\lambda_1}(\mathsf{A})$, $E_{\lambda_2}(\mathsf{A})$, …, $E_{\lambda_r}(\mathsf{A})$.

(c) Calculeu bases ortonormals B_1, B_2, …, B_r dels subespais propis.

(d) $B = B_1 \cup B_2 \cup \cdots \cup B_r$ és una base ortonormal de $\mathbb{R}^n$ formada per vectors propis de A.

EXEMPLE 24.3.

Diagonalitzeu la matriu $\mathsf{A} = \begin{bmatrix} 0 & -2 & 2 \\ -2 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}$

(a) L'equació característica és

$$
\begin{vmatrix} -\lambda & -2 & 2 \\ -2 & 3-\lambda & 1 \\ 2 & 1 & 3-\lambda \end{vmatrix} = (\lambda+2)(\lambda-4)^2 = 0
$$

així que els valors propis són $\lambda_1 = -2$ i $\lambda_2 = 4$, amb multiplicitats algèbriques 1 i 2.

(b) Els subespais propis són

$$
E_{-2}(\mathsf{A}) = \mathrm{Nul}\begin{bmatrix} 2 & -2 & 2 \\ -2 & 5 & 1 \\ 2 & 1 & 5 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \langle(-2,-1,1)\rangle
$$

$$
E_{4}(\mathsf{A}) = \mathrm{Nul}\begin{bmatrix} -4 & -2 & 2 \\ -2 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} = \mathrm{Nul}\begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \langle(1,0,2),(-1,2,0)\rangle
$$

Llavors,

$$
\mathsf{A} = \begin{bmatrix} -2 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}\begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}^{-1}
$$

Però, com que la matriu A és simètrica, podem trobar una base ortonormal,

(c) Calculem bases ortonormals dels subespais propis:

$$B_1 = \left\{ \frac{1}{\sqrt{6}}(-2,-1,1) \right\}, \quad B_2 = \left\{ \frac{1}{\sqrt{5}}(1,0,2), \frac{1}{\sqrt{30}}(-2,5,1) \right\}$$

(d) Per tant,

$$\mathsf{A} = \begin{bmatrix} -2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30} \\ -1/\sqrt{6} & 0 & 5/\sqrt{30} \\ 1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30} \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} -2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30} \\ -1/\sqrt{6} & 0 & 5/\sqrt{30} \\ 1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30} \end{bmatrix}^t$$

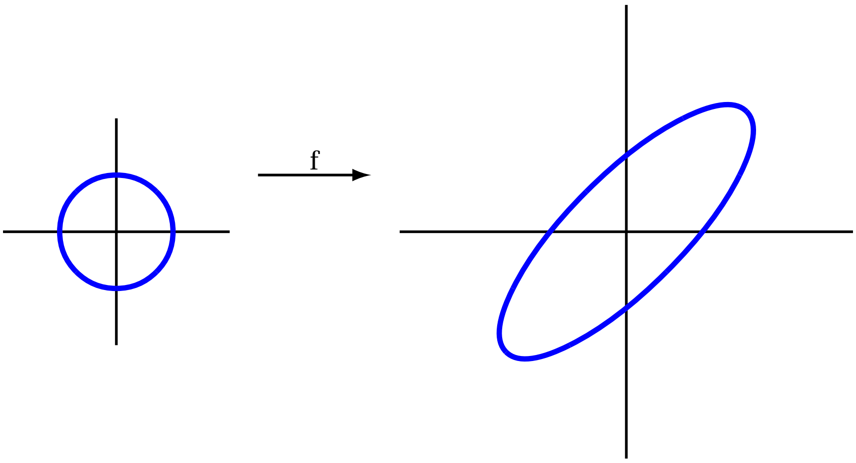
24.3. Endomorfismes simètrics

Un endomorfisme de $\mathbb{R}^n$ és *simètric* si la matriu canònica d'aquest endomorfisme és simètrica; com que acabem de veure que les matrius simètriques són diagonalitzables ortogonalment, els endomorfismes simètrics també ho són. En aquest apartat veurem alguns exemples que mostren el significat geomètric d'aquest fet.

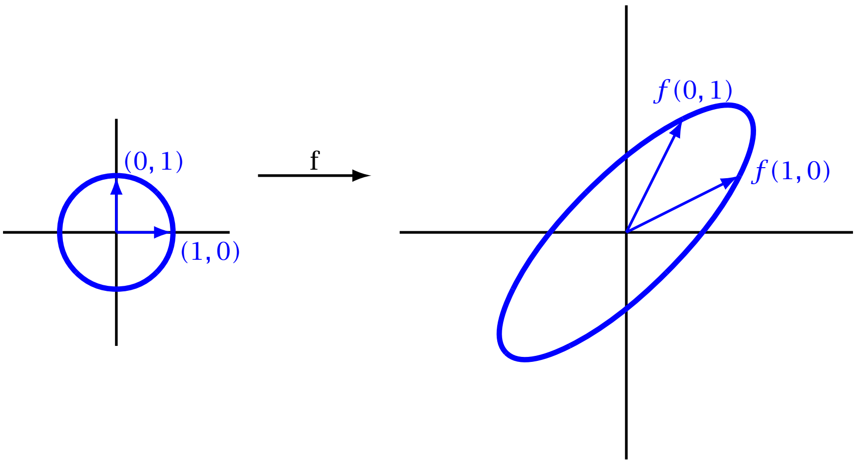
Exemple 24.4.

Diagonalització de l'endomorfisme simètric $f(\vec{x}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \vec{x}$.

Aquest endomorfisme transforma les circumferències de centre l'origen en el·lipses.



això vol dir que cada vector de posició de la circumferència es transforma en un vector de posició de l'el·lipse. Per exemple, els vectors de la base canònica es converteixen en $f(1,0) = (2,1)$ i $f(0,1) = (1,2)$,



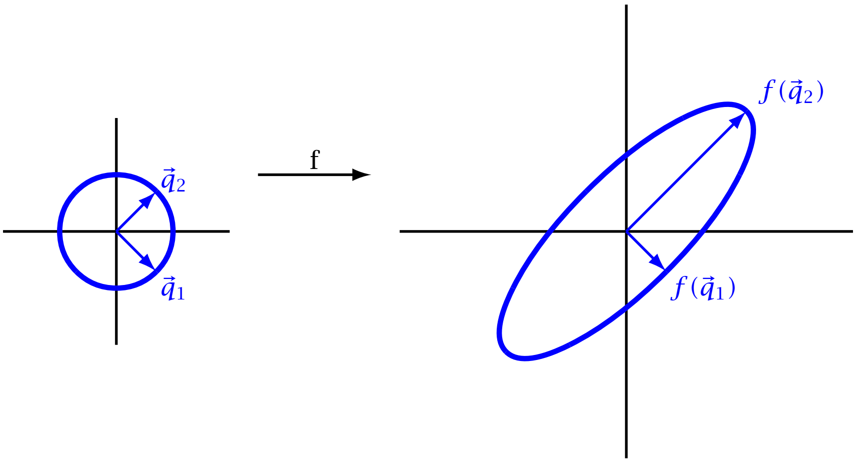
la qual cosa significa que aquests dos vectors canvien de longitud i també de direcció. Vegem el que passa amb els vectors propis.

La matriu $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ l'hem diagonalitzada en l'exemple 24.2., on hem trobat els valors propis $\lambda_1 = 1$ i $\lambda_2 = 3$ i una base ortonormal de vectors propis, $B = \{ \underbrace{\tfrac{1}{\sqrt{2}}(1,-1)}_{\vec{q}_1}, \underbrace{\tfrac{1}{\sqrt{2}}(1,1)}_{\vec{q}_2} \}$.

En conseqüència,

$$f(\vec{q}_1) = \vec{q}_1, \quad f(\vec{q}_2) = 3\vec{q}_2$$

de manera que quan l'aplicació f transforma els vectors $\vec{q}_1$ i $\vec{q}_2$ el segon canvia de longitud, però cap dels dos no canvia de direcció.



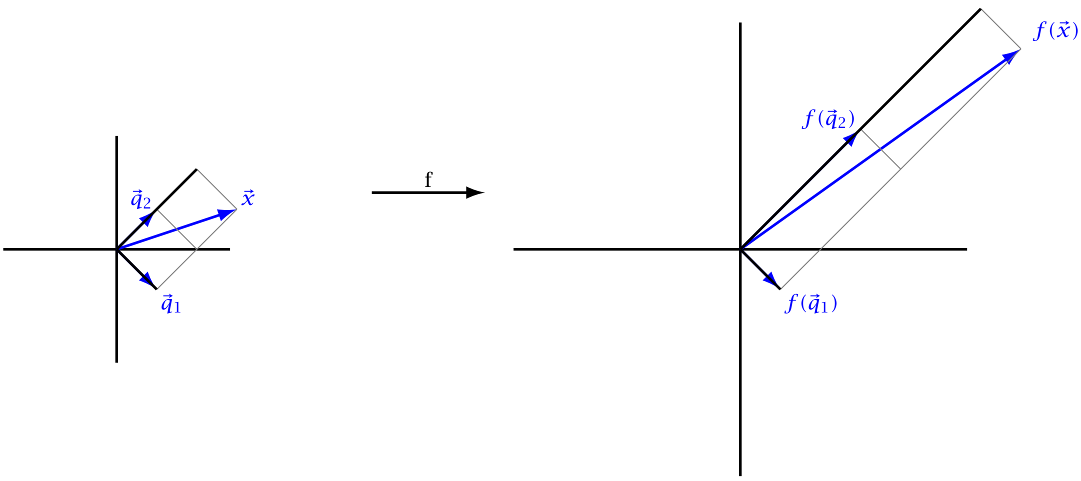
Com que $f(1,1) = 3(1,1)$, la longitud del vector $\vec{q}_2$ es multiplica per 3; en canvi, la grandària de $\vec{q}_1$ no canvia, perquè el valor propi corresponent és 1.

☞ Els valors propis mesuren la dilatació dels vectors propis.

Les rectes generades pels dos vectors propis s'anomenen *direccions principals* de f i són els eixos de simetria de l'el·lipse. Els valors propis 1 i 3 són els semieixos de l'el·lipse (les distàncies màxima i mínima de l'el·lipse al seu centre).

☞ El vectors propis $\vec{q}_1$ i $\vec{q}_2$ són, respectivament, el vector unitari que més es dilata en aplicar-li f i el que es dilata menys.

Si fem servir aquesta base, llavors la imatge del vector $\vec{x} = x_1\vec{q}_1 + x_2\vec{q}_2$ és $f(\vec{x}) = x_1\vec{q}_1 + 3x_2\vec{q}_2$, ço és, $f(\vec{x})_B = (x_1, 3x_2)$, és a dir, que cada coordenada es multiplica pel valor propi corresponent.

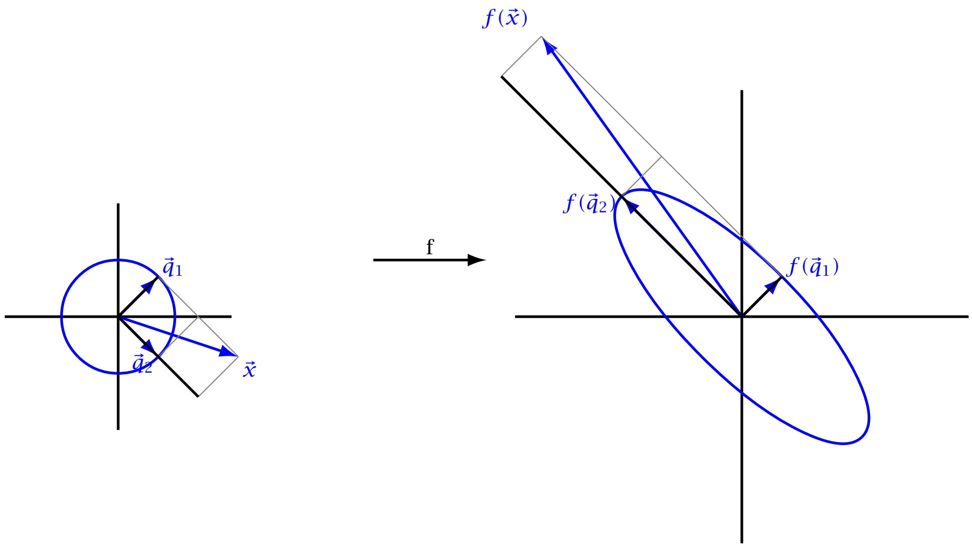


Exemple 24.5.

Diagonalització de l'endomorfisme simètric $f(\vec{x}) = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \vec{x}$.

En aquest cas, els valors propis són $\lambda_1 = 1$ i $\lambda_2 = -3$ amb vectors propis unitaris associats $\vec{q}_1 = \frac{1}{\sqrt{2}}(1,1)$ i $\vec{q}_1 = \frac{1}{\sqrt{2}}(1,-1)$, i la imatge del vector $\vec{x} = x_1\vec{q}_1 + x_2\vec{q}_2$ és

$$f(\vec{x}) = x_1\vec{q}_1 - 3x_2\vec{q}_2$$



Noteu que el valor propi negatiu $\lambda_2 = -3$ fa que el vector $f(\vec{q}_2)$, tot i conservar la direcció, canvia de sentit.

24.4. Exercicis

Exercici 24.1. Diagonalitzeu ortogonalment les matrius simètriques següents

(a) $\mathsf{A} = \begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}$ (b) $\mathsf{B} = \begin{bmatrix} 0 & 0 & 4 \\ 0 & -2 & 0 \\ 4 & 0 & 0 \end{bmatrix}$ (c) $\mathsf{C} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (d) $\mathsf{E} = \begin{bmatrix} 3 & 0 & -1 & 0 \\ 0 & 3 & 0 & -1 \\ -1 & 0 & 3 & 0 \\ 0 & -1 & 0 & 3 \end{bmatrix}$

Exercici 24.2. Diagonalitzeu ortogonalment la matriu simètrica

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Exercici 24.3. Expresseu la matriu $\begin{bmatrix} 6 & -2 & -2 \\ -2 & 7 & 0 \\ -2 & 0 & 5 \end{bmatrix}$ com una combinació lineal de matrius de rang un, $A = \lambda_1\vec{q}_1\vec{q}_1^{\,t} + \lambda_2\vec{q}_2\vec{q}_2^{\,t} + \lambda_3\vec{q}_3\vec{q}_3^{\,t}$, on $\lambda_1, \lambda_3, \lambda_3$ són els valors propis i $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ és una base ortonormal de vectors propis.

Exercici 24.4. Proveu que les úniques matrius reals que són diagonalitzables ortogonalment són les simètriques.

Exercici 24.5. Proveu que si A és una matriu real qualsevol, llavors tots els valors propis de la matriu $\mathsf{A}^t\mathsf{A}$ són reals no negatius.

25. Aplicacions de la diagonalització

Vam començar el curs presentant alguns problemes que es resolen mitjançant tècniques de l'àlgebra lineal. Alguns d'aquells problemes (i molts altres) es poden resoldre fàcilment si s'hi pot aplicar la diagonalització de matrius.

25.1. Potències d'una matriu

El càlcul de la potència A^n, si el nombre n no és molt petit, requereix moltes operacions (tinguem en compte que només per multiplicar dues matrius 3×3 cal fer 45 operacions, entre sumes i productes. Ara bé, si la matriu A és diagonalitzable, llavors tindrem

$$\begin{aligned} \mathsf{A} &= \mathsf{PDP}^{-1} \\ \mathsf{A}^2 &= \mathsf{PDP}^{-1}\mathsf{PDP}^{-1} = \mathsf{PD}^2\mathsf{P}^{-1} \\ \mathsf{A}^3 &= \mathsf{PD}^2\mathsf{P}^{-1}\mathsf{PDP}^{-1} = \mathsf{PD}^3\mathsf{P}^{-1} \end{aligned}$$

i, en general,

$$\mathsf{A}^n = \mathsf{PD}^{n-1}\mathsf{P}^{-1}\mathsf{PDP}^{-1} = \mathsf{PD}^n\mathsf{P}^{-1}$$

Com que el càlcul de D^n només suposa calcular les potències dels elements diagonals, aquesta fórmula significa que trobar A^n es redueix, esencialment, a multiplicar tres matrius.

Exemple 25.1.

Calculeu A^n, si $\mathsf{A} = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$

Aquesta matriu l'hem diagonalitzada en un exemple anterior:

$$\underbrace{\begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}_{\mathsf{P}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}}_{\mathsf{D}} \underbrace{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}^{-1}}_{\mathsf{P}^{-1}}$$

Per tant,

$$\begin{aligned} \mathsf{A}^n = \mathsf{PD}^n\mathsf{P}^{-1} &= \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}^n \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & (-3)^n \end{bmatrix} \frac{1}{5} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \\ &= \frac{1}{5}\left(2^n \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} + (-3)^n \begin{bmatrix} -2 \\ 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \end{bmatrix}\right) \\ &= \frac{2^n}{5} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} + \frac{(-3)^n}{5} \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \quad \square \end{aligned}$$

25.2. Equacions en diferències (o recurrències) lineals

Un dels exemples que presentavem a la primera unitat és aquest:

Exemple 25.2.

La successió $\{a_n\}$ es defineix de manera recurrent com

$$\begin{aligned} a_1 &= 0 \\ a_2 &= 1 \\ a_n - 3a_{n-1} + 2a_{n-2} &= 0 \end{aligned}$$

Trobeu una expressió explícita (no recurrent) del terme general.

Aquest tipus de problemes es coneixen con *equacions en diferències*. La nostra equació en concret és d'ordre 2, perquè un terme qualsevol a_n es pot calcular si ja es coneixen els dos termes anteriors ($a_n = -2a_{n-2} + 3a_{n-1}$). Aquesta equació és equivalent a una altra recurrència, d'ordre 1 però vectorial: si definim la successió de vectors $\vec{a}_n = (a_n, a_{n+1})$ tindrem

$$\vec{a}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \vec{a}_n = \begin{bmatrix} a_n \\ a_{n+1} \end{bmatrix} = \begin{bmatrix} a_n \\ -2a_{n-1} + 3a_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_n \end{bmatrix} = \mathsf{A}\vec{a}_{n-1}$$

La solució d'aquesta equació és, òbviament,

$$\vec{a}_n = \mathsf{A}^{n-1}\vec{a}_1$$

I si la matriu A és diagonalitzable serà molt fàcil calcular el terme general. L'equació característica és

$$\det \begin{bmatrix} -\lambda & 1 \\ -2 & 3-\lambda \end{bmatrix} = \lambda^2 - 3\lambda + 2 = 0$$

els valors propis són $\lambda_1 = 1$ i $\lambda_2 = 2$ i els espais propis

$$E_1(\mathsf{A}) = \operatorname{Nul} \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix} = \langle (1,1) \rangle$$
$$E_2(\mathsf{A}) = \operatorname{Nul} \begin{bmatrix} -2 & 1 \\ -2 & 1 \end{bmatrix} = \langle (1,2) \rangle$$

de manera que, si

$$\mathsf{P} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathsf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

llavors

$$\mathsf{A} = \mathsf{PDP}^{-1}$$
$$\mathsf{P}^{-1}\mathsf{A} = \mathsf{DP}^{-1}$$

i l'equació $a_n = \mathsf{A}\vec{a}_{n-1}$ és equivalent a $\mathsf{P}^{-1}\vec{a}_n = \mathsf{P}^{-1}\mathsf{A}\vec{a}_{n-1} = \mathsf{DP}^{-1}\vec{a}_{n-1}$.

Si fem el canvi $\vec{b}_n = \mathsf{P}^{-1}\vec{a}_n$ (o $\vec{a}_n = \mathsf{P}\vec{b}_n$) ens quedarà l'equació equivalent

$$\vec{b}_n = \mathsf{D}\vec{b}_{n-1}$$

la solució de la qual és

$$\vec{b}_n = \mathsf{D}^{n-1}\vec{b}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2^{n-1} \end{bmatrix} \vec{b}_1$$

i l'únic que ens resta és calcular el vector $\vec{b}_1$. Això ho podem fer sense necessitat de calcular la inversa de la matriu P: el vector $\vec{b}_1$ és la solució del sistema lineal $\mathsf{P}\vec{b}_1 = \vec{a}_1$, així que resolem aquest sistema i obtenim $\vec{b}_1 = (-1, 1)$,

$$\vec{b}_n = \mathsf{D}^{n-1}\vec{b}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2^{n-1} \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2^{n-1} \end{bmatrix}$$

i, finalment,

$$\vec{a}_n = \mathsf{P}\vec{b}_n = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2^{n-1} \end{bmatrix} = \begin{bmatrix} -1 + 2^{n-1} \\ -1 + 2^n \end{bmatrix}$$

de manera que el terme general de la successió original és $a_n = -1 + 2^{n-1}$. □

25.2.1. Els nombres de Fibonacci

La successió de Fibonacci es defineix mitjançant la recurrència

$$\begin{aligned} f_1 &= 1 \\ f_2 &= 1 \\ f_n &= f_{n-2} + f_{n-1} \end{aligned}$$

així que podem trobar-ne el terme general seguint un procés semblant al de l'exemple anterior. Definint la successió de vectors $\vec{f}_n = (f_n, f_{n+1})$ tindrem

$$\vec{f}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \vec{f}_n = \begin{bmatrix} f_n \\ f_{n+1} \end{bmatrix} = \begin{bmatrix} f_n \\ f_{n-1} + f_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_{n-1} \\ f_n \end{bmatrix} = \mathsf{A}\vec{f}_{n-1}$$

i, llavors, $\vec{f}_n = \mathsf{A}^{n-1}\vec{f}_1$. Diagonalitzem la matriu A: l'equació característica és

$$\det \begin{bmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} = \lambda^2 - \lambda - 1 = 0$$

de manera que els valors propis són $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$ i $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$ i els espais propis

$$\begin{aligned} E_{\lambda_1}(\mathsf{A}) &= \operatorname{Nul} \begin{bmatrix} -\lambda_1 & 1 \\ 1 & 1-\lambda_1 \end{bmatrix} = \langle (1, \lambda_1) \rangle \\ E_{\lambda_2}(\mathsf{A}) &= \operatorname{Nul} \begin{bmatrix} -\lambda_2 & 1 \\ 1 & 1-\lambda_2 \end{bmatrix} = \langle (1, \lambda_2) \rangle \end{aligned}$$

així que

$$\mathsf{A} = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}^{-1}$$

i

$$\begin{aligned} \mathsf{A}^{n-1} &= \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} \begin{bmatrix} \lambda_1^{n-1} & 0 \\ 0 & \lambda_2^{n-1} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}^{-1} \\ \mathsf{A}^{n-1} &= \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_1^{n-1}\lambda_2 - \lambda_1\lambda_2^{n-1} & -\lambda_1^{n-1} + \lambda_2^{n-1} \\ \lambda_1^{n}\lambda_2 + \lambda_1\lambda_2^{n} & -\lambda_1^{n} + \lambda_2^{n} \end{bmatrix} \end{aligned} \tag{25.1}$$

D'altra banda, calculant directament algunes potències de A,

$$\begin{aligned} \mathsf{A}^2 &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \\ f_2 & f_3 \end{bmatrix} \\ \mathsf{A}^3 &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 \\ f_2 & f_3 \end{bmatrix} = \begin{bmatrix} f_2 & f_3 \\ f_1 + f_2 & f_2 + f_3 \end{bmatrix} = \begin{bmatrix} f_2 & f_3 \\ f_3 & f_4 \end{bmatrix} \end{aligned}$$

s'observa que

$$\mathsf{A}^{n-1} = \begin{bmatrix} f_{n-2} & f_{n-1} \\ f_{n-1} & f_n \end{bmatrix}$$

Comparant aquesta igualtat amb (25.1) obtenim que $f_n = \frac{1}{\lambda_2 - \lambda_1}(-\lambda_1^n + \lambda_2^n)$, és a dir,

$$\boxed{f_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n} \quad \square$$

25.3. Equacions diferencials lineals

L'equació diferencial lineal homogènia de primer ordre i coeficient constant és l'equació

$$y' = ay$$

on $y : \mathbb{R} \longrightarrow \mathbb{R}$ és una funció desconeguda de la variable real t i a és una constant. Les funcions de la forma $y = Ce^{at}$, on C és una constant, són solucions d'aquesta equació, perquè

$$y = Ce^{at} \Longrightarrow y' = Cae^{at} = ay$$

El que no és tan evident, però també es pot provar, és que *totes* les solucions tenen aquesta forma, és a dir, que

La solució general de l'equació $y' = ay$ és $y = Ce^{at}$.

Ací ens interessen els sistemes d'equacions diferencials lineals (amb coeficients constants) de la forma

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ y_2' &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ &\dots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n \end{aligned}$$

o, en forma matricial,

$$\vec{y}' = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \vec{y}$$

$$\vec{y}' = \mathsf{A}\vec{y}$$

Si la matriu A és diagonalitzable, podrem escriure aquesta equació com

$$\vec{y}' = \mathsf{PDP}^{-1}\vec{y}$$
$$\mathsf{P}^{-1}\vec{y}' = \mathsf{DP}^{-1}\vec{y}$$

així que podem fer el canvi de funció $\vec{z} = \mathsf{P}^{-1}\vec{y}$ (o $\vec{y} = \mathsf{P}\vec{z}$) i obtindrem el sistema equivalent

$$\vec{z}' = \mathsf{D}\vec{z}$$

L'avantatge d'aquest sistema és que *és desacoplat*, en el sentit que cada equació conté únicament una funció incògnita:

$$\vec{z}' = \mathsf{D}\vec{z}$$

$$\vec{z}' = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \vec{z}$$

és a dir,

$$z_1' = \lambda_1 z_1, \qquad z_2' = \lambda_2 z_2, \qquad \dots, \quad z_n' = \lambda_n z_n$$

Les solucions generals de cadascuna d'aquestes equacions són, segons acabem de veure,

$$z_1 = C_1 e^{\lambda_1 t}, \quad z_2 = C_2 e^{\lambda_2 t}, \quad \dots, \quad z_n = C_n e^{\lambda_n t}$$

Si ara desfem el canvi $\vec{y} = \mathsf{P}\vec{z}$ obtindrem la solució general del sistema original $\vec{y}' = \mathsf{A}\vec{y}$:

$$\begin{aligned} \vec{y} &= \mathsf{P}\vec{z} \\ &= \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} \begin{bmatrix} C_1 e^{\lambda_1 t} \\ C_2 e^{\lambda_2 t} \\ \dots \\ C_n e^{\lambda_n t} \end{bmatrix} \end{aligned}$$

ço és,

$$\boxed{\vec{y} = C_1 e^{\lambda_1 t}\vec{p}_1 + C_2 e^{\lambda_2 t}\vec{p}_2 + \cdots + C_n e^{\lambda_n t}\vec{p}_n}$$

Exemple 25.3.

Trobeu la solució general del sistema d'equacions diferencials lineals

$$\begin{aligned} y_1' &= -5y_1 \qquad - 6y_3 \\ y_2' &= 3y_1 + y_2 + 3y_3 \\ y_3' &= 3y_1 \qquad + 4y_3 \end{aligned}$$

La matriu $A = \begin{bmatrix} -5 & 0 & -6 \\ 3 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}$ l'hem diagonalitzada a l'exemple 23.3. (pàgina 193),

$$\underbrace{\begin{bmatrix} -5 & 0 & -6 \\ 3 & 1 & 3 \\ 3 & 0 & 4 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1}}_{P^{-1}}$$

Per tant, la solució general del nostre sistema és

$$\vec{y} = C_1 e^{t}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + C_2 e^{t}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + C_3 e^{-2t}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

o

$$\begin{aligned} y_1 &= \qquad - C_2 e^{t} - 2C_3 e^{-2t} \\ y_2 &= C_1 e^{t} \qquad + \ C_3 e^{-2t} \\ y_3 &= \qquad + C_2 e^{t} + \ C_3 e^{-2t} \quad \square \end{aligned}$$

25.4. Formes quadràtiques

Una *forma quadràtica* és una aplicació $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ en la qual la imatge $f(x_1, x_2, \ldots, x_n)$ és un polinomi homogeni[2] de grau dos, per exemple,

$$\begin{aligned} f(x_1, x_2) &= x_1^2 - x_2^2 \\ g(x_1, x_2, x_3) &= 2x_1^2 - x_2^2 - 6x_3^2 - 10x_1x_2 + 6x_2x_3 \end{aligned}$$

Fins ara, només ens hem ocupat de les aplicacions lineals, perquè es poden calcular, fent servir matrius, com $f(\vec{x}) = A\vec{x}$. Les formes quadràtiques també podem expressar-les matricialment. Per exemple, l'aplicació f,

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - x_2^2$$

Aquesta representació matricial no és única; per exemple, en el cas de l'aplicació g, podem triar

$$g(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & -10 & 0 \\ 0 & -1 & 6 \\ 0 & 0 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1^2 - x_2^2 - 6x_3^2 - 10x_1x_2 + 6x_2x_3$$

o, també,

$$g(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 2 & -5 & 0 \\ -5 & -1 & 3 \\ 0 & 3 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 2x_1^2 - x_2^2 - 6x_3^2 - 10x_1x_2 + 6x_2x_3$$

[2]Un polinomi és homogeni si tots els monomis que en formen part són del mateix grau.

Això és degut a que

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$
$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + (a_{23} + a_{23})x_2x_3$$

de manera que, en el cas de l'aplicació g, podem elegir lliurement els elements no diagonals sempre que ens assegurem que $a_{12} + a_{21} = -10$, $a_{13} + a_{31} = 0$ i $a_{23} + a_{32} = 6$. Com que el que ens va a interessar és la diagonalització de la matriu, els triarem de manera que ens assegurem que la matriu resulte simètrica.

DEFINICIÓ 25.1.

Una forma quadràtica és una aplicació $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de la forma $f(\vec{x}) = \vec{x}^t \mathsf{A} \vec{x}$ on A és una matriu simètrica.

Com que la matriu A és simètrica, podem diagonalitzar-la ortogonalment, així que, si $\lambda_1, \lambda_2, \dots, \lambda_n$ són els valors propis, hi ha una base ortonormal de vectors propis $B = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$ de manera que

$$\mathsf{A} = \mathsf{PDP}^t = \begin{bmatrix} \vec{p}_1 & \vec{p}_2 & \dots & \vec{p}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{p}_1^t \\ \vec{p}_2^t \\ \dots \\ \vec{p}_n^t \end{bmatrix}$$

Si fem el canvi de base $\vec{x} = \mathsf{P}\vec{y}$ llavors,

$$f(\vec{x}) = \vec{x}^t \mathsf{A} \vec{x} = \vec{x}^t \mathsf{P}^t \mathsf{PDP}^t \mathsf{P} \vec{y} = \vec{y}^t \mathsf{D} \vec{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$
$$f(\vec{x}) = \vec{x}_B^t \mathsf{D} \vec{x}_B$$

Hem provat que qualsevol forma quadràtica es pot expressar sense termes mixtes (del tipus x_2x_3). Aquesta és la forma reduïda de la forma quadràtica.

EXEMPLE 25.4.

Trobeu la forma reduïda de la forma quadràtica

$$g(\vec{x}) = \vec{x}^t \begin{bmatrix} 2 & -5 & 0 \\ -5 & -1 & 3 \\ 0 & 3 & -6 \end{bmatrix} \vec{x} = 2x_1^2 - x_2^2 - 6x_3^2 - 10x_1x_2 + 6x_2x_3$$

Els valors propis de la matriu $\left[\begin{smallmatrix} 2 & -5 & 0 \\ -5 & -1 & 3 \\ 0 & 3 & -6 \end{smallmatrix}\right]$ són $\lambda_1 = -3$, $\lambda_2 = 6$ i $\lambda_3 = -8$ i, els subespais propis associats,

$$E_{-3} = \text{Nul}\,(1, 1, 1)$$
$$E_6 = \text{Nul}\,(-5, 4, 1)$$
$$E_{-8} = \text{Nul}\,(-1/3, -2/3, 1)$$

Així que

$$B = \left\{ \frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{42}}(-5, 4, 1), \frac{1}{\sqrt{14}}(-1/3, -2/3, 1) \right\}$$

és una base ortonormal de vectors propis. L'expressió de la forma quadràtica referida a aquesta base és

$$f(\vec{x}) = \vec{y}^t \begin{bmatrix} -3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & -8 \end{bmatrix} \vec{y} = -3y_1^2 + 6y_2^2 - 8y_3^3 \quad \square$$

25.5. Exercicis

Exercici 25.1. Resoleu el sistema d'equacions diferencials lineals

$$\begin{aligned} y_1' &= y_1 + y_2 + y_3 \\ y_2' &= y_1 + y_2 + y_3 \\ y_3' &= y_1 + y_2 + y_3 \end{aligned}$$

i trobeu la solució particular que compleix les condicions $y_1(0) = -1, y_2(0) = 2, y_3(0) = 1$.

Exercici 25.2. Resoleu l'equació (matricial) en diferències $\vec{x}_{n+1} = \mathsf{A}\vec{x}_n$, on $\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Exercici 25.3. Una forma quadràtica $f(\vec{x}) = \vec{x}^t\mathsf{A}\vec{x}$ és *definida positiva* si $f(\vec{x}) > 0$ per a tots els vectors no nuls. Proveu que això és equivalent a que tots els valors propis de la matriu A són positius. Si $f(\vec{x}) \geq 0$ per a tots els vectors, llavors la forma quadràtica és *semidefinida positiva*; i si la funció pren algun valor positiu i algun de negatiu, llavors es diu que la forma quadràtica és *indefinida.* Quan la forma quadràtica f és definida positiva, semidefinida positiva o indefinida, també diem que la matriu simètrica A ho és.

(a) Caracteritzeu les formes semidefinides positives i les indefinides segons els signes dels valors propis.
(b) Estudieu si cadascuna de les matrius següents són definides positives, semidefinides positives o indefinides.

$$\mathsf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \mathsf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \mathsf{C} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \quad \mathsf{D} = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \quad \mathsf{E} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad \mathsf{F} = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

26. Vectors singulars i valors singulars

Anem acabant el curs amb una última factorització de matrius, la descomposició en valors singulars. Veurem que qualsevol matriu $m \times n$, A, es pot escriure en la forma $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$, on U i V són matrius ortogonals i les entrades no diagonals de la matriu Σ són totes nul·les.[a] Això vol dir que podem *diagonalitzar* qualsevol matriu (o transformació lineal), però elegint bases diferents en els espais vectorials inicial i final.

[a] No podem dir que Σ és diagonal, perquè aquesta matriu només és quadrada quan ho és A.

26.1. Descomposició en valors singulars

Diagonalitzar la matriu quadrada A (o l'endomorfisme de $\mathbb{R}^n$ $f(\vec{x}) = \mathsf{A}\vec{x}$) vol dir trobar una base $B = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$ de vectors propis, és a dir, trobar n vectors linealment independents de manera que

$$\mathsf{A}\vec{p}_1 = \lambda_1\vec{p}_1,\ \mathsf{A}\vec{p}_2 = \lambda_2\vec{p}_2,\ \dots,\ \mathsf{A}\vec{p}_n = \lambda_n\vec{p}_n$$

O, en termes matricials,

$$\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^{-1}$$

on D és una matriu diagonal.

Com hem vist en les unitats anteriors, això no és possible sempre, però si en el cas particular de les matrius simètriques. A més, en aquest cas, la base B podem elegir-la ortonormal. Això vol dir que, si A és simètrica,

$$\mathsf{A} = \mathsf{Q}\mathsf{D}\mathsf{Q}^t$$

on D és una matriu diagonal i Q és ortogonal.

Si admetem la possibilitat d'elegir bases diferents als espais inicial i final, llavors resulta que qualsevol transformació lineal es pot *diagonalitzar* ortogonalment, sense cap restricció sobre la matriu A. El problema, en el cas més general, es planteja d'aquesta manera:

El problema de la descomposició en valors singulars

Suposem que A és una matriu real $m \times n$. Factoritzeu aquesta matriu com $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$ on

(a) U és una matriu ortogonal $m \times m$,

(b) V és una matriu ortogonal $n \times n$,

(c) Σ és una matriu $m \times n$ que té aquesta forma:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{bmatrix}$$

on Σ_{11} és una matriu diagonal que no conté zeros a la diagonal.

Per exemple,

$$\underbrace{\begin{bmatrix} -3 & 6 & 0 \\ -5 & 2 & -4 \\ 5 & -2 & 4 \\ 3 & -6 & 0 \end{bmatrix}}_{\mathsf{A}} = \underbrace{\left[\begin{array}{rr|rr} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \end{array}\right]}_{\mathsf{U}} \underbrace{\left[\begin{array}{cc|c} 12 & 0 & 0 \\ 0 & 6 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]}_{\Sigma} \underbrace{\left[\begin{array}{rrr} -2/3 & 2/3 & -1/3 \\ 1/3 & 2/3 & 2/3 \\ \hline -2/3 & -1/3 & 2/3 \end{array}\right]}_{\mathsf{V}^t}$$

El nostre objectiu és provar que qualsevol matriu es pot factoritzar d'aquesta manera. El primer que cal observar és el fet que, com que les matrius U i V són invertibles, el rang de A és el mateix que el de Σ_{11}: si $\operatorname{rang}\mathsf{A} = r$, llavors Σ és una matriu $r \times r$. A més a més, si $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$ llavors, $\mathsf{AV} = \mathsf{U}\Sigma$, així que $\mathsf{A}\vec{v}_1 = \sigma_1\vec{u}_1$, $\mathsf{A}\vec{v}_2 = \sigma_2\vec{u}_2$, ..., $\mathsf{A}\vec{v}_r = \sigma_r\vec{u}_r$ (on $\sigma_1, \sigma_2, \dots$ són les entrades diagonals de Σ_{11}) i $\mathsf{A}\vec{v}_{r+1} = \cdots = \mathsf{A}\vec{v}_n = 0$.

26.2. Construcció de la descomposició en valors singulars

Per a obtenir una descomposició en valors singulars de la matriu A hem de trobar dues bases ortonormals, $B_1 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$, base de $\mathbb{R}^n$, i $B_2 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$, base de $\mathbb{R}^m$, de manera que

$$\mathsf{A}\vec{v}_1 = \sigma_1\vec{u}_1,\ \mathsf{A}\vec{v}_2 = \sigma_2\vec{u}_2,\ \dots,\ \mathsf{A}\vec{v}_r = \sigma_r\vec{u}_r \tag{26.1}$$

$$\mathsf{A}\vec{v}_{r+1} = \vec{0},\ \mathsf{A}\vec{v}_{r+2} = \vec{0},\ \dots,\ \mathsf{A}\vec{v}_n = \vec{0} \tag{26.2}$$

Perquè passe això, cal que $\{\vec{v}_{r+1}, \vec{v}_{r+2}, \dots, \vec{v}_n\}$ siga una base de l'espai nul de A i, atés que $\operatorname{Nul}\mathsf{A}^\perp = \operatorname{Fil}\mathsf{A}$, que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ siga una base de l'espai fila de A. Així, doncs, la base B_1 ha de ser la unió de

(a) una base ortonormal de Fil A, $B_{\operatorname{Fil}\mathsf{A}} = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$ i

(b) una base ortonormal de Nul A, $B_{\operatorname{Nul}\mathsf{A}} = \{\vec{v}_{r+1}, \vec{v}_{r+2}, \dots, \vec{v}_n\}$.

Ara bé, la base $B_{\operatorname{Fil}\mathsf{A}}$ ha de complir una condició addicional: el conjunt dels vectors imatge, $\{\mathsf{A}\vec{v}_1, \mathsf{A}\vec{v}_2, \dots, \mathsf{A}\vec{v}_r\}$ ha de ser ortogonal (si aconseguim això, després triarem els nombres σ_i perquè els vectors $\vec{u}_i$ siguen unitaris). Per tant, el problema important és aquest:

> Elegiu una base ortonormal de l'espai fila, $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r\}$, de manera que el conjunt $\{\mathsf{A}\vec{v}_1, \mathsf{A}\vec{v}_2, \dots, \mathsf{A}\vec{v}_r\}$ siga ortogonal.

Necessitem, doncs, que $\mathsf{A}\vec{v}_i \cdot \mathsf{A}\vec{v}_j = 0$ (si $i \neq j$), és a dir, que $\vec{v}_i^t\mathsf{A}^t\mathsf{A}\vec{v}_j = 0$. O que la matriu

$$\begin{bmatrix} \vec{v}_1^t \\ \vec{v}_2^t \\ \dots \\ \vec{v}_r^t \end{bmatrix} \mathsf{A}^t\mathsf{A} \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_r \end{bmatrix}$$

siga diagonal; per tant, el que necessitem és diagonalitzar la matriu $\mathsf{A}^t\mathsf{A}$. Afortunadament, com que aquesta matriu és simètrica, podem diagonalitzar-la ortogonalment: existeix una matriu ortogonal V, *les columnes de la qual són vectors propis de* $\mathsf{A}^t\mathsf{A}$, i una matriu diagonal D, *la diagonal dels valors propis de* $\mathsf{A}^t\mathsf{A}$, de manera que

$$\mathsf{A}^t\mathsf{A} = \mathsf{V}\mathsf{D}\mathsf{V}^t$$

o, equivalentment,

$$\mathsf{V}^t\mathsf{A}^t\mathsf{A}\mathsf{V} = \mathsf{D}$$

A més a més, com que $\operatorname{rang}\mathsf{A}^t\mathsf{A} = \operatorname{rang}\mathsf{A} = r$, podem ordenar els valors i els vectors propis de manera que tinguem

$$\left[\begin{array}{c} \vec{v}_1^t \\ \vec{v}_2^t \\ \dots \\ \vec{v}_r^t \\ \hline \vec{v}_{r+1}^t \\ \dots \\ \vec{v}_n^t \end{array}\right] \mathsf{A}^t\mathsf{A} \left[\begin{array}{cccc|ccc} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_r & \vec{v}_{r+1} & \dots & \vec{v}_n \end{array}\right] = \left[\begin{array}{cccc|ccc} \lambda_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & \lambda_r & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right]$$

($\vec{v}_{r+1}, \dots, \vec{v}_n$ són els vectors propis associats al valor propi zero). Elegint d'aquesta manera els vectors, tindrem que el conjunt $\{\mathsf{A}\vec{v}_1, \mathsf{A}\vec{v}_2, \dots, \mathsf{A}\vec{v}_r\}$ és ortogonal, perquè

$$\mathsf{A}\vec{v}_i \cdot \mathsf{A}\vec{v}_j = \vec{v}_i^t\mathsf{A}^t\mathsf{A}\vec{v}_j = \vec{v}_i^t(\lambda_j\vec{v}_j) = 0 \quad \text{si } i \neq j$$

Finalment, per tal d'aconseguir un conjunt ortonormal, definim els vectors unitaris

$$\vec{u}_i = \frac{1}{\|\mathsf{A}\vec{v}_i\|}\mathsf{A}\vec{v}_i, \qquad (1 \leq i \leq r)$$

I, com que $\|\mathsf{A}\vec{v}_i\|^2 = \vec{v}_i^t\mathsf{A}^t\mathsf{A}\vec{v}_i = \lambda_i$, podem assegurar que $\lambda_i > 0$ i tindrem

$$\vec{u}_i = \frac{1}{\sqrt{\lambda_i}}\mathsf{A}\vec{v}_i, \qquad (1 \leq i \leq r)$$

DEFINICIÓ 26.1.

> *Els valors singulars de la matriu* A *són les arrels quadrades positives dels valors propis no nuls de* $\mathsf{A}^t\mathsf{A}$.
> *Els valors singulars es representen habitualment amb la lletra* σ ($\sigma_i = \sqrt{\lambda_i}$) *i es presenten ordenats de més gran a més menut.*

Els vectors $\vec{u}_i$ formen una base ortonormal de l'espai columna $\operatorname{Col}\mathsf{A}$ i, tal com els hem definits,

$$\vec{u}_i = \frac{1}{\sqrt{\lambda_i}}\mathsf{A}\vec{v}_i \xrightarrow{\sigma_i=\sqrt{\lambda_i}} \mathsf{A}\vec{v}_i = \sigma_i\vec{u}_i, \qquad (1 \le i \le r)$$

Per tant, construirem les bases B_1 i B_2 de la manera següent:

(a) La base B_1 és la unió de

- (a1) una base ortonormal de $\operatorname{Fil}\mathsf{A}$, $B_{\operatorname{Fil}\mathsf{A}} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$, *formada per vectors propis de* $\mathsf{A}^t\mathsf{A}$ *associats als valors propis no nuls* i
- (a2) una base ortonormal de $\operatorname{Nul}\mathsf{A}$, $B_{\operatorname{Nul}\mathsf{A}} = \{\vec{v}_{r+1}, \vec{v}_{r+2}, \ldots, \vec{v}_n\}$.

(b) La base B_2 és la unió de

- (b1) una base ortonormal de $\operatorname{Col}\mathsf{A}$, $B_{\operatorname{Col}\mathsf{A}} = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$, formada pels vectors $\vec{u}_i = \frac{1}{\sigma_i}\mathsf{A}\vec{v}_i$, i
- (b2) una base ortonormal de $\operatorname{Nul}\mathsf{A}^t$, $B_{\operatorname{Nul}\mathsf{A}^t} = \{\vec{u}_{r+1}, \vec{u}_{r+2}, \ldots, \vec{u}_m\}$.

Només ens queda construir la matriu Σ:

$$\begin{aligned}
\mathsf{AV} &= \begin{bmatrix} \mathsf{A}\vec{v}_1 & \mathsf{A}\vec{v}_2 & \ldots\mathsf{A}\vec{v}_r & \mathsf{A}\vec{v}_{r+1} & \ldots & \mathsf{A}\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1\vec{u}_1 & \sigma_2\vec{u}_2 & \ldots\sigma_r\vec{u}_r & \vec{0} & \ldots & \vec{0} \end{bmatrix} \\
&= \mathsf{U}\left[\begin{array}{cccc|ccc}
\sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & \sigma_r & 0 & \ldots & 0 \\ \hline
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\ldots & & & & & & \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{array}\right]
\end{aligned}$$

DEFINICIÓ 26.2. (DESCOMPOSICIÓ EN VALORS SINGULARS)

> *Suposem que* A *és una matriu* $m \times n$ *de rang* r. *Una descomposició en valors singulars d'aquesta matriu és una factorització* $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$ *on*
>
> *(a)* U *és una matriu ortogonal* $m \times m$,
>
> *(b)* V *és una matriu ortogonal* $n \times n$,
>
> *(c)* Σ *és una matriu* $m \times n$ *que té aquesta forma:*
>
> $$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \mathrm{O} \\ \hline \mathrm{O} & \mathrm{O} \end{array}\right] = \begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 & 0 & \ldots & 0 \\
\ldots & & & & & & \\
0 & 0 & \ldots & \sigma_r & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\
\ldots & & & & & & \\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{bmatrix}$$
>
> *A més a més,* $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$

El que hem provat en aquest apartat és que sempre es pot trobar una descomposició en valors singulars fent servir l'algorisme següent:

Algorisme de càlcul d'una descomposició en valors sigulars de la matriu A

(a) Calculeu els valors propis no nuls de la matriu $\mathsf{A}^t\mathsf{A}$. Ordeneu-los de més gran a més petit: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$.

Les arrels quadrades positives, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$, d'aquests valors propis són els valors singulars de la matriu A. Això determina la matriu diagonal Σ_{11}.

(b) Calculeu bases ortonormals dels subespais propis de $\mathsf{A}^t\mathsf{A}$ corresponents als valors propis no nuls i una base ortonormal de l'espai Nul A. Si B és la unió ordenada d'aquestes bases, la matriu V és M_B.

(c) Calculeu els vectors $\mathsf{u}_i = (1/\sigma_i)\mathsf{A}\vec{v}_i$, $1 \leq i \leq r$ i afegiu-hi una base ortonormal de l'espai Nul A^t. Això ens proporciona la matriu U.

Exemple 26.1.

Trobeu una descomposició en valors singulars de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Calculant el producte $\mathsf{A}^t\mathsf{A}$ obtenim

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix}$$

Els valors propis d'aquesta matriu són $\lambda_1 = 25$ i $\lambda_2 = 0$, així que l'únic valor singular de A és $\sigma_1 = 5$. El subespai propi de $\mathsf{A}^t\mathsf{A}$ associat al valor propi λ_1 és

$$\mathrm{Nul}(\mathsf{A}^t\mathsf{A} - 25\mathrm{I}) = \mathrm{Nul}\begin{bmatrix} -20 & 10 \\ 10 & -5 \end{bmatrix} = \langle (1,2) \rangle$$

Elegim el vector $\vec{v}_1 = \frac{1}{\sqrt{5}}(1,2)$ perquè siga unitari. El conjunt $\{\vec{v}_1\}$ és la base ortonormal de l'espai fila que necessitem. El completem amb una base ortonormal de l'espai nul de A:

$$\mathrm{Nul}\,\mathsf{A} = \mathrm{Nul}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \langle (-2,1) \rangle$$

Podem elegir el vector $\vec{v}_2 = \frac{1}{\sqrt{5}}(-2,1)$. D'aquesta manera, la matriu V serà aquesta: $\mathsf{V} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$.

Ara calcularem els vectors $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ i $\vec{u}_4$. El primer ha de ser

$$\vec{u}_1 = \frac{1}{\sigma_1}\mathsf{A}\vec{v}_1 = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

I l'hem de completar amb una base ortonormal de l'espai nul esquerre:

$$\mathrm{Nul}\,\mathsf{A}^t = \mathrm{Nul}\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \left\langle \frac{1}{\sqrt{5}}(-2,1,0,0), (0,0,1,0), (0,0,0,1) \right\rangle$$

Amb tot això ja podem armar la descomposició en valors singulars:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \left[\begin{array}{c|ccc} 1/\sqrt{5} & -2/\sqrt{5} & 0 & 0 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c|c} 5 & 0 \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{cc} 1/\sqrt{5} & 2/\sqrt{5} \\ \hline 2/\sqrt{5} & -1/\sqrt{5} \end{array}\right] \qquad \square$$

EXEMPLE 26.2.

Trobeu una descomposició en valors singulars de la matriu $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix}$.

Com que aquesta matriu és la transposada de la de l'exemple anterior, podem trobar la descomposició sense cap càlcul addicional, perquè

$$A = U\Sigma V^t \Leftrightarrow A^t = V\Sigma^t U^t$$

És a dir,

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 & 0 \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \square$$

Aquest exemple mostra que els valors singulars de A^t són els mateixos que els de A.

26.2.1. VALORS SINGULARS I VECTORS SINGULARS

Hem definit els valors singulars de la matriu A com els valors propis de la matriu A^tA. Els vectors singulars es defineixen de manera anàloga:

DEFINICIONS 26.3.

Els vectors singulars drets de la matriu A són els vectors propis associats als valors propis no nuls de A^tA.
Els vectors singulars esquerres de la matriu A són els vectors propis associats als valors propis no nuls de AA^t.

Es pot provar que, si $A = U\Sigma V^t$ és una descomposició en valors singulars, llavors, les entrades no nul·les de Σ són els valors singulars de A, les columnes corresponents de U són vectors singulars esquerres i, les de V, vectors singulars drets.

26.2.2. DESCOMPOSICIÓ EN VALORS SINGULARS REDUÏDA

En la pràctica es pot fer servir una forma reduïda de la descomposició en valors singulars, sense necessitat de calcular bases dels subespais nuls: si particionem les matrius U i V en la forma $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ i $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ on els blocs U_2 i V_2 corresponen a les bases dels dos espais nuls, tindrem que

$$\begin{aligned} A = U\Sigma V^t &= \left[U_1 \mid U_2 \right] \left[\begin{array}{c|c} \Sigma_{11} & O \\ \hline O & O \end{array} \right] \begin{bmatrix} V_1^t \\ V_2^t \end{bmatrix} \\ &= \left[U_1\Sigma_{11} + U_2O \mid U_10 + U_2O \right] \begin{bmatrix} V_1^t \\ V_2^t \end{bmatrix} = \left[U_1\Sigma_{11} \mid O \right] \begin{bmatrix} V_1^t \\ V_2^t \end{bmatrix} \\ &= U_1\Sigma_{11}V_1^t + OV_2^t \\ A &= U_1\Sigma_{11}V_1^t \end{aligned}$$

EXEMPLE 26.3.

Trobeu una descomposició en valors singulars reduïda de la matriu $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

En l'exemple 26.1. hem trobat la descomposició en valors singulars

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \left[\begin{array}{c|ccc} 1/\sqrt{5} & -2/\sqrt{5} & 0 & 0 \\ 2/\sqrt{5} & 1/\sqrt{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \left[\begin{array}{c|c} 5 & 0 \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array} \right] \left[\begin{array}{cc} 1/\sqrt{5} & 2/\sqrt{5} \\ \hline 2/\sqrt{5} & -1/\sqrt{5} \end{array} \right] \qquad \square$$

així que aquesta seria una descomposició reduïda:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \quad \square$$

En aquest exemple no hem hagut de fer cap càlcul, perquè ja coneixíem una descomposició en valors singulars completa. En general, però, i tal com hem dit adés, en calcular la descomposició reduïda ens estalviarem el càlcul de les bases dels dos espais nuls:

Algorisme de càlcul de la descomposició en valors sigulars reduïda de la matriu A

(a) Calculeu els valors propis no nuls de la matriu $\mathsf{A}^t\mathsf{A}$. Ordeneu-los de més gran a més petit: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$.

Les arrels quadrades positives, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$, d'aquests valors propis són els valors singulars de la matriu A. Això determina la matriu diagonal Σ_{11}.

(b) Calculeu bases ortonormals dels subespais propis de $\mathsf{A}^t\mathsf{A}$ corresponents als valors propis no nuls. Si B_1 és la unió ordenada d'aquestes bases, la matriu V_1 és M_{B_1}.

(c) Calculeu els vectors $\mathsf{u}_i = (1/\sigma_i)\mathsf{A}\vec{v}_i$. Això ens proporciona la matriu U_1.

Exemple 26.4.

Trobeu una factorització en valors singulars reduïda de la matriu $\mathsf{A} = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$.

(a) Els valors propis de la matriu $\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 8 & 2 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ són 9 (doble), 4 i 0. Així que els valors singulars de A, ordenats del més gran al més petit, són $3, 3, 2$. La matriu Σ_{11} és $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(b) Els espais propis de $\mathsf{A}^t\mathsf{A}$ associats als valors propis no nuls són

$$\mathrm{Nul}(\mathsf{A}^t\mathsf{A} - 9\mathrm{I}) = \mathrm{Nul} \begin{bmatrix} -1 & 2 & 0 & 0 \\ 2 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -9 \end{bmatrix} = \mathrm{Nul} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \left\langle (2/\sqrt{5}, 1/\sqrt{5}, 0, 0), (0, 0, 1, 0) \right\rangle$$

$$\mathrm{Nul}(\mathsf{A}^t\mathsf{A} - 4\mathrm{I}) = \mathrm{Nul} \begin{bmatrix} 4 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix} = \mathrm{Nul} \begin{bmatrix} 1 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \langle (-1, 2, 0, 0) \rangle \left\langle (-1/\sqrt{5}, 2/\sqrt{5}, 0, 0) \right\rangle$$

Els conjunts

$$\left\{ (2/\sqrt{5}, 1\sqrt{5}, 0, 0), (0, 0, 1, 0) \right\}, \left\{ (-1/\sqrt{5}, 2/\sqrt{5}, 0, 0) \right\}$$

són bases ortonormals dels dos espais propis.

La matriu V_1 és aquesta:

$$\mathsf{V}_1 = \begin{bmatrix} 2/\sqrt{5} & 0 & -1/\sqrt{5} \\ 1/\sqrt{5} & 0 & 2/\sqrt{5} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(c) Per a trobar la matriu U_1 calculem

$$\begin{aligned}
\vec{u}_1 &= \frac{1}{3}\mathsf{A}\vec{v}_1 = \frac{1}{3}\left(\frac{3}{\sqrt{5}}, \frac{6}{\sqrt{5}}, 0\right) = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right)\\
\vec{u}_2 &= \frac{1}{3}\mathsf{A}\vec{v}_2 = \frac{1}{3}(0,0,3) = (0,0,1)\\
\vec{u}_3 &= \frac{1}{3}\mathsf{A}\vec{v}_3 = \frac{1}{2}\left(\frac{-4}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right) = \left(\frac{-2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right)
\end{aligned}$$

Així, hem obtingut la descomposició

$$\underbrace{\begin{bmatrix} 2 & -1 & 0 & 0\\ 2 & 2 & 0 & 0\\ 0 & 0 & 3 & 0\end{bmatrix}}_{\mathsf{A}} = \underbrace{\begin{bmatrix} 1/\sqrt{5} & 0 & -2/\sqrt{5}\\ 2/\sqrt{5} & 0 & 1/\sqrt{5}\\ 0 & 1 & 0\end{bmatrix}}_{\mathsf{U}_1}\underbrace{\begin{bmatrix} 3 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 2\end{bmatrix}}_{\Sigma_{11}}\underbrace{\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} & 0 & 0\\ 0 & 0 & 1 & 0\\ -1/\sqrt{5} & 2/\sqrt{5} & 0 & 0\end{bmatrix}}_{\mathsf{V}_1^t} \qquad \square$$

26.3. El significat dels valors singulars en una transformació lineal

Perquè la transformació lineal $f(\vec{x}) = \mathsf{A}\vec{x}$ siga bijectiva el que ha de passar és que la matriu A siga invertible; en aquest cas, l'aplicació f^{-1} també és lineal i es pot calcular com $f(\vec{y}) = \mathsf{A}^{-1}\vec{y}$. Si f no és bijectiva, no hi ha una aplicació inversa perquè algun vector de l'espai d'arribada no té cap antiimatge, o algun en té més d'una (o les dues coses a l'hora). Ara bé, al capítol 20. hem vist que la restricció als espais fila i columna

$$\begin{aligned} f\colon \operatorname{Fil}\mathsf{A} &\longrightarrow \operatorname{Col}\mathsf{A}\\ \vec{x} &\rightsquigarrow f(\vec{x}) = \mathsf{A}\vec{x}\end{aligned}$$

sí que és bijectiva. A més, aquesta aplicació restringida determina perfectament la transformació f, perquè qualsevol vector $\vec{v}$ de l'espai original és la suma $\vec{v} = \vec{v}_{\operatorname{Fil}\mathsf{A}} + \vec{v}_{\operatorname{Nul}\mathsf{A}^t}$, on $\vec{v}_{\operatorname{Fil}\mathsf{A}}$ és un vector de l'espai fila i $\vec{v}_{\operatorname{Nul}\mathsf{A}^t} \in \operatorname{Nul}\mathsf{A}^t$, de manera que

$$f(\vec{v}) = f(\vec{v}_{\operatorname{Fil}\mathsf{A}}) + f(\vec{v}_{\operatorname{Nul}\mathsf{A}^t}) = f(\vec{v}_{\operatorname{Fil}\mathsf{A}})$$

Si $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$ és una descomposició en valors singulars i A és una matriu de rang r, llavors les r primeres columnes de V són una base de l'espai fila i les r primeres de U ho són de l'espai columna i

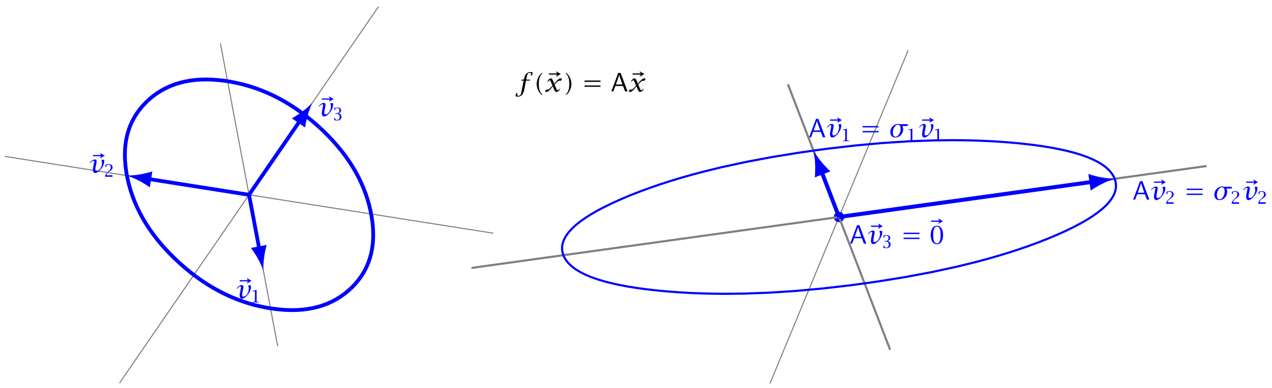
$$\begin{aligned} f(\vec{v}) &= f(\vec{v}_{\operatorname{Fil}\mathsf{A}}) = \mathsf{A}\vec{v}_{\operatorname{Fil}\mathsf{A}} = \mathsf{A}\Big((\vec{v}\cdot\vec{v}_1)\,\vec{v}_1 + (\vec{v}\cdot\vec{v}_2)\,\vec{v}_2 + \cdots + (\vec{v}\cdot\vec{v}_r)\,\vec{v}_r\Big)\\ &= (\vec{v}\cdot\vec{v}_1)\,\mathsf{A}\vec{v}_1 + (\vec{v}\cdot\vec{v}_2)\,\mathsf{A}\vec{v}_2 + \cdots + (\vec{v}\cdot\vec{v}_r)\,\mathsf{A}\vec{v}_r\end{aligned}$$

I tenint en compte que $\mathsf{A}\vec{v}_1 = \sigma_1\vec{u}_1, \mathsf{A}\vec{v}_2 = \sigma_2\vec{u}_2, \ldots$, obtenim una fórmula explícita per a calcular els valors de qualsevol transformació lineal:

$$\boxed{f(\vec{v}) = \sigma_1(\vec{v}\cdot\vec{v}_1)\,\vec{u}_1 + \sigma_2(\vec{v}\cdot\vec{v}_2)\,\vec{u}_2 + \cdots + \sigma_r(\vec{v}\cdot\vec{v}_r)\,\vec{u}_r} \tag{26.3}$$

Aquesta fórmula descriu perfectament la transformació lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$: si A és la matriu canònica de f, llavors la imatge de qualsevol vector és la mateixa que la de la projecció d'aquest vector sobre $\operatorname{Fil}\mathsf{A}$; a més, es poden trobar bases ortonormals de $\operatorname{Fil}\mathsf{A}$ i $\operatorname{Col}\mathsf{A}$, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$ i $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\}$, de manera que $f(\vec{v}_i) = \sigma_i\vec{u}_i$ i la imatge de qualsevol altre vector és una combinació lineal de les imatges dels vectors $\vec{v}_i$. Els valors singulars mesuren el pes dels vectors singulars en la transformació.

La figura adjunta representa l'aplicació $f(\vec{x}) = \mathsf{A}\vec{x}$ amb $\mathsf{A} = \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \vec{u}_3\end{bmatrix}\begin{bmatrix}\sigma_1 & 0 & 0\\ 0 & \sigma_2 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\vec{v}_1^t\\ \vec{v}_2^t\\ \vec{v}_3^t\end{bmatrix}$.



La circumferència de centre l'origen que conté els radis $\vec{v}_1$ i $\vec{v}_2$ es transforma en una el·lipse en el pla generat per $\vec{u}_1$ i $\vec{u}_2$, amb semieixos σ_1 i σ_2.[3]

El teorema següent resumeix aquests fets. Observeu que aquest teorema completa el teorema fonamental de l'àlgebra lineal i generalitza el teorema espectral.

Teorema 26.1.

Suposem que $f(\vec{x}) = \mathsf{A}\vec{x}$ és una transformació lineal i que el rang de A és r. Siga

$$\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t = \left[\vec{u}_1 \quad \vec{u}_2 \quad \dots \quad \vec{u}_r \;\middle|\; \vec{u}_{r+1} \quad \dots \quad \vec{u}_m\right] \left[\begin{array}{cccc|ccc} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right] \left[\begin{array}{c} \vec{v}_1^t \\ \vec{v}_2^t \\ \dots \\ \vec{v}_r^t \\ \hline \vec{v}_{r+1}^t \\ \dots \\ \vec{v}_n^t \end{array}\right]$$

una descomposició en valors singulars de A. Llavors

(a) La matriu A es pot expressar com una combinació lineal de matrius de rang 1,

$$\mathsf{A} = \sigma_1\vec{u}_1\vec{v}_1^t + \sigma_2\vec{u}_2\vec{v}_2^t + \cdots + \sigma_r\vec{u}_r\vec{v}_r^t$$

(b) Els escalars $\sigma_1, \sigma_2, \dots, \sigma_r$ són els valors singulars de la matriu A.

(c) $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ són vectors singulars drets, i les seues imatges,

$$f(\vec{v}_1) = \sigma_1\vec{u}_1, f(\vec{v}_2) = \sigma_2\vec{u}_2, \dots, f(\vec{v}_r) = \sigma_r\vec{u}_r$$

són vectors singulars esquerres.

(d) Les imatges de la resta de vectors $\vec{v}_i$ són nul·les: $f(\vec{v}_i) = \vec{0}$, si $i > r$

(e) La imatge de qualsevol vector $\vec{v}$ és

$$f(\vec{v}) = \sigma_1(\vec{v}\cdot\vec{v}_1)\vec{u}_1 + \sigma_2(\vec{v}\cdot\vec{v}_2)\vec{u}_2 + \cdots + \sigma_r(\vec{v}\cdot\vec{v}_r)\vec{u}_r \quad \square$$

26.4. Exercicis

Exercici 26.1. Trobeu la descomposició en valors singulars de la matriu

$$\mathsf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \end{bmatrix}$$

[3] En el cas general, si el rang de la matriu A és r, l'aplicació $f(\vec{x}) = \mathsf{A}\vec{x}$ transforma la hiperesfera unitària r-dimensional definida pels vectors singulars drets en un hiperel·lipsoide, els semieixos del qual són els valors singulars.

Exercici 26.2. Trobeu una descomposició en valors singulars reduïda de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$.

Exercici 26.3. Proveu que, si $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$ és una descomposició en valors singulars, llavors, les columnes de U són vectors singulars esquerres, les de V, vectors singulars drets, i les entrades no nul·les de Σ, els valors singulars de A.

Exercici 26.4. Proveu que, si la matriu real A és simètrica definida (o semidefinida) positiva, llavors, els valors propis no nuls de A són valors singulars. Què passa, en general, el cas de les matrius simètriques? Quina relació hi ha entre els vectors propis i els vectors singulars, en una matriu simètrica?

27. Aplicacions dels valors singulars. La pseudoinversa

En aquesta lliçó generalitzem la idea de transformació inversa i de matriu inversa quan la transformació $f(\vec{x}) = \mathsf{A}\vec{x}$ no és bijectiva, és a dir, quan la matriu A no és invertible i estudiem algunes aplicacions de la descomposició en valors singulars.

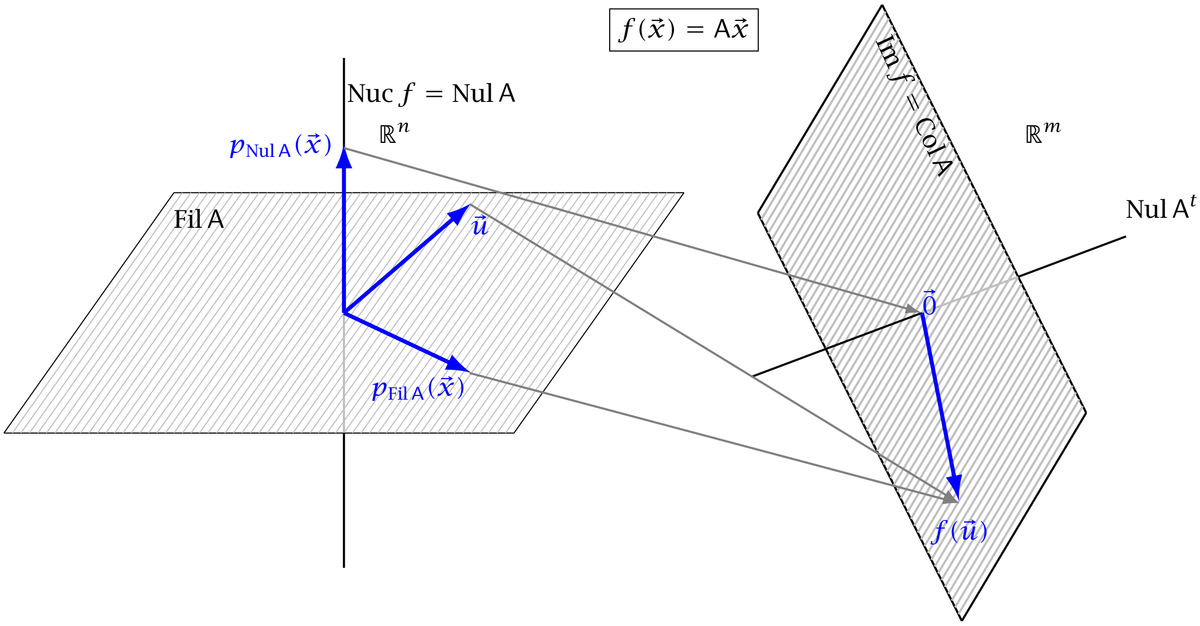
27.1. La pseudoinversa

Qualsevol transformació lineal

$$\begin{aligned} f\colon \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \vec{x} &\rightsquigarrow f(\vec{x}) = \mathsf{A}\vec{x} \end{aligned}$$

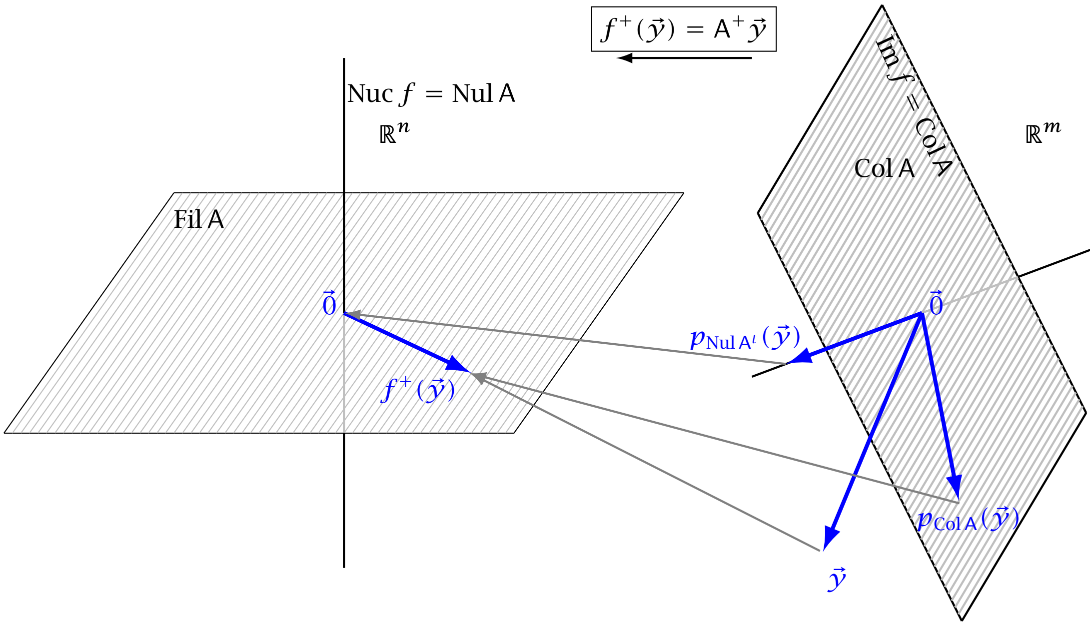
és bijectiva quan la restringim als espais fila i columna de la matriu A, i aplica l'espai nul en el vector zero. Si el vector $\vec{x}$ es descompon com $\vec{x} = p_{\mathrm{Fil}\,\mathsf{A}}(\vec{x}) + p_{\mathrm{Nul}\,\mathsf{A}}(\vec{x})$ llavors,

$$f(\vec{x}) = f\left(p_{\mathrm{Fil}\,\mathsf{A}}(\vec{x})\right) + f\left(p_{\mathrm{Nul}\,\mathsf{A}}(\vec{x})\right) = f\left(p_{\mathrm{Fil}\,\mathsf{A}}(\vec{x})\right)$$



Per tant, és raonable definir una *falsa inversa* que aplique l'espai columna de A en l'espai fila i l'espai nul esquerre en el vector zero. Com que la suma $\mathrm{Col}\,\mathsf{A} \oplus \mathrm{Nul}\,\mathsf{A}^t$ és directa, això defineix perfectament una aplicació lineal: si $\vec{y} = p_{\mathrm{Col}\,\mathsf{A}}(\vec{y}) + p_{\mathrm{Nul}\,\mathsf{A}^t}(\vec{y})$ llavors, només hi ha un vector $\vec{x}_1$ en $\mathrm{Fil}\,\mathsf{A}$ que és antiimatge de $p_{\mathrm{Col}\,\mathsf{A}}(\vec{y})$, i podem definir

$$f^+(\vec{y}) = f^+\left(p_{\mathrm{Col}\,\mathsf{A}}(\vec{y})\right) = \vec{x}_1$$



La matriu canònica associada a aquesta aplicació és la *pseudoinversa* o *inversa Moore-Penrose* de A. La representarem com A^+.

Podem fer servir els valors i els vectors singulars per a trobar una fórmula explícita per calcular l'aplicació f^+: si coneixem una descomposició en valors singulars de la matriu A, $\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^t$, sabem que, si $r = \operatorname{rang} \mathsf{A}$,

$$f(\vec{v}_1) = \sigma_1\vec{u}_1, f(\vec{v}_2) = \sigma_2\vec{u}_2, \dots f(\vec{v}_r) = \sigma_r\vec{u}_r,$$
$$f(\vec{v}_{r+1}) = \vec{0}, f(\vec{v}_{r+2}) = \vec{0}, \dots, f(\vec{v}_n) = \vec{0}$$

Per tant, com que $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r\}$ és una base de Col A i $\{\vec{u}_{r+1}, \vec{u}_{r+2}, \dots, \vec{u}_m\}$ és una base de Nul A^t, la transformació lineal $f^+ : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ és definida per les igualtats

$$\begin{gathered} f^+(\vec{u}_1) = (1/\sigma_1)\vec{v}_1, f^+(\vec{u}_2) = (1/\sigma_2)\vec{v}_2, \dots f^+(\vec{u}_r) = (1/\sigma_r)\vec{v}_r, \\ f^+(\vec{u}_{r+1}) = \vec{0}, f^+(\vec{v}_{u+2}) = \vec{0}, \dots, f^+(\vec{u}_m) = \vec{0} \end{gathered} \tag{27.1}$$

La matriu A^+ encara no la coneixem, per ara, però sí que sabem quina és la matriu de f^+ respecte a les bases $B_2 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m\}$ i $B_1 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$. Tenint en compte les fórmules (27.1),

$$\mathsf{M}(f, B_2, B_1) = \left[\begin{array}{cccc|ccc} 1/\sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1/\sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & 1/\sigma_r & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array}\right] = \begin{bmatrix} \Sigma_{11}^{-1} & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{bmatrix}$$

Aleshores, com que $\mathsf{M}(f, B_2, B_1) = \mathsf{M}_{B_1}^{-1}\mathsf{A}^+\mathsf{M}_{B_2}$, resulta que

$$\mathsf{A}^+ = \mathsf{M}_{B_1}\mathsf{M}(f, B_2, B_1)\mathsf{M}_{B_2}^{-1} = \mathsf{V}\begin{bmatrix} \Sigma_{11}^{-1} & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{bmatrix}\mathsf{U}^{-1} = \mathsf{V}\begin{bmatrix} \Sigma_{11}^{-1} & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{bmatrix}\mathsf{U}^t$$

o, millor,

$$\mathsf{A}^+ = \mathsf{V}_1\Sigma_{11}^{-1}\mathsf{U}_1^t$$

DEFINICIÓ 27.1.

Si $\mathsf{A} = \mathsf{U}_1\Sigma_{11}\mathsf{V}_1^t$ *és una descomposició en valors singulars reduïda de la matriu* A*, llavors la matriu pseudoinversa (o inversa Moore-Penrose) de* A *és* $\mathsf{A}^+ = \mathsf{V}_1\Sigma_{11}^{-1}\mathsf{U}_1^t$.

La pseudoinversa és una generalització de la matriu inversa en el sentit que si A és una matriu invertible $n \times n$, llavors les matrius U i V són ortogonals, Σ és invertible i $\mathsf{A}^+ = \mathsf{V}\Sigma^{-1}\mathsf{U}^t = (\mathsf{U}\Sigma\mathsf{V}^t)^{-1} = \mathsf{A}^{-1}$.

EXEMPLE 27.1.

Calculeu la pseudoinversa de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix}$.

En un exemple anterior hem obtingut aquesta descomposició en valors singulars:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 & 0 \\ -2/\sqrt{5} & 1/\sqrt{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La versió reduïda serà

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}}_{\mathsf{U}_1} \underbrace{\begin{bmatrix} 5 \end{bmatrix}}_{\Sigma_{11}} \underbrace{\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} & 0 & 0 \end{bmatrix}}_{\mathsf{V}_1^t}$$

i, per tant, la pseudoinversa és

$$\mathsf{A}^+ = \underbrace{\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \\ 0 \end{bmatrix}}_{\mathsf{V}_1} \underbrace{\begin{bmatrix} 1/5 \end{bmatrix}}_{\Sigma_{11}^{-1}} \underbrace{\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}} = \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \square$$

27.2. APLICACIÓ AL PROBLEMA DE MÍNIMS QUADRATS

Si la matriu A és invertible, llavors la solució del sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ és $\vec{x} = \mathsf{A}^{-1}\vec{b}$. La pseudoinversa resol el problema de la mateixa manera (en el sentit dels mínims quadrats):

PROPIETAT 27.1.

El vector $\vec{x} = \mathsf{A}^+\vec{b}$ *és una solució per mínims quadrats del sistema lineal* $\mathsf{A}\vec{x} = \vec{b}$.

Demostració: Volem provar és que $\mathsf{A}\mathsf{A}^+\vec{b} = p_{\mathrm{Col}\,\mathsf{A}}(\vec{b})$. Ara bé, si $\mathsf{A} = \mathsf{U}_1\Sigma_{11}\mathsf{V}_1^t$ és una descomposició en valors singulars, com que les columnes de U_1 són una base ortonormal de $\mathrm{Col}\,\mathsf{A}$, $p_{\mathrm{Col}\,\mathsf{A}}(\vec{b}) = \mathsf{U}\mathsf{U}^t\vec{b}$, així que el que hem de provar és que $\mathsf{A}\mathsf{A}^+\vec{b} = \mathsf{U}\mathsf{U}^t\vec{b}$:

$$\mathsf{A}\mathsf{A}^+\vec{b} = \mathsf{U}_1\Sigma_{11}\mathsf{V}_1^t\mathsf{V}_1\Sigma_{11}^{-1}\mathsf{U}_1^t\vec{b} = \mathsf{U}_1\mathsf{U}_1^t\vec{b} \qquad \square$$

EXEMPLE 27.2.

Si $\mathsf{A} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{bmatrix}$ i $\vec{b} = (30, 61)$, trobeu una solució per mínims quadrats del sistema lineal $\mathsf{A}\vec{x} = \vec{b}$ i trobeu-ne l'error de mínims quadrats.

La pseudoinversa de la matriu A és $\mathsf{A}^+ = \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. Així que el vector

$$\vec{x} = \mathsf{A}^+\vec{b} = \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 30 \\ 61 \end{bmatrix} = \begin{bmatrix} 6{,}08 \\ 12{,}16 \end{bmatrix}$$

és una solució del problema de mínims quadrats. L'error és

$$\left\|\vec{b} - \mathsf{A}\vec{x}\right\| = \|(0{,}4, -0{,}2)\| = \sqrt{0{,}2} \approx 0{,}45 \quad \square$$

27.3. Aproximacions de rang limitat

Com que la matriu A pot expressar-se com una combinació lineal de matrius de rang 1,

$$\mathsf{A} = \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_r \vec{u}_r \vec{v}_r^t$$

on els pesos són els valors singulars, podem considerar les sumes parcials

$$\begin{aligned}
\mathsf{A}_1 &= \sigma_1 \vec{u}_1 \vec{v}_1^t \\
\mathsf{A}_2 &= \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t \\
&\dots \\
\mathsf{A}_k &= \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_k \vec{u}_k \vec{v}_k^t \\
&\dots \\
\mathsf{A} = \mathsf{A}_r &= \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_k \vec{u}_k \vec{v}_k^t + \cdots + \sigma_r \vec{u}_r \vec{v}_r^t
\end{aligned}$$

com aproximacions parcials a la matriu A, de rangs $1, 2, \dots, k, \dots, r$. Intuïtivament, quan fem una suma parcial estem eliminant les matrius que menys *pesen* en la combinació (perquè els valors singulars estan ordenats de més gran a més menut), de manera que A_k deu ser una bona aproximació de A. En realitat, és la millor possible: si definim la norma d'una matriu d'aquesta manera (anàloga a la definició de la norma d'un vector),

Definició 27.2.

La norma de Frobenius de la matriu A *és*

$$\|\mathsf{A}\| = \sqrt{\sum \left|a_{ij}\right|^2}$$

i la distància entre dues matrius $m \times n$, A i B, és $\|\mathsf{A} - \mathsf{B}\|$, aleshores,

Teorema de Eckart–Young

Si $\mathsf{A} = \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_r \vec{u}_r \vec{v}_r^t$ és una descomposició de la matriu A en valors singulars, llavors la matriu de rang k més pròxima a A és

$$\mathsf{A}_k = \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_k \vec{u}_k \vec{v}_k^t \quad \square$$

Això vol dir que $\|\mathsf{A} - \mathsf{A}_k\| \leq \|\mathsf{A} - \mathsf{B}\|$, per a qualsevol matriu B de rang k. Convé notar que el valor singular més gran de la matriu error,

$$\mathsf{A} - \mathsf{A}_k = \sigma_{k+1} \vec{u}_{k+1} \vec{v}_{k+1}^t + \sigma_{k+2} \vec{u}_{k+2} \vec{v}_{k+2}^t + \cdots + \sigma_r \vec{u}_r \vec{v}_r^t$$

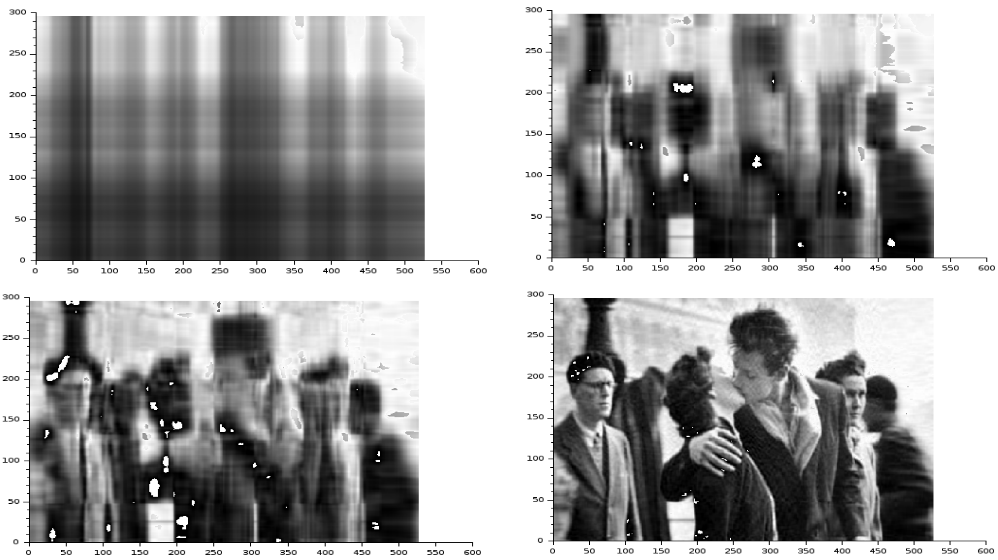
és precisament σ_{k+1}, així que el primer valor singular que menyspreem ens dona una mida de la magnitud de l'error.

27.3.1. Aplicació a la compressió d'imatges

Una imatge gràfica, com ara, una fotografia, és una matriu de punts, cadascun dels quals té un color determinat. Per simplificar, suposarem que la imatge només conté tons de gris, de manera que el color de cada punt es pot representar com un nombre: per exemple, fent servir nombres enters de 8 bits, 255 si el punt és blanc, 0 si és negre, o un nombre entre zero i 255, si el punt és d'un gris més o menys fosc (50 és un gris fosc, 225 és molt clar). Per exemple, aquesta imatge,[4]

[4]Detall de la famosa fotografia *Le baiser de l'Hôtel de Ville*, de Robert Doisneau.

conté 526 punts d'amplada per 296 d'alçada, de manera que la podem representar amb una matriu A d'aquestes dimensions (296×526), és a dir, per $296 \cdot 526 = 155696$ nombres. Fent servir l'Scilab (un sistema informàtic de càlcul numèric especialment indicat per al càlcul matricial) hem calculat els valors i els vectors singulars i hem trobat que aquesta matriu té rang màxim (296). Les imatges següents mostren les aproximacions de rangs 1, 5, 10, 50 i 75.



Notem que la informació que cal emmagatzemar per desar la matriu

$$A_{75} = \sigma_1 \vec{u}_1 \vec{v}_1^t + \sigma_2 \vec{u}_2 \vec{v}_2^t + \cdots + \sigma_{75} \vec{u}_{75} \vec{v}_{75}^t$$

consisteix en 75 valors singulars i 150 vectors singulars (75 esquerres i 75 drets), és a dir, $75(1 + 526 + 296 = 61725$ nombres (aproximadament, el 40% dels 155696 de la imatge original).

27.4. Exercicis

Exercici 27.1. Trobeu la pseudoinversa de la matriu $\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ i apliqueu el resultat per trobar la solució del problema de mínims quadrats $\mathsf{A}\vec{x} = (1, 1, 1, 2)$.

Exercici 27.2. Siga A una matriu real $m \times n$. Proveu les següents propietats a prop de la pseudoinversa:
(a) $\mathsf{A}\mathsf{A}^{+}\mathsf{A} = \mathsf{A}$
(b) $\mathsf{A}^{+}\mathsf{A}\mathsf{A}^{+} = \mathsf{A}^{+}$
(c) $\mathsf{A}\mathsf{A}^{+}$ és simètrica
(d) $\mathsf{A}^{+}\mathsf{A}$ és simètrica

Es pot provar que aquestes quatre propietats caracteritzen la pseudoinversa.

Exercici 27.3. Quina condició s'ha de complir perquè la pseudoinversa de la matriu A siga una inversa esquerra? I perquè siga una inversa dreta?

TAULA