Curs
Algebră Liniară, Geometrie Analitică şi Ecuaţii Diferenţiale

Conf. dr. Adriana Balan

Semestrul I 2021/2022

Cuprins

Capitolul I

Algebră liniară

> *Algebra is the offer made by the devil to the mathematician. The devil says: "I will give you this powerful machine, and it will answer any question you like. All you need to do is give me your soul: give up geometry and you will have this marvellous machine."*
>
> Michael Atiyah. *Mathematics in the 20th Century*, American Mathematical Monthly, 108:7(2001), 654-666

1 Vectori şi matrice

1.1 Introducere

Informal, un vector este o listă finită ordonată de numere şi scrisă pe coloană sau pe linie, cu sau fără paranteze rotunde/pătrate, cu virgulă sau fără, cum ar fi de exemplu

$$\mathbf{v} = \begin{pmatrix} 1.2 + \mathrm{i} \\ 0 \\ -3.16\,\mathrm{i} \\ 4 \end{pmatrix}$$

În funcţie de domeniul în care se întâlnesc, vectorii se mai numesc şi secvenţe, tupluri, şiruri, liste, cuvinte, etc.

Conceptul de vector a fost utilizat în învăţământul preuniversitar la Geometrie sau la Fizică. În acest curs vom aborda vectorii sub aspect **algebric**, urmând să utilizăm exemplele geometrice ca suport intuitiv.

Până la finalul acestui capitol, veţi întâlni (cel puţin!) trei importante exemple de aplicare în Computer Science a noţiunilor studiate: compresia imaginilor, recunoaşterea feţelor şi algoritmul PageRank utilizat în motorul de căutare Google pentru sortarea rezultatelor căutării în funcţie de importanţa acestora.

Vector

Definiţie 1.1. Vom numi **vector** o matrice cu o singură coloană.

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Componentele vectorului sunt $v_1, \ldots, v_n$, iar $n \in \mathbb{N}^*$ este *numărul de componente* al vectorului (sau *dimensiunea* vectorului).

În exemplul cu care am început capitolul, $\mathbf{v}$ este un vector cu 4 componente, iar componentele sale sunt numere complexe; de exemplu, a treia componentă a sa este $v_3 = -3.16\,\mathrm{i}$.

Vom nota vectorii cu simboluri îngroşate $\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$ etc., iar componentele acestora cu $u_i, v_i, w_i, \ldots$.

Doi vectori sunt egali $\mathbf{u} = \mathbf{v}$ dacă şi numai dacă au acelaşi număr de componente $n \in \mathbb{N}^*$, şi aceleaşi componente $u_i = v_i$, $\forall\, i \in \{1, \ldots, n\}$.

Fie $\Bbbk^n$ mulţimea vectorilor cu n componente din mulţimea $\Bbbk$. Pentru Algebra liniară, vom avea nevoie de o structură de corp comutativ pe $\Bbbk$. Reamintim pe scurt că un corp comutativ este o mulţime $\Bbbk$ înzestrată cu două operaţii notate uzual $+$ şi $\cdot$, şi două constante 0 şi 1, astfel încât $(\Bbbk, +, 0)$ este grup abelian, $(\Bbbk, \cdot, 1)$ este monoid comutativ, cele două operaţii distribuie una faţă de celalată, c si fiecare element din $\Bbbk$ diferit de 0 este inversabil în raport cu $\cdot$. Vom numi **scalari** elementele corpului $\Bbbk$ (în particular, şi componentele vectorilor). Vom lucra cel mai des cu corpul numerelor reale $\mathbb{R}$, dar sunt cazuri în care vom utiliza corpul numerelor complexe $\mathbb{C}$, sau corpul claselor de resturi $\mathbb{Z}_p$, unde p este un număr prim.

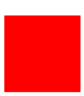
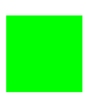
Revenind la cazul general al unui corp comutativ $\Bbbk$, vom numi vector **nul** vectorul cu toate componentele egale cu 0, $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. Dacă va fi necesar, vom scrie $\mathbf{0_n}$ pentru a evidenţia numărul de componente n ale vectorului nul.

Un vector **unitar** este un vector cu toate componentele nule, mai puţin una singură, egală cu 1. Iată mai jos vectorii unitari cu n componente:

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \tag{1}$$

Un vector se numeşte **rar** ("sparse", în limba engleză) dacă majoritatea componentelor sale sunt nule. Vectorii unitari sunt exemple de vectori rari. Cel mai rar vector este, evident, vectorul nul.

Exemplu 1.2. 1) **RGB**. Un vector cu trei componente reale (între 0 şi 255) poate reprezenta o culoare, componentele vectorului indicând intensităţile culorilor constituente – roşu (Red), verde (Green) şi albastru (Blue).

 $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  $\begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$  $\begin{pmatrix} 1 \\ 0.5 \\ 0.5 \end{pmatrix}$

Această reprezentare a culorilor se scrie uneori sub formă procentuală (de exemplu, nuanţa de roz de mai sus este $(100\%, 50\%, 50\%)$) sau convertită binar, cu valori de la 0 la 255 (deci $(255, 128, 128)$).

2) **Analiza formală a conceptelor**.[1] În multe aplicaţii un vector colectează informaţii referitoare la proprietăţi diferite ale unui singur **obiect**. Componentele vectorului sunt numite de

[1] Interfaţa Om Calculator, curs anul IV.

obicei **atribute**. Se realizează astfel o ierarhie conceptuală plecând de la o mulţime de obiecte şi proprietăţile acestora. Fiecare concept din ierarhie reprezintă obiectele care partajează o mulţime de proprietăţi. Analiza formăla a conceptelor este un instrument utilizat de exemplu în extragerea de cunoştinţe din date (data mining) sau învăţare automată, utilizând teoria matematică a structurilor ordonate şi a laticelor, pe care o vom întâlni în partea a treia a cursului.

3) **Prelucrarea limbajului natural**.[2]Documentele pot fi percepute ca vectori într-un spaţiu multi-dimensional, în care fiecare dimensiune corespunde unui cuvânt posibil. Mai precis, componenta respectivă a documentului/vectorului indică de numărul de apariţii ale cuvântului corespunzător în document (1 dacă apare, 0 dacă nu). Clusterele de documente astfel obţinute corespund deseori grupurilor de documente din acelaşi domeniu.

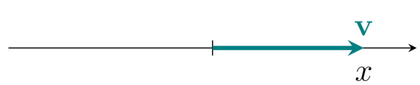
1.2 Interpretarea geometrică a vectorilor

> *"Four man, who had been blind from birth, wanted to know what an elephant was like, so they asked an elephant driver for information. He led them to an elephant, and invited them to examine it, so one man felt the elephant's leg, another its trunk, another its tail and the fourth its ear. Then they attempted to describe the elephant to one another. The first man said 'The elephant is like a tree'. 'No', said the second, 'the elephant is like a snake'. 'Nonsense', said the third, 'the elephant is like a broom'. 'You are all wrong', said the fourth, 'the elephant is like a fan'. And so they went on arguing amongst themselves, while the elephant stood watching them quietly."*
>
> P.T. Johnstone, *Sketches of an Elephant: A Topos Theory Compendium*, Oxford University Press (2002).

Vectorii coloană cu componente reale $\mathbf{v} = \big(v_i\big)_{i=\overline{1,n}} \in \mathbb{R}^n$ pot fi puşi în corespondenţă bijectivă cu punctele unui spaţiu n-dimensional. Astfel, pentru $n \leq 3$, se obţin următoarele:

1) $n = 0$: o mulţime cu un singur punct $\{0\}$.[3]

2) $n = 1$: dreapta reală $\mathbb{R}$. Fiecare vector $\mathbf{v} = \big(x\big)$ are o singură componentă $x \in \mathbb{R}$, şi se poate identifica cu aceasta.

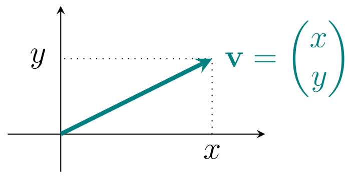


3) $n = 2$: planul euclidian $\mathbb{R}^2$. Mai precis, dacă notăm un vector $\mathbf{v} \in \mathbb{R}^2$ prin $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$, atunci acestuia îi corespunde un punct din plan (vârful vectorului) şi reciproc, fiecărui punct (x, y)

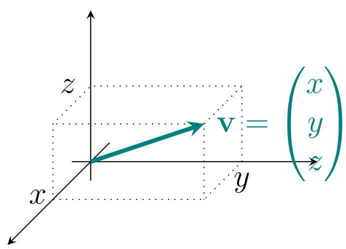
[2]Inteligenţă Artificială, curs anul IV.

[3]Vom vedea mai târziu, în partea a III-a a cursului, că mulţimea $\mathbb{R}^n$ este un exemplu de *produs cartezian*, şi motivaţia pentru care $\mathbb{R}^0$ este o mulţime cu un singur element este determinată de cardinalitate.

din plan îi corespunde un vector având originea în $(0,0)$ şi vârful în (x,y), numit **vectorul de poziţie** asociat.



4) $n = 3$: spaţiul euclidian (tridimensional) $\mathbb{R}^3$. Ca mai sus, există o corespondenţă bijectivă între punctele din spaţiu (x,y,z) şi vectorii de poziţie asociaţi $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$.



Intuitiv, acestea sunt două reprezentări echivalente ale aceluiaşi concept. În funcţie însă de ceea ce dorim să obţinem, putem utiliza abordarea algebrică (vectori coloană) sau geometrică (puncte în spaţiul euclidian).

1.3 Operaţii cu vectori.

Vectorii sunt matrice cu o singură coloană; în particular, operaţiile cu matrice induc operaţii cu vectori. Vom menţiona pentru moment doar două operaţii:

1) Adunarea vectorilor

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \ \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \implies \mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

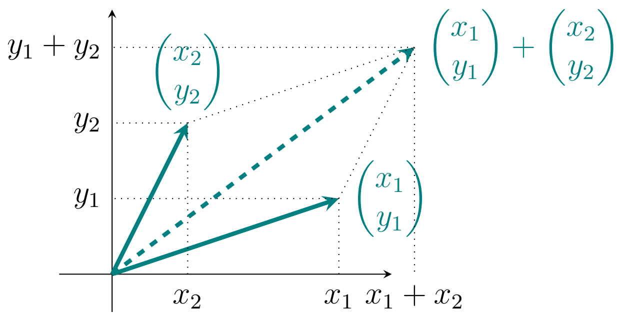
2) Înmulţirea vectorilor cu scalari

$$\alpha \in \mathbb{k}, \ \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{k}^n \implies \alpha \cdot \mathbf{v} = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix}$$

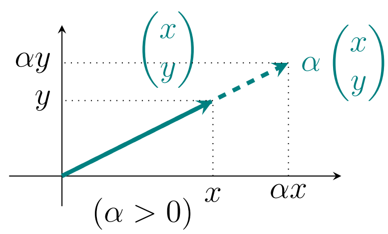
Dacă nu există pericol de confuzie, vom nota cu $\alpha\mathbf{v}$ rezultatul înmulţirii unui vector $\mathbf{v}$ cu un scalar α, în loc de $\alpha \cdot \mathbf{v}$.

1.4 Interpretarea geometrică a operaţiilor cu vectori

1) Adunarea vectorilor $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}$ corespunde geometric regulii paralelogramului:



2) Înmulţirea unui vector $\begin{pmatrix} x \\ y \end{pmatrix}$ cu un scalar α, şi anume $\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix}$ conduce, din punct de vedere geometric, la un vector coliniar cu vectorul $\begin{pmatrix} x \\ y \end{pmatrix}$, având acelaşi sens, dacă scalarul α este pozitiv, respectiv de sens opus, dacă $\alpha < 0$. Pentru $\alpha = 0$ se obţine vectorul nul: $0 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \mathbf{0}$.



1.5 Proprietăţile operaţiilor cu vectori

Întrucât operaţiile cu vectori sunt de fapt operaţii cu matrice, verifică aceleaşi proprietăţi, determinate de proprietăţile operaţiilor din corpul $\Bbbk$:

1) (a) Comutativitatea adunării vectorilor

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \forall\ \mathbf{u}, \mathbf{v} \in \Bbbk^n$$

(b) Asociativitatea adunării vectorilor

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \forall\ \mathbf{u}, \mathbf{v}, \mathbf{w} \in \Bbbk^n$$

(c) Existenţa elementului neutru în raport cu operaţia de adunare a vectorilor (vectorul nul)

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \Bbbk^n,$$

$$\mathbf{v} + \mathbf{0}_n = \mathbf{0}_n + \mathbf{v} = \mathbf{v}, \forall\ \mathbf{v} \in \Bbbk^n$$

Dacă nu există pericol de confuzie, vom nota simplu $\mathbf{0}$ în loc de $\mathbf{0}_n$.

(d) Existenţa simetricului fiecărui vector în raport cu operaţia de adunare a vectorilor: pentru orice $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, vectorul $-\mathbf{v} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$ verifică relaţiile

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}_n = (-\mathbf{v}) + \mathbf{u}$$

2) (a) "Asociativitatea"[4] înmulţirii scalarilor faţă de înmulţirea vectorilor cu scalari

$$(\alpha\mu)\mathbf{v} = \alpha(\mu\mathbf{v}), \forall\ \alpha, \mu \in \mathbb{k}, \forall\ \mathbf{v} \in \mathbb{k}^n$$

(b) Înmulţirea vectorilor cu scalarul 1 (elementul neutru pentru operaţia de înmulţire a scalarilor)

$$1\mathbf{v} = \mathbf{v}, \forall\ \mathbf{v} \in \mathbb{k}^n$$

3) (a) Distributivitatea înmulţirii cu scalari faţă de adunarea vectorilor

$$\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}, \forall\ \alpha \in \mathbb{k}, \forall\ \mathbf{u}, \mathbf{v} \in \mathbb{k}^n$$

(b) Distributivitatea adunării scalarilor faţă de înmulţirea cu scalari a vectorilor

$$(\alpha + \mu)\mathbf{v} = \alpha\mathbf{v} + \mu\mathbf{v}, \forall\ \alpha, \mu \in \mathbb{k}, \forall\ \mathbf{v} \in \mathbb{k}^n$$

[4] Formal, asociativitatea este o proprietate caracteristică unei operaţii binare interne, de exemplu multiplicarea scalarilor.

2 Spaţii vectoriale

2.1 Spaţii vectoriale – definiţie, exemple

Generalizând modelul vectorilor coloană, se obţine noţiunea de **spaţiu vectorial**; informal, un spaţiu vectorial este o colecţie de obiecte numite vectori, înzestrată cu două operaţii: adunare şi multiplicare (scalare) cu numere, numite scalari.

Spaţiu vectorial

Definiţie 2.1. Fie $\Bbbk$ un corp comutativ. Un **spaţiu vectorial** peste corpul $\Bbbk$ ($\Bbbk$-spaţiu vectorial) este o mulţime nevidă V, ale cărei elemente le vom numi **vectori**, înzestrată cu două operaţii:

1) **adunarea** vectorilor $+ : V \times V \to V$

$$\mathbf{u}, \mathbf{v} \in V \Longrightarrow \mathbf{u} + \mathbf{v} \in V$$

2) **multiplicarea** vectorilor cu scalari $\cdot : \Bbbk \times V \to V$

$$\alpha \in \Bbbk, \mathbf{v} \in V \Longrightarrow \alpha \cdot \mathbf{v} \in V$$

care verifică următoarele relaţii:

1) (a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
   (b) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
   (c) $\exists\ \mathbf{0} \in V$ astfel încât $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}$
   (d) $\forall\ \mathbf{v} \in V \implies \exists\ -\mathbf{v} \in V$ astfel încât $(-\mathbf{v}) + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
2) (a) $\alpha \cdot (\mu \cdot \mathbf{v}) = (\alpha\mu) \cdot \mathbf{v}$
   (b) $1 \cdot \mathbf{v} = \mathbf{v}$
3) (a) $\alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$
   (b) $(\alpha + \mu) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \mu \cdot \mathbf{v}$

pentru orice $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $\alpha, \mu \in \Bbbk$

Se observă cu uşurinţă că cele opt proprietăţi de mai sus coincid cu proprietăţile enumerate în secţiunea 1.5 pentru vectorii coloană. Dar pentru că acum noţiunile de spaţiu vectorial/vector/scalar sunt abstracte, vom verifica în cele ce urmează validitatea următoarelor proprietăţi:

Proprietăţi ale operaţiilor cu vectori

Propoziţie 2.2. În orice $\Bbbk$-spaţiu vectorial V, au loc următoarele relaţii:

1) Vectorul nul $\mathbf{0}$ este unic.
2) Simetricul unui vector $\mathbf{v}$ este unic.

3) $0 \cdot \mathbf{v} = \mathbf{0}$

4) $(-1) \cdot \mathbf{v} = -\mathbf{v}$

5) $\alpha \cdot \mathbf{0} = \mathbf{0}$

6) $\alpha \cdot \mathbf{v} = \mathbf{0} \iff \alpha = 0$ sau $\mathbf{v} = \mathbf{0}$

7) $(-\alpha) \cdot \mathbf{v} = \alpha \cdot (-\mathbf{v}) = -(\alpha \cdot \mathbf{v})$

Demonstraţie. 1) Să presupunem că există doi vectori nuli, notaţi $\mathbf{0}_1$ şi $\mathbf{0}_2$. Atunci:

$$\begin{aligned} \mathbf{0}_1 &= \mathbf{0}_1 + \mathbf{0}_2 && (\mathbf{0}_2 \text{ este element neutru pentru “+”}) \\ &= \mathbf{0}_1 && (\mathbf{0}_1 \text{ este element neutru pentru “+”}) \end{aligned}$$

2) Dacă vectorul $\mathbf{v}$ admite atât pe $-\mathbf{v}$ cât şi pe $\widetilde{-\mathbf{v}}$ drept simetrice în raport cu operaţia “+”, atunci

$$\begin{aligned} -\mathbf{v} &= -\mathbf{v} + \mathbf{0} && (\mathbf{0} \text{ este element neutru pentru “+”}) \\ &= -\mathbf{v} + (\mathbf{v} + \widetilde{-\mathbf{v}}) && (\widetilde{-\mathbf{v}} \text{ este simetricul lui } \mathbf{v} \text{ în raport cu “+”}) \\ &= (-\mathbf{v} + \mathbf{v}) + \widetilde{-\mathbf{v}} && (\text{asociativitatea “+”}) \\ &= \mathbf{0} + \widetilde{-\mathbf{v}} && (-\mathbf{v} \text{ este simetricul lui } \mathbf{v} \text{ în raport cu “+”}) \\ &= \widetilde{-\mathbf{v}} && (\mathbf{0} \text{ este element neutru pentru “+”}) \end{aligned}$$

3) Pentru orice $\mathbf{v} \in V$, avem $\mathbf{v} = 1 \cdot \mathbf{v} = (0+1) \cdot \mathbf{v} = 0 \cdot \mathbf{v} + 1 \cdot \mathbf{v} = 0 \cdot \mathbf{v} + \mathbf{v}$, şi analog $\mathbf{v} = \mathbf{v} + 0 \cdot \mathbf{v}$. Cum elementul neutru în raport cu operaţia de adunare a vectorilor (vectorul nul) este unic, rezultă $0 \cdot \mathbf{v} = \mathbf{0}$.

4) Rezultă din unicitatea simetricului unui vector şi din relaţia $\mathbf{0} = 0 \cdot \mathbf{v} = (1 + (-1)) \cdot \mathbf{v}$, pentru orice $\mathbf{v} \in V$.

5) Avem

$$\begin{aligned} \alpha \cdot \mathbf{0} &= \alpha \cdot (\mathbf{0} + \mathbf{0}) \\ &= \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0} \end{aligned}$$

de unde, adunând în ambii membri $-(\alpha \cdot \mathbf{0})$ şi utilizând asociativitatea, obţinem $\alpha \cdot \mathbf{0} = \mathbf{0}$.

6) Fie $\alpha \in \Bbbk$ şi $\mathbf{v} \in V$ astfel încât $\alpha \cdot \mathbf{v} = \mathbf{0}$. Să presupunem că $\alpha \neq 0$; atunci α este inversabil în raport cu înmulţirea scalarilor. Înmulţind relaţia $\alpha \cdot \mathbf{v} = \mathbf{0}$ cu α^{-1} şi ţinând cont de proprietăţile 2) (a) şi 2) (b) din Definiţia 2.1 şi de formula demonstrată la punctul anterior, rezultă

$$\mathbf{v} = 1 \cdot \mathbf{v} = (\alpha^{-1}\alpha) \cdot \mathbf{v} = \alpha^{-1} \cdot (\alpha \cdot \mathbf{v}) = \alpha^{-1} \cdot \mathbf{0} = \mathbf{0}$$

7) Exerciţiu.

□

Pentru simplificare, vom scrie în continuare $\alpha\mathbf{v}$ în loc de $\alpha \cdot \mathbf{v}$ pentru multiplicarea cu scalari.

Exemplu 2.3. Am întâlnit deja spaţiul vectorial al vectorilor coloană cu n componente $\Bbbk^n$. Vom vedea ulterior că acesta este prototipic: orice spaţiu vectorial *finit dimensional* este *izomorf*[5] cu un spaţiu vectorial de tip $\Bbbk^n$. Până atunci, să menţionăm câteva de exemple naturale de spaţii vectoriale la care vom face referire în mod continuu pe parcursul acestui curs:

1) Spaţiul vectorial al matricelor $\mathcal{M}_{m,n}(\Bbbk)$, cu operaţiile uzuale de adunare.

2) Spaţiul vectorial al polinoamelor $\Bbbk[X]$, cu operaţiile de adunare a polinoamelor şi multiplicare cu scalari.

3) Fie M o mulţime arbitrară. Vom nota cu $\Bbbk^M$ mulţimea funcţiilor $f : M \to \Bbbk$. Atunci $\Bbbk^M$ devine spaţiu vectorial peste corpul $\Bbbk$, cu operaţiile definite punctual:
$$(f+g)(x) = f(x) + g(x) \qquad (\alpha f)(x) = \alpha f(x) \quad x \in M$$
unde $f, g \in \Bbbk^M,\ \alpha \in \Bbbk$.

4) Fiecare corp $\Bbbk$ este spaţiu vectorial peste el însuşi (se identifică cu spaţiul vectorial al vectorilor coloană cu o singură componentă $\Bbbk^1$). În particular, corpul numerelor complexe $\mathbb{C}$ este $\mathbb{C}$-spaţiu vectorial. Păstrând operaţia de adunare a numerelor complexe, dar restricţionând multiplicarea numerelor complexe doar cu numere reale, $\mathbb{C}$ devine $\mathbb{R}$-spaţiu vectorial. Vom vedea că aceste două ultime exemple, $\mathbb{C}$ ca spaţiu vectorial complex şi $\mathbb{C}$ ca spaţiu vectorial real, au proprietăţi diferite, deşi au aceeaşi mulţime suport $\mathbb{C}$.

2.2 Combinaţii liniare

Combinaţie liniară

Definiţie 2.4. Fie V un $\Bbbk$-spaţiu vectorial. Dacă $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt vectori din V şi $\alpha_1, \ldots \alpha_n \in \Bbbk$ sunt scalari, vectorul
$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$$
se numeşte **combinaţie liniară** a vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_n$, cu coeficienţii $\alpha_1, \ldots \alpha_n$.

Observaţie 2.5. 1) Fie $V = \Bbbk^n$; atunci orice vector $\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \Bbbk^n$ se poate scrie ca o combinaţie liniară a vectorilor unitari $\mathbf{e}_1, \ldots, \mathbf{e}_n$:
$$\mathbf{v} = v_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + v_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = v_1 \mathbf{e}_1 + \ldots + v_n \mathbf{e}_n$$
De exemplu,
$$\begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-3) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

[5] Noţiunile de dimensiune a unui spaţiu vectorial şi de izomorfism între două spaţii vectoriale vor fi discutate în Secţiunile 4.1,5.

2) Vectorul nul se poate scrie întotdeauna ca o combinaţie liniară a oricăror vectori $\mathbf{v}_1, \ldots, \mathbf{v}_n$, luând toţi coeficienţii nuli:
$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_n$$

3) Orice vector $\mathbf{v}$ se poate scrie ca o combinaţie liniară a sa, împreună cu oricare alţi vectori $\mathbf{v}_1, \ldots, \mathbf{v}_n$:
$$\mathbf{v} = 1\mathbf{v} + 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_n$$

4) Să menţionăm şi următoarele combinaţii liniare speciale ale vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_n$: mai întâi, luând toţi scalarii egali cu unitatea $\alpha_1 = \ldots = \alpha_n = 1$, obţinem *suma* vectorilor
$$\mathbf{v}_1 + \ldots + \mathbf{v}_n$$
Al doilea exemplu se obţine pentru scalarii $\alpha_1 = \ldots = \alpha_n = \frac{1}{n}$, şi anume *media* vectorilor
$$\frac{1}{n}(\mathbf{v}_1 + \ldots + \mathbf{v}_n)$$
Acesta din urmă este un caz particular de *combinaţie liniară convexă* (cunoscută şi sub denumirea de medie ponderată):
$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n, \text{ unde } \alpha_1, \ldots, \alpha_n \in [0,1], \alpha_1 + \ldots + \alpha_n = 1$$
De exemplu, $0.2\mathbf{v}_1 + 0.35\mathbf{v}_2 + 0.45\mathbf{v}_3$ este o combinaţie liniară convexă a vectorilor $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Exemplu 2.6. 1) **Culori RGB şi combinatii liniare.** În exemplul 1.2.1 am văzut cum vectorii servesc la reprezentarea numerică a culorilor. Putem vizualiza fiecare asemenea vector în spaţiul tridimensional. De exemplu, la adresa `https://www.geogebra.org/m/Dq2A7aRv` puteţi accesa o aplicaţie online realizată în GeoGebra pe tema reprezentării vectoriale a culorilor: aici $(255, 0, 0)$ reprezintă roşu, $(0, 255, 0)$ - verde, $(0, 0, 255)$ - albastru. Orice altă culoare se obţine ca o **combinaţie liniară** a acestor culori elementare.

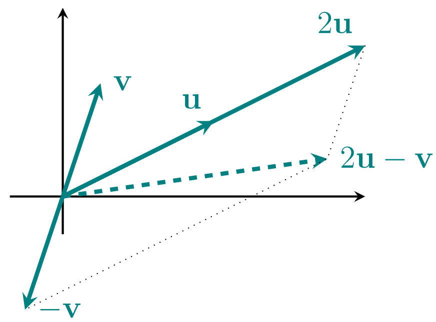
2) **Sisteme cuantice.** Dacă unitatea de informaţie în calculatoarele clasice este bitul, cu doar două valori posibile 0 şi 1, calculatoarele cuantice lucrează cu **qubiţi**, reprezentaţi matematic sub forma unei **combinaţii liniare** (superpoziţie de stări): $\alpha\,|0\rangle + \beta\,|1\rangle$ a qubiţilor elementari $|0\rangle$ şi $|1\rangle$, unde $\alpha, \beta \in \mathbb{C}$, $|\alpha|^2 + |\beta|^2 = 1$. În urma unei măsurători asupra qubitului, rezultatul va fi $|0\rangle$, cu probabilitate $|\alpha|^2$, sau $|1\rangle$, cu probabilitate $|\beta|^2$.[6]

2.3 Interpretarea geometrică a combinaţiilor liniare

În spaţiul vectorial real $\mathbb{R}^2$, o combinaţie liniară formată cu un singur vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ este un vector de forma $\alpha\mathbf{v} = \begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix}$, unde $\alpha \in \mathbb{R}$. Geometric, vectorii $\mathbf{v}$ şi $\alpha\mathbf{v}$ sunt coliniari (au aceeaşi dreaptă suport) – a se vedea figura din Exemplul 1.4.2. Prin extensie, vom numi **vectori coliniari** orice vectori de forma $\mathbf{v}, \alpha\mathbf{v}$, unde $\mathbf{v}$ aparţine unui spaţiu vectorial oarecare V peste un corp $\mathbb{k}$ (deci $\alpha \in \mathbb{k}$).

[6] Introducere în Quantum Computing, curs master anii V-VI.

O combinaţie liniară a doi vectori $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$, de forma $\alpha\mathbf{u} + \beta\mathbf{v}$, unde $\alpha, \beta \in \mathbb{R}$, geometric se interpretează ca vectorul obţinut în urma aplicării regulii paralelogramului pentru vectorii $\alpha\mathbf{u}$ şi $\beta\mathbf{v}$ (observaţi că $\alpha\mathbf{u}$ este coliniar cu $\mathbf{u}$, respectiv $\beta\mathbf{v}$ este coliniar cu cu $\mathbf{v}$). În figura de mai jos am ilustrat combinaţia liniară $2\mathbf{u} - \mathbf{v}$:



Exerciţiu 2.7. 1) Fie $\mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix}, \mathbf{v} = \begin{pmatrix} z \\ t \end{pmatrix} \in \mathbb{R}^2$ vectori nenuli. Arătaţi că $\mathbf{u}$ şi $\mathbf{v}$ sunt coliniari dacă şi numai dacă matricea $\begin{pmatrix} x & z \\ y & t \end{pmatrix}$ are rangul 1.

2) Deduceţi că dacă $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ sunt nenuli şi necoliniari, atunci mulţimea tuturor combinaţiilor liniare ce se pot forma cu aceşti doi vectori este chiar $\mathbb{R}^2$.

3) Ce interpretare geometrică puteţi da unei combinaţii liniare convexe $\alpha\mathbf{u} + \beta\mathbf{v}$, unde $\mathbf{u}, \mathbf{v}$ sunt vectori nenuli din $\mathbb{R}^2$, iar scalarii α şi β verifică $\alpha, \beta \geq 0$ şi $\alpha + \beta = 1$? Dar în cazul particular $\alpha = \beta = \frac{1}{2}$?

2.4 Liniar independenţa unei mulţimi (finite) de vectori

Formalizarea conceptelor geometrice de necoliniaritate şi necoplanaritate în cadrul algebrei liniare se realizează cu ajutorul noţiunii de *liniar independenţă*.

Liniar (in)dependenţă

Definiţie 2.8. Vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ sunt **liniar independenţi** dacă ecuaţia

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}, \text{ unde } \alpha_1, \ldots, \alpha_n \in \mathbb{k}$$

are doar soluţia $\alpha_1 = \ldots = \alpha_n = 0$.
În caz contrar, vectorii se numesc **liniar dependenţi**.
Noţiunea de liniar independenţă se poate extinde şi pentru mulţimi *infinite* de vectori, prin condiţia ca orice submulţime *finită* să conţină doar vectori liniar independenţi.

Observaţii 2.9. 1) Liniar (in)dependenţa unei mulţimi de vectori nu depinde de ordinea acestora.

2) Un singur vector $\mathbf{v} \in V$ este liniar independent dacă şi numai dacă $\mathbf{v} \neq \mathbf{0}$.

3) O mulţime de vectori care conţine vectorul nul $\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{0}\}$ este întotdeauna liniar dependentă:

$$0 \cdot \mathbf{v}_1 + \ldots + 0 \cdot \mathbf{v}_n + 1 \cdot \mathbf{0} = \mathbf{0}$$

4) Vectorii $\mathbf{v}_1, \ldots \mathbf{v}_n$ sunt **liniar dependenţi** dacă şi numai dacă cel puţin unul dintre vectori este *o combinaţie liniară* a celorlalţi vectori.

   Într-adevăr, să presupunem că $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt liniar dependenţi. Atunci există $\alpha_1, \ldots, \alpha_n \in \mathbb{k}$, nu toţi nuli, astfel încât
$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0}$$
   Întrucât cel puţin un scalar este nenul, putem presupune $\alpha_n \neq 0$ (eventual renumerotând vectorii). Multiplicând relaţia de mai sus cu α_n^{-1}, observăm că $\mathbf{v}_n$ este o combinaţie liniară a celorlalţi vectori:
$$\mathbf{v}_n = (-\alpha_1 \alpha_n^{-1})\mathbf{v}_1 + \ldots + (-\alpha_{n-1}\alpha_n^{-1})\mathbf{v}_{n-1}$$
   Reciproc, dacă unul dintre vectori este o combinaţie liniară a celorlalţi, putem din nou, eventual renumerotând, presupune că
$$\mathbf{v}_n = \beta_1 \mathbf{v}_1 + \ldots \beta_{n-1}\mathbf{v}_{n-1}$$
   de unde obţinem combinaţia liniară nulă
$$\beta_1 \mathbf{v}_1 + \ldots \beta_{n-1}\mathbf{v}_{n-1} + (-1)\mathbf{v}_n = 0$$
   în care cel puţin un scalar este nenul, deci $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt liniar dependenţi.

5) **Combinaţii liniare de vectori liniar independenţi.** Dacă $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt vectori liniar independenţi în V şi dacă
$$\alpha_1 \mathbf{v}_1 + \ldots \alpha_n \mathbf{v}_n = \beta_1 \mathbf{v}_1 + \ldots + \beta_n \mathbf{v}_n$$
   atunci $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$.

   Deci dacă un vector se poate scrie ca o combinaţie liniară a unor vectori liniar independenţi, atunci această combinaţie liniară este **unică**.

   Într-adevăr, din relaţia de mai sus rezultă
$$(\alpha_1 - \beta_1)\mathbf{v}_1 + \ldots + (\alpha_n - \beta_n)\mathbf{v}_n = \mathbf{0}$$
   de unde, ţinând cont de independenţa vectorilor, rezultă $\alpha_1 - \beta_1 = 0, \ldots, \alpha_n - \beta_n = 0$,adicaă $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$.

6) Dacă $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ sunt liniar dependenţi, atunci şi $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k} \in V$ sunt vectori liniar dependenţi, pentru orice $\mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k} \in V$.

   Într-adevăr, dacă $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ sunt liniar dependenţi, atunci există $\alpha_1, \ldots, \alpha_n \in \mathbb{k}$ nu toţi nuli, astfel încât $\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0}$. Atunci pentru orice alţi vectori $\mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k} \in V$, putem scrie combinaţia liniară nulă
$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n + 0\mathbf{v}_{n+1} + \ldots + 0\mathbf{v}_{n+k} = \mathbf{0}$$
   în care cel puţin un scalar este nenul, deci vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k}$ sunt liniar dependenţi.

   Afirmaţia 6 de mai sus este echivalentă cu următoarea: orice submulţime a unei mulţimi finite de vectori liniar independenţi, este de asemenea formată cu vectori liniar independenţi.

Exemplu 2.10. 1) Să cercetăm liniar independenţa vectorilor

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 4 \\ 1 \\ -1 \\ 2 \end{pmatrix} \in \mathbb{R}^4$$

Fie $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ astfel încât $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}_3$. Echivalent, scalarii $\alpha_1, \alpha_2, \alpha_3$ verifică

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2 \\ 3 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \iff \begin{cases} \alpha_1 - 2\alpha_2 + 4\alpha_3 = 0 \\ 2\alpha_1 + 3\alpha_2 + \alpha_3 = 0 \\ \alpha_2 - \alpha_3 = 0 \\ \alpha_1 + 2\alpha_3 = 0 \end{cases}$$

de unde obţinem $\alpha_1 = -2\alpha_3, \alpha_2 = \alpha_3$. Întrucât sistemul de mai sus nu admite doar soluţia banală $\alpha_1 = \alpha_2 = \alpha_3 = 0$, rezultă că vectorii $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sunt liniar dependenţi. O relaţie de dependenţă liniară este $(-2)\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}_3$.

2) Monoamele $1, X, X^2, \ldots, X^n, \ldots$ sunt liniar independente în $\Bbbk[X]$. Întrucât

$$\{1, X, X^2, \ldots, X^n, \ldots\}$$

este o mulţime infinită de vectori, conform Definiţiei 2.8, va trebui să verificăm dacă orice submulţime finită a sa este liniar independentă.

Fie în acest sens o submulţime finită de monoame $\{X^{n_1}, \ldots, X^{n_k}\}$ şi să presupunem că $n_1 < \ldots < n_k$. Fie $\alpha_1, \ldots, \alpha_k \in \Bbbk$ astfel încât $\alpha_1 X^{n_1} + \ldots + \alpha_k X^{n_k} = 0$, unde în membrul drept avem polinomul nul. Întrucât gradele $n_1, \ldots, n_k$ sunt distincte, rezultă $\alpha_1 = \ldots = \alpha_k = 0$.

Deoarece submulţimea $\{X^{n_1}, \ldots, X^{n_k}\}$ a fost aleasă arbitrar, rezultă că orice submulţime finită de monoame este liniar independentă, deci $1, X, X^2, \ldots, X^n, \ldots$ sunt liniar independenţi în $\Bbbk[X]$.

3) Să verificăm liniar (in)dependenţa matricelor $A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{C})$ peste corpul numerelor complexe: fie $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ astfel încât

$$\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Rezultă

$$\alpha_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \iff \begin{cases} \alpha_3 = 0 \\ \alpha_1 - \alpha_2 i = 0 \\ \alpha_1 + \alpha_2 i = 0 \\ -\alpha_3 = 0 \end{cases}$$

de unde $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Matricele A_1, A_2, A_3 se numesc matrice Pauli şi împreună cu matricea identică, formează o bază[7] în spaţiul matricelor unitare şi hermitice[8], adică a observabilelor cuantice pentru sisteme cuantice bipartite.[9]

[7]Noţiunea de bază va fi introdusă în Secţiunea 4.1.

[8]Matricele pătratice simetrice şi hermitice vor fi discutate în Secţiunea 8.

[9]M. A. Nielsen, I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press (2000).

2.5 Verificarea liniar independenţei vectorilor coloană

Vom vedea în continuare că studiul liniar independenţei vectorilor coloană se reduce la rezolvarea unui sistem de ecuaţii liniar şi omogen, respectiv la determinarea rangului matricei formate cu aceşti vectori coloane:

Liniar (in)dependenţa vectorilor coloană

Propoziţie 2.11. Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^m$ vectori şi $A = (\mathbf{v}_1 \mid \mathbf{v}_2 \mid \ldots \mid \mathbf{v}_n)$ matricea ale cărei coloane sunt vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (să observăm că $A \in \mathcal{M}_{m,n}(\Bbbk)$).

Atunci vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt **liniar independenţi** dacă şi numai dacă sistemul de ecuaţii liniare şi omogene

$$A\mathbf{x} = \mathbf{0}_m$$

admite doar soluţia banală $\mathbf{x} = \mathbf{0}_n$.

Demonstraţie. Vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt **liniar independenţi** dacă şi numai dacă ecuaţia

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0} \tag{2}$$

unde $\alpha_1, \ldots, \alpha_n \in \Bbbk$ sunt scalari, are doar soluţia banală $\alpha_1 = \ldots = \alpha_n = 0$.

Să observăm că ecuaţia (2) se mai scrie echivalent

$$\underbrace{(\mathbf{v}_1 \mid \mathbf{v}_2 \mid \ldots \mid \mathbf{v}_n)}_{A} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \mathbf{0}$$

adică $A\mathbf{x} = \mathbf{0}$, unde $\mathbf{x} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \Bbbk^n$. Deci liniar independenţa vectorilor este echivalentă cu *unicitatea* soluţiei sistemului $A\mathbf{x} = \mathbf{0}$ (fiind un sistem liniar omogen, acesta va avea mereu cel puţin o soluţie, şi anume soluţia banală). □

Liniar (in)dependenţă şi rang

Corolar 2.12. 1) Vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^m$ sunt liniar independenţi dacă şi numai dacă $m \geq n$ (numărul componentelor fiecărui vector este mai mare decât numărul vectorilor) şi $\mathsf{rang}\,(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) = n$.

2) În particular, dacă $m = n$, vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^n$ sunt liniar independenţi dacă şi numai dacă $\mathsf{det}\,(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \neq 0$.

Exemplu 2.13. 1) Vectorii $\mathbf{e}_1, \ldots, \mathbf{e}_n$ sunt liniar independenţi în $\Bbbk^n$. Într-adevăr, avem

$$\mathsf{det}\,(\mathbf{e}_1 \mid \mathbf{e}_2 \mid \ldots \mid \mathbf{e}_n) = \mathsf{det} \left(\begin{array}{c|c|c|c} 1 & 0 & \ldots & 0 \\ 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{array}\right) = 1 \neq 0$$

2) Vectorii $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sunt liniar dependenţi: sunt $n = 3$ vectori cu $m = 2$ componente. Întrucat numărul componentelor este mai mic decât numărul vectorilor, conform corolarului de mai sus, rezultă că $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sunt liniar dependenţi.

3) Vectorii $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ -1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} \in \mathbb{R}^3$ sunt liniar dependenţi. Pentru a verifica aceasta, să observăm pentru început că numărul vectorilor este $n = 3$, iar numărul componentelor este $m = 4$. Relaţia $m \geq n$ nu ne spune nimic despre liniar independenţă; mergem mai departe şi calculăm

$$\mathsf{rang}\,(\mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_n) = \mathsf{rang}\left(\begin{array}{r|r|r} 1 & 4 & 2 \\ 2 & 5 & 1 \\ 3 & 6 & 0 \\ -1 & 1 & 3 \end{array}\right) = 2$$

Întrucât $\mathsf{rang}\,(\mathbf{v}_1 \mid \mathbf{v}_2 \mid \mathbf{v}_n) = 2 < n$, rezultă că $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sunt liniar dependenţi.

Observaţie 2.14. Propoziţia şi Corolarul de mai sus produc doar un *criteriu de decizie*: vectorii sunt liniar independenţi sau nu; în caz de liniar dependenţă, dacă este nevoie efectiv de relaţia de dependenţă (combinaţia liniară nulă a vectorilor, cu cel puţin un scalar nenul), atunci sunt diverse metode de obţinere a acesteia. Să luăm de exemplu cazul vectorilor din Exemplul 2.13.2): $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Căutăm $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ astfel încât $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$, adică

$$\begin{cases} \alpha_1 + 3\alpha_2 & = 0 \\ -2\alpha_1 - \alpha_2 + \alpha_3 & = 0 \end{cases}$$

Obţinem $\alpha_1 = -3\alpha_2, \alpha_3 = -5\alpha_2$, cu $\alpha_2 \in \mathbb{R}$. Deci relaţia de dependenţă căutată este $-3\mathbf{v}_1 + \mathbf{v}_2 - 5\mathbf{v}_3 = \mathbf{0}$.

2.6 Interpretarea geometrică a liniar (in)dependenţei

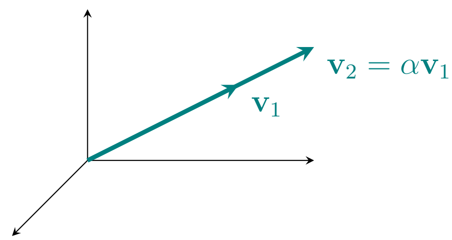
Observaţia 2.9, împreună cu Corolarul 2.12 sunt cheia interpretării geometrice a liniar (in)dependenţei vectorilor.

Vom discuta cazul vectorilor din $\mathbb{R}^3$. Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^3$.

1) $n = 1$: conform Observaţiei 2.9.2, $\mathbf{v}_1$ este liniar independent dacă şi numai dacă este nenul.

2) $n = 2$: vectorii $\mathbf{v}_1, \mathbf{v}_2$ sunt liniar dependenţi dacă cel puţin unul dintre aceştia este nul (Observaţia 2.9.3).

   Să presupunem în continuare că $\mathbf{v}_1, \mathbf{v}_2$ sunt nenuli şi liniar dependenţi. Atunci conform Observaţiei 2.9.4, unul dintre aceştia este un multiplu scalar al celuilalt, de exemplu $\mathbf{v}_2 = \alpha\mathbf{v}_1$, adică sunt **coliniari**.



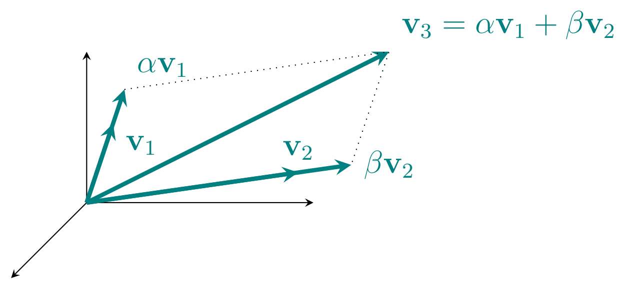
Deci pentru $n = 2$ vectori, avem:

$$\text{liniar dependenţă} \quad \Longleftrightarrow \quad \text{coliniaritate}$$

şi bineînţeles, afirmaţia contrară

$$\text{liniar independenţă} \quad \Longleftrightarrow \quad \text{necoliniaritate}$$

3) $n = 3$: urmărind acelaşi raţionament în cazul precedent, vom discuta doar cazul vectorilor $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^2$ nenuli şi liniar dependenţi. Atunci unul dintre aceştia, de exemplu $\mathbf{v}_3$, este o combinaţie liniară a celorlalţi doi: $\mathbf{v}_3 = \alpha\mathbf{v}_1 + \beta\mathbf{v}_2$. Deducem că $\mathbf{v}_3$ aparţine planului determinat de $\mathbf{v}_1$ şi $\mathbf{v}_2$, adică cei trei vectori sunt **coplanari**.



Deci pentru $n = 3$ vectori, avem:

$$\text{liniar dependenţă} \quad \Longleftrightarrow \quad \text{coplanaritate}$$

şi bineînţeles, afirmaţia contrară

$$\text{liniar independenţă} \quad \Longleftrightarrow \quad \text{necoplanaritate}$$

4) $n \geq 4$: conform Corolarului 2.12, orice n vectori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vor fi liniar dependenţi.

3 Subspaţii vectoriale

Subspaţiu vectorial

Definiţie 3.1. Fie V un $\Bbbk$-spaţiu vectorial. O submulţime *nevidă* $U \subseteq V$ se numeşte **subspaţiu vectorial** al lui V dacă

$$\forall\ \alpha, \beta \in \Bbbk,\ \forall\ \mathbf{v}, \mathbf{w} \in U \implies \alpha\mathbf{v} + \beta\mathbf{w} \in U \tag{3}$$

Exerciţiu 3.2. Arătaţi că o submulţime nevidă $U \subseteq V$ este subspaţiu vectorial dacă şi numai dacă verifică următoarele condiţii:

1) $\forall\ \mathbf{v}, \mathbf{w} \in U \implies \mathbf{v} + \mathbf{w} \in U$;

2) $\forall\ \alpha \in \Bbbk, \mathbf{v} \in U \implies \alpha\mathbf{v} \in U$.

Observaţii 3.3. 1) Fie U un subspaţiu vectorial al $\Bbbk$-spaţiului vectorial V. Întrucât U este o mulţime nevidă, există $\mathbf{u} \in U$. Atunci $\mathbf{0} = \mathbf{0} + \mathbf{0} = 0\mathbf{u} + 0\mathbf{u} \in U$. Deci vectorul nul $\mathbf{0}$ aparţine oricărui subspaţiu U. În particular, submulţimea vidă $\emptyset \subseteq V$ nu poate fi subspaţiu vectorial.

2) $\{\mathbf{0}\}$, V sunt subspaţii ale oricărui spaţiu vectorial V, numite subspaţiile *universale*.

3) Orice subspaţiu vectorial U este stabil la combinaţii liniare:

$$\mathbf{v}_1, \ldots, \mathbf{v}_n \in U,\ \alpha_1, \ldots, \alpha_n \in \Bbbk \implies \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n \in U \tag{4}$$

(prin inducţie după $n \in \mathbb{N}^*$).

Exemplu 3.4. În fizică, proprietatea (4) este cunoscută sub denumirea de *Principiul superpoziţiei*: răspunsul unui sistem dinamic *liniar* la primirea unui număr finit de input-uri se obţine prin suprapunerea output-urilor generate de fiecare input în parte.
În mecanica cuantică şi în Quantum Computing, superpoziţia cuantica reprezintă combinarea tuturor stărilor cuantice ale sistemului (de exemplu, poziţiile posibile ale particulelor subatomice). Am întâlnit în Exemplul 2.6.2 cel mai simplu exemplu de stare cuantică, şi anume **qubitul**: $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ - o superpoziţie (combinaţie liniară) a qubiţilor $|0\rangle$ şi $|1\rangle$, unde $\alpha, \beta \in \mathbb{C}$, $|\alpha|^2 + |\beta|^2 = 1$.

Exemplu 3.5. Vom ilustra conceptul de subspaţiu vectorial printr-un exemplu pentru fiecare dintre spaţiile vectoriale prezentate în Exemplul 2.3. Cazul vectorilor coloană va fi discutat separat în secţiunea următoare.

1) Pentru $n \in \mathbb{N}$, fie $\Bbbk_n[X]$ mulţimea polinoamelor de grad cel mult n cu coeficienţi din corpul $\Bbbk$. Se observă uşor că $\Bbbk_n[X]$ este un subspaţiu vectorial al lui $\Bbbk[X]$.

2) Fie $\mathsf{Sym}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A^t = A\}$ şi $\mathsf{AntiSym}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A^t = -A\}$ mulţimea matricelor pătratice reale simetrice, respectiv antisimetrice. Atunci $\mathsf{Sym}_n(\mathbb{R})$ şi $\mathsf{AntiSym}_n(\mathbb{R})$ sunt subspaţii vectoriale în $\mathcal{M}_n(\mathbb{R})$. Vom întâlni în acest curs şi subspaţiul vectorial al matricelor complexe **hermitice**: $\mathsf{Herm}_n = \{A \in \mathcal{M}_n(\mathbb{C}) \mid A^t = \overline{A}\}$, unde $\overline{A}$ se obţine din A prin conjugarea (complexă) a tuturor componentelor sale.

3) În exemplul 2.3.3, fie $M = \mathbb{N}$ şi $\Bbbk = \mathbb{R}$. Atunci $\mathbb{R}^{\mathbb{N}}$ este spaţiul vectorial al şirurilor reale. Exemple importante de subspaţii vectoriale sunt:

   (a) Şirurile de numere reale convergente $\mathsf{c} = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \exists a \in \mathbb{R}\ a_n \to a\}$

   (b) Şirurile de numere reale absolut sumabile $\ell^1 = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_n |a_n| < \infty\}$

   (c) Şirurile de numere reale de pătrat absolut sumabil $\ell^2 = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_n |a_n|^2 < \infty\}$

   (d) Şirurile de numere reale mărginite $\ell^\infty = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_n |a_n| < \infty\}$[10]

Exerciţiu 3.6. Verificaţi că toate exemplele menţionate mai sus sunt într-adevăr subspaţii vectoriale în spaţiile vectoriale indicate.

3.1 Subspaţii vectoriale: operaţii şi relaţii

Observaţie 3.7. Fie U_1, U_2 subspaţii vectoriale ale unui spaţiu vectorial V. În particular, acestea sunt submulţimi ale lui V, deci putem efectua cu acestea operaţiile uzuale de reuniune, intersecţie, etc. Sunt mulţimile rezultate stabile în raport cu cele două operaţii de adunare şi multiplicare cu scalari (adică sunt subspaţii vectoriale)? Să observăm următoarele:

1) $U_1 \cap U_2$ este subspaţiu vectorial.

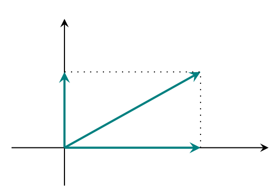
2) $U_1 \cup U_2$ este subspaţiu vectorial dacă şi numai dacă $U_1 \subseteq U_2$ sau dacă $U_2 \subseteq U_1$.

Demonstraţie. 1) Fie $\mathbf{v}, \mathbf{w} \in U_1 \cap U_2$ şi $\alpha, \beta \in \Bbbk$. Atunci $\mathbf{v}, \mathbf{w} \in U_1$ şi $\mathbf{v}, \mathbf{w} \in U_2$. Deoarece U_1 este subspaţiu, rezultă $\alpha\mathbf{v} + \beta\mathbf{w} \in U_1$. Analog, $\alpha\mathbf{v} + \beta\mathbf{w} \in U_2$, deci $\alpha\mathbf{v} + \beta\mathbf{w} \in U_1 \cap U_2$.

2) Dacă $U_1 \subseteq U_2$, rezultă $U_1 \cup U_2 = U_2$ care este subspaţiu vectorial. Analog pentru cazul $U_2 \subseteq U_1$. Să presupunem acum că $U_1 \cup U_2$ este subspaţiu vectorial, dar $U_1 \not\subseteq U_2$ şi şi nici $U_2 \not\subseteq U_1$. Atunci există $\mathbf{v}_1 \in U_1, \mathbf{v}_1 \notin U_2$ şi $\mathbf{v}_2 \in U_2, \mathbf{v}_2 \notin U_1$. Acum $\mathbf{v}_1 \in U_1 \subseteq U_1 \cup U_2$ şi analog $\mathbf{v}_2 \in U_2 \subseteq U_1 \cup U_2$, deci ambii vectori aparţin subspaţiului $U_1 \cup U_2$. În particular, $\mathbf{v}_1 + \mathbf{v}_2 \in U_1 \cup U_2$. Atunci $\mathbf{v}_1 + \mathbf{v}_2 \in U_1$ sau $\mathbf{v}_1 + \mathbf{v}_2 \in U_2$. Să analizăm primul caz. Deoarece U_1 este subspaţiu şi $\mathbf{v}_1 \in U_1$, rezultă $-\mathbf{v}_1 = (-1)\mathbf{v}_1 \in U_1$. Deducem $\mathbf{v}_2 = (\mathbf{v}_1 + \mathbf{v}_2) + (-\mathbf{v}_1) \in U_1$, contradicţie. La fel, situaţia $\mathbf{v}_1 + \mathbf{v}_2 \in U_2$ conduce la o contradicţie.

□

Exemplu 3.8. Rezultatele de mai sus referitoare la reuniune sunt mult mai uşor de observat geometric. Să luăm ca exemplu în $\mathbb{R}^2$, subspaţiile $U_1 = \{\begin{pmatrix} x \\ y \end{pmatrix} \mid x \in \mathbb{R}, y = 0\}$ şi $U_2 = \{\begin{pmatrix} x \\ y \end{pmatrix} \mid x = 0, y \in \mathbb{R}\}$[11] Acestea corespund în planul cartezian celor două axe.



Atunci reuniunea $U_1 \cup U_2$ nu formează un subspaţiu: suma dintre un vector din U_1 şi unul din U_2 nu este în general un vector aparţinând reuniunii $U_1 \cup U_2$.

[10] Notaţiile $\mathsf{c}, \ell^1, \ell^\infty$ sunt notaţiile standard ale acestor subspaţii vectoriale în Matematică şi Computer Science.

[11] Verificaţi că U_1 şi U_2 sunt subspaţii vectoriale.

Având în vedere cele de mai sus, este natural să ne punem întrebarea dacă există un subspaţiu vectorial care să includă cele două subspaţii şi să fie cel mai mic subspaţiu cu această proprietate (în raport cu relaţia de incluziune). Răspunsul este pozitiv şi detaliat în cele ce urmează.

Suma a două subspaţii vectoriale

Definiţie 3.9. Fie U_1, U_2 subspaţii ale unui $\Bbbk$-spaţiu vectorial V. Mulţimea

$$\{\mathbf{u}_1 + \mathbf{u}_2 \mid \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}$$

se numeşte **suma** subspaţiilor U_1 şi U_2 şi se notează $U_1 + U_2$.

Exemplu 3.10. Fie $V = \mathcal{M}_n(\mathbb{R})$ spaţiul vectorial real al matricelor pătratice cu n linii şi n coloane, şi $\mathsf{Sym}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A = A^t\}$, $\mathsf{AntiSym}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) \mid A = -A^t\}$ subspaţiile vectoriale ale matricelor simetrice, respectiv antisimetrice (Exemplul 3.5.2). Atunci

$$\mathsf{Sym}_n(\mathbb{R}) + \mathsf{AntiSym}_n(\mathbb{R}) = \mathcal{M}_n(\mathbb{R})$$

Într-adevăr, fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice arbitrară. Observăm că A se poate scrie ca suma dintre o matrice simetrică şi o matrice antisimetrică $A = A_1 + A_2$, unde

$$A_1 = \frac{1}{2}(A + A^t) \quad A_2 = \frac{1}{2}(A - A^t)$$

Să observăm în plus că $\mathsf{Sym}_n(\mathbb{R}) \cap \mathsf{AntiSym}_n(\mathbb{R}) = \{0\}$ – spunem în acest caz că suma e directă şi scriem $\mathsf{Sym}_n(\mathbb{R}) \oplus \mathsf{AntiSym}_n(\mathbb{R}) = \mathcal{M}_n(\mathbb{R})$ (a se vedea Corolarul 3.12).

Suma a două subspaţii vectoriale este un subspaţiu vectorial

Propoziţie 3.11. Fie U_1, U_2 subspaţii ale unui $\Bbbk$-spaţiu vectorial V. Atunci $U_1 + U_2$ este subspaţiu vectorial şi este cel mai mic subspaţiu vectorial care conţine pe U_1 şi pe U_2.

Demonstraţie. Fie $\alpha, \beta \in \Bbbk$ şi $\mathbf{v}, \mathbf{w} \in U_1 + U_2$. Atunci există $\mathbf{v}_1, \mathbf{w}_1 \in U_1$ şi $\mathbf{v}_2, \mathbf{w}_2 \in U_2$ astfel încât $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ şi $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$. Atunci

$$\alpha\mathbf{v} + \beta\mathbf{w} = \alpha(\mathbf{v}_1 + \mathbf{v}_2) + \beta(\mathbf{w}_1 + \mathbf{w}_2) = \underbrace{(\alpha_b v_1 + \beta\mathbf{w}_1)}_{\in U_1} + \underbrace{(\alpha_b v_2 + \beta\mathbf{w}_2 z)}_{\in U_2} \in U_1 + U_2$$

Cum $\alpha, \beta \in \Bbbk$ şi $\mathbf{v}, \mathbf{w} \in U_1 + U_2$ au fost arbitrari, rezultă că $U_1 + U_2$ este subspaţiu vectorial al lui V.

Pentru a proba incluziunea $U_1 \subseteq U_1 + U_2$, să observăm că vectorul nul aparţine oricărui subspaţiu; atunci dacă $\mathbf{v}$ este un vector arbitrar din U, avem $\mathbf{v} = \mathbf{v} + \mathbf{0} \in U_1 + U_2$. Analog rezultă şi $U_2 \subseteq U_1 + U_2$.

Deci $U_1 + U_2$ este un subspaţiu vectorial care conţine atât pe U_1 cât şi pe U_2. Să arătăm acum că este cel mai mic subspaţiu cu această proprietate.

Fie $W \subseteq V$ un subspaţiu vectorial cu proprietatea că $U_1, U_2 \subseteq W$. Fie $\mathbf{v} \in U_1 + U_2$ un vector arbitrar. Atunci $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, cu $\mathbf{v}_1 \in U_1, \mathbf{v}_2 \in U_2$. Rezultă $\mathbf{v}_1 \in U_1 \subseteq W$, $\mathbf{v}_2 \in U_2 \subseteq W$, şi cum W este subspaţiu, $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 \in W$. Cum $\mathbf{v}$ a fost ales arbitrar, obţinem $U_1 + U_2 \subseteq W$. $\square$

Sumă directă

Corolar 3.12. Dacă $U_1 \cap U_2 = \{\mathbf{0}\}$, atunci reprezentarea unui vector din $U_1 + U_2$ ca sumă $\mathbf{v}_1 + \mathbf{v}_2$, unde $\mathbf{v}_1 \in U_1, \mathbf{v}_2 \in U_2$, este unică.
În acest caz spunem că suma subspaţiilor U_1 şi U_2 este **directă** şi scriem $U_1 \oplus U_2$.
Dacă în plus $U_1 + U_2 = V$, subspaţiile U_1, U_2 se numesc **complementare** (mai spunem că U_2 este un complement al lui U_1 şi invers).

Demonstraţie. Exerciţiu. □

Exerciţiu 3.13. Fie $U_1 = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x + 2y - z = 0\}$, $U_2 = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x - 2y + 3z = 0\}$, $U_3 = \{\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid x = y = 3z\}$.

1) Verificaţi că U_1, U_2, U_3 sunt subspaţii în $\mathbb{R}^3$.
2) Calculaţi intersecţia şi suma fiecărei perechi de subspaţii de mai sus.
3) Deduceţi că complementul unui subspaţiu nu este unic (mai târziu, în secţiunile următoare, vom vedea că putem întotdeauna găsi un complement pentru fiecare subspaţiu vectorial).

3.2 Subspaţiile vectoriale asociate unei matrice.

Vom introduce în această secţiune două subspaţii vectoriale asociate unei matrice arbitrare. Ulterior, vom vedea că orice subspaţiu vectorial al spaţiului $\mathbb{k}^n$ poate fi reprezentat sub această formă.

Nucleul unei matrice

Definiţie 3.14. Să considerăm o matrice $A \in \mathcal{M}_{m,n}(\mathbb{k})$. Mulţimea

$$\mathsf{Ker}(A) = \{\mathbf{x} \in \mathbb{k}^n \mid A\mathbf{x} = \mathbf{0}_m\} \tag{5}$$

se numeşte **nucleul** matricei A.

Conform (5), un vector $\mathbf{x}$ cu n componente aparţine nucleului matricei A (n este numărul de coloane ale matricei!) dacă şi numai dacă $A\mathbf{x} = \mathbf{0}_m$, unde m este numărul de linii ale matricei. Dacă notăm $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ şi $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, atunci relaţia precedentă este echivalentă cu

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0 \end{cases}$$

Deci $\mathsf{Ker}(A)$ este *mulţimea soluţiilor sistemului* **omogen** asociat matricei A.

Nucleul unei matrice este subspaţiu vectorial

Propoziţie 3.15. $\mathsf{Ker}(A)$ este subspaţiu vectorial în $\mathbb{k}^n$.

Demonstraţie. Fie $\mathbf{x}, \mathbf{y} \in \mathsf{Ker}(A)$ şi $\alpha, \beta \in \mathbb{k}$. Atunci $A(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha(A\mathbf{x}) + \beta(A\mathbf{y}) = \alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0}$, deci $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathsf{Ker}(A)$. □

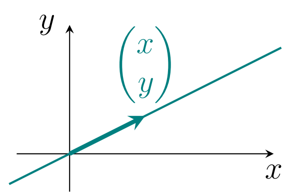
Exemplu 3.16. 1) Fie $A = \mathbf{0}_{m,n} \in \mathcal{M}_{m,n}(\mathbb{k})$ matricea nulă. Atunci $\mathsf{Ker}(A) = \mathbb{k}^n$.

2) Dacă A este matricea identică $A = I_n \in \mathcal{M}_n(\mathbb{k})$, atunci nucleul său este trivial: $\mathsf{Ker}(A) = \{\mathbf{0}_n\}$. Mai general, dacă A este o matrice pătratică inversabilă, atunci $\mathsf{Ker}(A) = \{\mathbf{0}_n\}$.

3) Fie $A = \begin{pmatrix} a & b \end{pmatrix} \in \mathcal{M}_{1,2}(\mathbb{R})$ o matrice nenulă cu o singură linie. Atunci nucleul său este

$$\mathsf{Ker}(A) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid ax + by = 0 \right\}$$

Interpretare geometrică: $ax + by = 0$ este ecuaţia unei **drepte** în plan, care trece prin origine. Deci nucleul matricei A conţine toţi vectorii ce pleacă din origine şi au vârful pe dreapta de ecuaţie $ax + by = 0$. Se obţine astfel o corespondenţă bijectivă între mulţimea subspaţiilor vectoriale de tip $\mathsf{Ker}(A)$, unde $A \in \mathcal{M}_{1,2}(\mathbb{R})$, $A \neq \mathbf{0}$, şi mulţimea dreptelor din plan care trec prin origine. Ulterior, pentru simplificare, vom identifica dreapta respectivă care trece prin origine cu subspaţiul vectorial asociat.



4) Analog, pentru matricea nenulă $A = \begin{pmatrix} a & b & c \end{pmatrix} \in \mathcal{M}_{1,3}(\mathbb{R})$, se obţine

$$\mathsf{Ker}(A) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid ax + by + cz = 0 \right\}$$

Geometric, $ax + by + cz = 0$ este ecuaţia unui **plan** în spaţiu care trece prin origine.

5) Fie $A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R})$. Atunci

$$\mathsf{Ker}(A) = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{cases} ax + by + cz = 0 \\ a'x + b'y + c'z = 0 \end{cases} \right\}$$

Interpretare geometrică: Dacă $\mathsf{rang}(A) = 2$, obţinem ecuaţiile unei **drepte** din spaţiu (intersecţia a două plane) care trece prin origine. Dar dacă $\mathsf{rang}(A) \leq 1$?

Imaginea unei matrice

Definiţie 3.17. Fie $A \in \mathcal{M}_{m,n}(\Bbbk)$ o matrice. Mulţimea

$$\mathsf{Im}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \Bbbk^n\} = \{\mathbf{b} \in \Bbbk^m \mid \exists \mathbf{x} \in \Bbbk^n \ . \ A\mathbf{x} = \mathbf{b}\}$$

se numeşte **imaginea** matricei A.

Observaţie 3.18. Fie $\mathbf{b} \in \Bbbk^m$ un vector arbitrar. Notând $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, să observăm că $\mathbf{b} \in \mathsf{Im}(A)$ dacă şi numai dacă există $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \Bbbk^n$ astfel încât $A\mathbf{x} = \mathbf{b}$. Scriind pe componente, obţinem sistemul

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

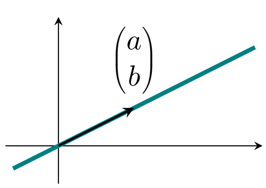
Deci un vector $\mathbf{b}$ aparţine imaginii matricei A dacă şi numai dacă sistemul de ecuaţii liniare asociat $A\mathbf{x} = \mathbf{b}$ este compatibil (are soluţie).

Imaginea unei matrice este subspaţiu vectorial

Propoziţie 3.19. $\mathsf{Im}(A)$ este subspaţiu vectorial în $\Bbbk^m$.

Demonstraţie. Exerciţiu. □

Exemplu 3.20. Fie $A = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathcal{M}_{2,1}(\mathbb{R})$ o matrice nenulă. Atunci $\mathsf{Im}(A) = \left\{ \begin{pmatrix} ax \\ bx \end{pmatrix} \mid x \in \mathbb{R} \right\}$ este mulţimea vectorilor coliniari cu vectorul $\begin{pmatrix} a \\ b \end{pmatrix}$:



Interpretare geometrică: o **dreaptă** care trece prin origine.

Observaţie 3.21. Am obţinut pentru o construcţie geometrică (dreapta care trece prin origine) două reprezentări algebrice: ca nucleul unei matrice, respectiv ca imaginea unei matrice. Vom vedea în Secţiunile următoare că această dualitate a reprezentării se poate extinde pentru toate subspaţiile vectoriale ale $\Bbbk^n$.

3.3 Subspaţiul generat de o mulţime finită de vectori

În secţiunea precendentă, am observat că dreptele şi planele care trec prin origine se pot identifica cu subspaţii vectoriale din $\mathbb{R}^2$, respectiv din $\mathbb{R}^3$. Acum vom vedea cum slogane de genul "două drepte concurente determină un plan" admit o interpretare algebrică.

Subspaţiul generat de o mulţime finită de vectori

Propoziţie 3.22. Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vectori în V. Mulţimea tuturor combinaţiilor liniare

$$\{\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n \mid \alpha_1, \ldots \alpha_n \in \mathbb{k}\}$$

formează un subspaţiu vectorial în V, numit **subspaţiul generat** de vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ şi notat $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

Demonstraţie. Exerciţiu. □

În vederea familiarizării cu noţiunea introdusă mai sus, propunem spre rezolvare următorul exerciţiu:

Exerciţiu 3.23. Să se verifice relaţiile de mai jos:

1) $\mathsf{Sp}\{\mathbf{0}\} = \{\mathbf{0}\}$
2) $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathsf{Sp}\{\mathbf{v}_1, \ldots \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k}\}$
3) $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{0}\} = \mathsf{Sp}\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$
4) $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \alpha\mathbf{v}_n\} = \mathsf{Sp}\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$
5) $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n-1} + \mathbf{v}_n\} = \mathsf{Sp}\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$

unde $\mathbf{v}_1, \ldots, \mathbf{v}_n, \ldots \in V$ şi $\alpha \in \mathbb{k}$.

3.4 Combinaţii liniare şi imaginea unei matrice

Orice mulţime finită de vectori coloană $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{k}^m$ determină o matrice $A = (\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \in \mathcal{M}_{m,n}(\mathbb{k})$ şi reciproc. Vom vedea acum ce legătură există între imagine unei matrice (reamintim că aceasta este subspaţiul termenilor liberi $\mathbf{b}$ pentru care sistemul $A\mathbf{x} = \mathbf{b}$ este compatibil) şi coloanele matricei A.

Imaginea unei matrice este generată de coloanele acesteia

Propoziţie 3.24. Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{k}^m$ şi $A = (\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \in \mathcal{M}_{m,n}(\mathbb{k})$. Atunci imaginea matricei A coincide cu subspaţiul generat de vectorii coloană $\mathbf{v}_1, \ldots, \mathbf{v}_n$:

$$\mathsf{Im}(A) = \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

Demonstraţie. Cu notaţiile din Definiţia 3.17, avem $\mathbf{b} \in \mathsf{Im}(A) \iff \exists \mathbf{x} \in \Bbbk^n . \mathbf{b} = A\mathbf{x}$. Dacă notăm $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ atunci $\mathbf{b} = A\mathbf{x} = (\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{v}_1 + \ldots + x_n\mathbf{v}_n$, adică $\mathbf{b} \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. □

Vectori generatori şi rang

Corolar 3.25. 1) Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^m$. Atunci $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \Bbbk^m$ (spunem în acest caz că vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ **generează** spaţiul vectorial $\Bbbk^m$) dacă şi numai dacă $m \leq n$ şi $\mathsf{rang}\,(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) = m$.

2) *(cazul $m = n$)* Fie $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \Bbbk^n$. Atunci $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \Bbbk^n$ dacă şi numai dacă $\mathsf{det}\,(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \neq 0$.

Demonstraţie. 1) "⇒" Fie vectorii $\mathbf{e}_1, \ldots, \mathbf{e}_m \in \Bbbk^m$. Cum $\Bbbk^m = \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, există scalarii $\alpha_{11}, \ldots, \alpha_{mn} \in \Bbbk$ astfel încât

$$\mathbf{e}_i = \sum_{j=1}^{n} \alpha_{ji}\mathbf{v}_j \ \forall i = \overline{1, m}$$

Dacă notăm cu $A \in \mathcal{M}_{m,n}(\Bbbk)$ matricea ale cărui coloane sunt $\mathbf{v}_1, \ldots, \mathbf{v}_n$ şi cu $B = (\alpha_{ji})_{i,j} \in \mathcal{M}_{n,m}(\Bbbk)$ matricea scalarilor, atunci relaţia de mai sus se scrie matriceal $I_m = AB$:

$$\underbrace{\begin{array}{|c|c|c|} \mathbf{e}_1 & \ldots & \mathbf{e}_m \end{array}}_{I_m} = \underbrace{\begin{array}{|c|c|c|} \mathbf{v}_1 & \ldots & \mathbf{v}_n \end{array}}_{A} \underbrace{\begin{array}{|ccc|} \alpha_{11} & \ldots & \alpha_{m1} \\ \vdots & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{nm} \end{array}}_{B}$$

În particular,

$$m = \mathsf{rang} I_m \leq \min(\mathsf{rang} A, \mathsf{rang} B) \leq \mathsf{rang} A \leq \min(m, n)$$

de unde obţinem $m \leq n$ şi $\mathsf{rang} A = m$.

"⇐" Deoarece $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \Bbbk^n$, va trebui să verificăm doar incluziunea inversă. Fie atunci $\mathbf{v} \in \Bbbk^n$. Atunci $\mathbf{v} \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \mathsf{Im}(A)$ dacă şi numai dacă sistemul $A\mathbf{x} = \mathbf{v}$ este compatibil. Dar $\mathsf{rang} A = m \leq n$ şi rangul matricei extinse $(A \mid \mathbf{v})$ este cel mult m (de ce?), deci $\mathsf{rang} A = \mathsf{rang}\,(A \mid \mathbf{v}) = m$, şi conform teoremei Kronecker-Cappelli sistemul $A\mathbf{x} = \mathbf{v}$ este compatibil.

2) Clar.

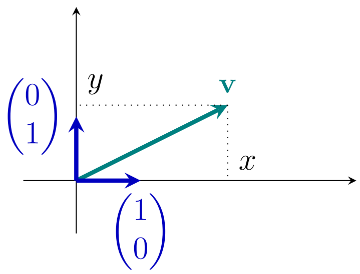
□

Exerciţiu 3.26. Comparaţi corolarul precedent cu Corolarul 2.12. Ce remarcaţi?

4 Baze, dimensiune şi coordonate

Am introdus vectorii (coloană) ca un caz particular de matrice şi i-am interpretat geometric. Orice asemenea vector este unic determinat de componentele sale relativ la un reper (se poate scrie ca o combinaţie liniară a vectorilor unitari (1)). Apare naturală întrebarea: cum se schimbă vectorii dacă se modifică reperul? Vom vedea că nu se întâmplă nimic cu vectorii în sine, ci doar cu componentele lor.

Să considerăm cazul unui vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ din planul cartezian:



Se observă uşor că $\mathbf{v}$ se poate scrie ca o combinaţie liniară $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. În plus, vectorii $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sunt *necoliniari (liniar independenţi)*. Vom vedea în această Secţiune cum putem generaliza rolul jucat de aceşti vectori în $\mathbb{R}^2$ pentru un spaţiu vectorial arbitrar V.

4.1 Bază a unui spaţiu vectorial

Bază

Definiţie 4.1. Fie V un $\Bbbk$-spaţiu vectorial. O mulţime $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ este o **bază** a spaţiului vectorial V dacă sunt verificate următoarele condiţii:

1) Vectorii $\mathbf{v}_1, \dots, \mathbf{v}_n$ sunt liniar independenţi.

2) $\mathsf{Sp}\{\mathbf{v}_1, \dots, \mathbf{v}_n\} = V$ (în acest caz spunem că vectorii $\mathbf{v}_1, \dots, \mathbf{v}_n$ **generează** spaţiul vectorial V, sau că formează **un sistem de generatori**).

Observaţie 4.2. Definiţia se poate extinde şi pentru baze infinite, astfel: spunem că o mulţime arbitrară de vectori $\mathcal{B} \subseteq V$ este o bază dacă vectorii din $\mathcal{B}$ sunt liniar independenţi (orice submulţime *finită* de vectori din $\mathcal{B}$ este liniar independentă) şi generează pe V (orice vector din V se poate scrie ca o combinaţie liniară *finită* de vectori din $\mathcal{B}$).

În cadrul cursului vom considera doar cazul bazelor finite (având un număr finit de vectori), menţionând exemple cu spaţii vectoriale cu baze infinite doar pentru a sublinia anumite proprietăţi sau aplicaţii ale acestora.

Existenţa bazei

Teoremă 4.3. Orice $\Bbbk$-spaţiu vectorial V nenul are cel puţin o bază.

Demonstraţia acestei teoreme depăşeste nivelul cursului şi o vom omite.[12].

Prin convenţie, spaţiul vectorial nul $\{\mathbf{0}\}$ are bază vidă $\emptyset$.

Exemplu 4.4 (**Baze canonice**). În fiecare dintre cazurile următoare, vom pune în evidenţă în mod *natural* o bază, numită **baza canonică** a spaţiului vectorial respectiv:

1) Fie $V = \mathbb{k}^n$. În Observaţia 2.5.1, am văzut că orice vector din $\mathbb{k}^n$ se poate scrie ca o combinaţie liniară a vectorilor unitari $\mathbf{e}_1, \ldots, \mathbf{e}_n$; cu alte cuvinte, aceşti vectori *generează* spaţiul $\mathbb{k}^n$. Ulterior, am arătat că sunt şi liniar independenţi (Exemplul 2.13.1). Deci vectorii

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

formează o bază, pe care o vom numi, după cum am menţionat mai sus, *baza canonică* a spaţiului vectorial $\mathbb{k}^n$.

În particular, pentru $n = 1$, rezultă că $\mathbb{k}$-spaţiul vectorial $\mathbb{k}$ (a se vedea Exemplul 2.3.**??**) admite baza $\{1\}$.

2) Plecând de la observaţia că orice polinom $P \in \mathbb{k}[X]$ se poate scrie $P(X) = a_0 1 + a_1 X + a_2 X^2 + \ldots + a_n X^n$, unde $n = \mathsf{grad} P$, deducem că monoamele $1, X, X^2, \ldots, X^n, \ldots$ generează spaţiul vectorial $V = \mathbb{k}[X]$. Întrucât sunt şi liniar independente (Exemplul 2.13.2), rezultă că acestea formează o bază (infinită!) în $\mathbb{k}[X]$.

3) În spaţiul vectorial $V = \mathcal{M}_{m,n}(\mathbb{k})$, matricele

$$E_{ij} = \begin{array}{c} \\ \\ \\ i \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} \ \ & & j & & \ \ \end{array} \\ \begin{pmatrix} 0 & \ldots & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 1 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{pmatrix} \end{array} \quad i = \overline{1,m}, j = \overline{1,n}$$

formează o bază. Pentru a verifica aceasta, să observăm că orice matrice $A \in \mathcal{M}_{m,n}(\mathbb{k})$, $A = (a_{ij})_{i=\overline{1,m}, j=\overline{1,n}}$ se scrie

$$A = \sum_{i=\overline{1,m}, j=\overline{1,n}} a_{ij} E_{ij}$$

ceea ce înseamnă că matricele E_{ij} generează spaţiul vectorial al matricelor. Lăsăm liniar indepedenţa acestor matrice ca exerciţiu.

Exerciţiu 4.5. Arătaţi că vectorii următori formează o bază în spaţiul vectorial indicat.

1) $\mathbb{k} = \mathbb{R}$, $V = \mathbb{R}^3$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$. În particular, deduceţi de aici că într-un spaţiu vectorial, **baza nu este unică** (de exemplu, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ este de asemenea o bază în $\mathbb{R}^3$)

[12]Se poate arăta că demonstraţia este echivalentă cu validitatea Axiomei Alegerii din teoria mulţimilor.

2) $\Bbbk = \mathbb{C}$, $V = \mathbb{C}^2$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ \mathsf{i} \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1-\mathsf{i} \\ 0 \end{pmatrix}$.

Formează $\mathbf{v}_1$ şi $\mathbf{v}_2$ o bază în $\mathbb{C}^2$ ca spaţiu vectorial peste corpul numerelor reale $\Bbbk = \mathbb{R}$?

3) $\Bbbk = \mathbb{R}$, $V = \mathbb{R}^4$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$.

Această bază se numeşte baza Haar, şi se întâlneşte în teoria undinelor (în limba engleză: *wavelet*) [6].

4.2 Baze şi coordonate

Criteriul bazei

Propoziţie 4.6. Mulţimea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ este bază în $\Bbbk$-spaţiul vectorial V dacă şi numai dacă pentru orice vector $\mathbf{v} \in V$, **există şi sunt unici** scalarii $\alpha_1, \ldots, \alpha_n \in \Bbbk$ astfel încât

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$$

Scalarii $\alpha_1, \ldots, \alpha_n$ se numesc **coordonatele** vectorului $\mathbf{v}$ în baza $\mathcal{B}$.

Demonstraţie. "$\Rightarrow$" Fie $\mathbf{v} \in V$. Deoarece $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = V$, rezultă că există scalarii $\alpha_1, \ldots, \alpha_n \in \Bbbk$ astfel încât $\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$. Unicitatea scalarilor rezultă din liniar independenţa vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_n$, utilizând Observaţia 2.9.5.

"$\Leftarrow$" Pasul I: verificăm liniar independenţa vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_n$: fie $\alpha_1, \ldots, \alpha_n \in \Bbbk$ astfel încât $\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}$. Dar vectorul nul se mai poate scrie şi $\mathbf{0} = 0\mathbf{v}_1 + \ldots + 0\mathbf{v}_n$, deci din ipoteza de unicitate a scalarilor obţinem $\alpha_1 = \ldots = \alpha_n = 0$. Pasul II: vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ genereazaă V, adică $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = V$. Aceasta rezultă din ipoteză. □

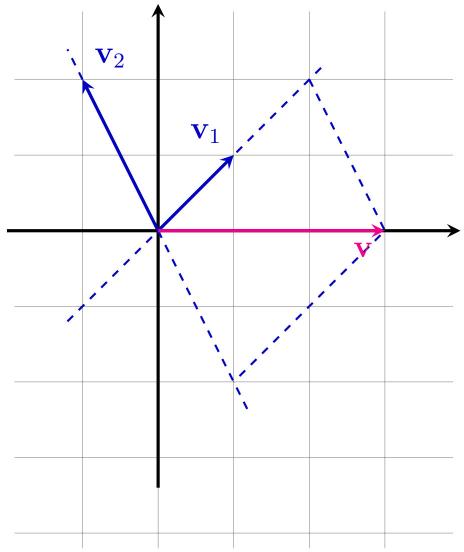
Vom nota cu $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \Bbbk^n$ vectorul coloană al coordonatelor vectorului $\mathbf{v}$ (presupunem aleasă o ordine a vectorilor din bază). Deci cu ajutorul *Criteriului bazei* putem stabili o corespondenţă bijectivă

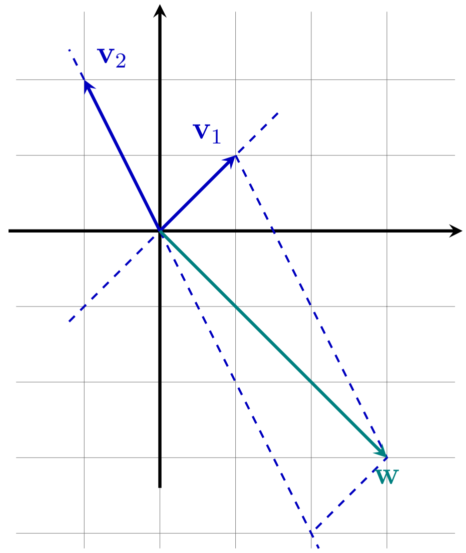
$$\text{vector (abstract) } \mathbf{v} \in V \qquad \Longleftrightarrow \qquad \text{vector coloană } [\mathbf{v}]_{\mathcal{B}} \in \Bbbk^n$$

Exerciţiu 4.7. 1) Fie $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ şi $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$. Arătaţi că $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ este o bază în $\mathbb{R}^2$.

2) Aflaţi coordonatele vectorului $\mathbf{v} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ în raport cu baza $\mathcal{B}$.

3) Determinaţi vectorul $\mathbf{w} \in \mathbb{R}^2$ care în baza $\mathcal{B}$ are coordonatele $[\mathbf{w}]_{\mathcal{B}} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.





Observaţie 4.8. Criteriul bazei are următoarele consecinţe importante, care permit reducerea raţionamentelor (şi calculelor) la cazul vectorilor coloană: fie V un $\Bbbk$-spaţiu vectorial şi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ o bază în V. Atunci:

1) Doi vectori sunt **egali** $\mathbf{v} = \mathbf{w}$ dacă şi numai dacă au **aceleaşi coordonate** $[\mathbf{v}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{B}}$ în raport cu baza $\mathcal{B}$.

2) Coordonatele vectorului nul $\mathbf{0}$ în orice bază $\mathcal{B}$ sunt **nule**: $[\mathbf{0}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. În particular, un vector este nul $\mathbf{v} = \mathbf{0} \iff [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

3) Coordonatele vectorilor din baza $\mathcal{B}$ în raport cu propria bază $\mathcal{B}$ sunt:
$$[\mathbf{v}_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ [\mathbf{v}_2]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \ldots, \ [\mathbf{v}_n]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

4) Coordonatele sumei a doi vectori se obţin însumând coordonatele celor doi vectori $[\mathbf{v}+\mathbf{w}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$; analog, coordonatele multiplicării unui vector cu un scalar sunt $[\alpha\mathbf{v}]_{\mathcal{B}} = \alpha[\mathbf{v}]_{\mathcal{B}}$. Deci operaţiile abstracte din V se reduc la operaţiile uzuale de adunare şi multiplicare cu scalari ale vectorilor coloană.

Exerciţiu 4.9. Să se determine coordonatele vectorului $\mathbf{v} = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 3 \end{pmatrix}$ în raport cu baza canonică din $\mathbb{R}^4$; aceeaşi cerinţă pentru polinomul $P(X) = 1 - 2X + 3X^3$ în raport cu baza canonică din $\mathbb{R}_3[X]$, respectiv pentru matricea $A = \begin{pmatrix} 1 & -2 \\ 0 & 3 \end{pmatrix}$ în $\mathcal{M}_2(\mathbb{R})$. Ce observaţi?

4.3 Bază/liniar independenţă/generatori

Am văzut în Corolarul 2.12 că numărul maxim posibil de vectori liniar independenţi în $\mathbb{k}^n$ este n. Aceasta nu este o coincidenţă – există o legătură strânsă între numărul vectorilor dintr-o bază şi numărul maxim de vectori liniar independenţi, pe care o vom discuta în această secţiune.

Numărul de vectori liniar independenţi nu depăşeşte numărul de vectori dintr-o bază

Propoziţie 4.10. Presupunem că în $\mathbb{k}$-spaţiul vectorial V există o bază cu n vectori. Atunci orice mulţime de vectori liniar independenţi conţine **cel mult** n **vectori**.

Demonstraţie. Fie vectorii $\mathbf{w}_1, \dots, \mathbf{w}_m \in V$, cu $m > n$. Vom arăta că $\mathbf{w}_1, \dots, \mathbf{w}_m$ sunt liniar dependenţi.

Fie $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ o bază cu n vectori în V. Exprimând fiecare dintre vectorii $\mathbf{w}_1, \dots, \mathbf{w}_m$ sub forma unei combinaţii liniare a vectorilor din bază, obţinem

$$\begin{aligned} &\mathbf{w}_1 = \alpha_{11}\mathbf{v}_1 + \ldots + \alpha_{1n}\mathbf{v}_n \\ &\vdots \\ &\mathbf{w}_m = \alpha_{m1}\mathbf{v}_1 + \ldots + \alpha_{mn}\mathbf{v}_n \end{aligned}$$

unde $\alpha_{11}, \dots, \alpha_{mn} \in \mathbb{k}$.

Fie acum $\beta_, \dots, \beta_m \in \mathbb{k}$ astfel încât

$$\beta_1\mathbf{w}_1 + \ldots + \beta_m\mathbf{w}_m = \mathbf{0}$$

Substituind în relaţia de mai sus $\mathbf{w}_1, \dots, \mathbf{w}_m$, obţinem

$$\beta_1(\alpha_{11}\mathbf{v}_1 + \ldots + \alpha_{1n}\mathbf{v}_n) + \ldots + \beta_m(\alpha_{m1}\mathbf{v}_1 + \ldots + \alpha_{mn}\mathbf{v}_n) = \mathbf{0}$$

Regrupând termenii, relaţia devine

$$(\beta_1\alpha_{11} + \ldots + \beta_m\alpha_{m1})\mathbf{v}_1 + \ldots + (\beta_1\alpha_{1n} + \ldots + \beta_m\alpha_{mn})\mathbf{v}_n = \mathbf{0}$$

Dar $\mathbf{v}_1, \dots, \mathbf{v}_n$ formează o bază, în particular sunt liniar independenţi, deci

$$\begin{cases} \beta_1\alpha_{11} + \ldots + \beta_m\alpha_{m1} = 0 \\ \vdots \\ \beta_1\alpha_{1n} + \ldots + \beta_m\alpha_{mn} = 0 \end{cases}$$

Am obţinut un sistem liniar omogen cu n ecuaţii şi m necunoscute $\beta_1, \dots, \beta_m$. Deoarece $m > n$, sistemul admite şi soluţii nebanale. Deci $\mathbf{w}_1, \dots, \mathbf{w}_m$ sunt liniar dependenţi. $\square$

Toate bazele au acelaşi număr de elemente

Teoremă 4.11. Dacă $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ sunt baze în V, atunci $m = n$.

Demonstraţie. $\mathcal{B}$ este o bază cu n vectori, iar $\mathbf{w}_1, \ldots, \mathbf{w}_m$ sunt liniar indepedenţi. Rezultă din propoziţia precentă că $m \leq n$. Reluând raţionamentul pentru baza $\mathcal{B}'$ şi vectorii liniar independenţi $\mathbf{v}_1, \ldots, \mathbf{v}_n$, obţinem $n \leq m$. Deci $m = n$. □

Conform celor obţinute mai sus, rezultă că dacă un spaţiu vectorial admite o bază cu un număr finit de vectori, atunci orice altă bază va fi de asemenea finită, cu acelaşi număr de vectori.

Dimensiunea unui spaţiu vectorial

Definiţie 4.12. Fie V un $\Bbbk$-spaţiu vectorial care admite o bază cu n vectori. Numărul vectorilor din bază se numeşte **dimensiunea** spaţiului vectorial V şi se notează $\mathbf{dim}_{\Bbbk} V$.

Exemplu 4.13. Următoarele baze, introduse în Exemplul 4.4 apar în mod natural în spaţiile vectoriale indicate (de unde şi denumirea de baze canonice) şi determină dimensiunea acestora:

1) $\dim_{\Bbbk} \Bbbk^n = n$, baza $\mathcal{B} = \{\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\}$

2) $\dim_{\Bbbk} \Bbbk_n[X] = n + 1$, baza $\{1, X, X^2, \ldots, X^n\}$

3) $\dim_{\Bbbk} \mathcal{M}_{m,n}(\Bbbk) = mn$, baza $\{E_{ij}\}_{i=\overline{1,m}, j=\overline{1,n}}$

Fie acum $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ o mulţime de vectori şi $1 \leq r \leq n$. Vom spune că $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ este o (sub)mulţime **maximală** de vectori liniar independenţi dacă $\mathbf{v}_1, \ldots, \mathbf{v}_r$ sunt liniar independenţi, şi pentru orice $\mathbf{v}_i$ cu $i > r$, vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_i$ sunt liniar dependenţi.

Din orice sistem de generatori se poate extrage o bază

Propoziţie 4.14. Fie V un $\Bbbk$-spaţiu vectorial şi $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ cu proprietatea că $V = \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Să presupunem, eventual renumerotând vectorii, că $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ este o mulţime **maximală** de vectori liniar independenţi. Atunci $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ este bază în V.

Demonstraţie. Conform ipotezei, vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_r$ sunt liniar independenţi. Trebuie doar să verificăm că aceştia generează pe V.

Pasul I: Vom arăta că pentru orice $r < i \leq n$, $\mathbf{v}_i \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. Fie deci i cu $r < i \leq n$. Întrucât $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ este o mulţime **maximală** de vectori liniar independenţi, $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_i$ vor fi liniar dependenţi. Deci există scalarii $\alpha_1, \ldots, \alpha_r, \beta \in \Bbbk$, nu toţi nuli, astfel încât

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_r \mathbf{v}_r + \beta \mathbf{v}_i = \mathbf{0}$$

Dacă $\beta = 0$, atunci $\alpha_1\mathbf{v}_1 + \ldots + \alpha_r\mathbf{v}_r = \mathbf{0}$, şi ţinând cont de liniar independenţa vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_r$, rezultă $\alpha_1 = \ldots = \alpha_r = 0$, contradicţie. Deci $\beta \neq 0$, de unde

$$\mathbf{v}_i = (-\alpha_1\beta^{-1})\mathbf{v}_1 + \ldots + (-\alpha_r\beta^{-1})\mathbf{v}_r \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$$

Pasul II: $V = \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. Fie $\mathbf{v} \in V$. Deoarece $V = \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, există $\alpha_1, \ldots, \alpha_n \in \Bbbk$ astfel încât

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n$$

De la Pasul I ştim că $\mathbf{v}_i \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ pentru orice $r < i \leq n$, deci

$$\begin{aligned}
\mathbf{v}_{r+1} &= \beta_{1(r+1)}\mathbf{v}_1 + \ldots + \beta_{r(r+1)}\mathbf{v}_r \\
&\vdots \\
\mathbf{v}_n &= \beta_{1n}\mathbf{v}_1 + \ldots + \beta_{rn}\mathbf{v}_r
\end{aligned}$$

pentru nişte scalari $\beta_{1(r+1)}, \ldots, \beta_{rn} \in \Bbbk$. Atunci

$$\begin{aligned}
\mathbf{v} &= \alpha_1\mathbf{v}_1 + \ldots + \alpha_r\mathbf{v}_r + \alpha_{r+1}(\beta_{1(r+1)}\mathbf{v}_1 + \ldots + \beta_{r(r+1)}\mathbf{v}_r) + \ldots + \alpha_n(\beta_{1n}\mathbf{v}_1 + \ldots + \beta_{rn}\mathbf{v}_r) \\
&= (\alpha_{r+1}\beta_{1(r+1)} + \ldots + \alpha_n\beta_{1n})\mathbf{v}_1 + \ldots + (\alpha_{r+1}\beta_{r(r+1)} + \ldots + \alpha_n\beta_{rn})\mathbf{v}_r \\
&\in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}
\end{aligned}$$

□

Corolar 4.15. 1) Fie $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq V$ o mulţime maximală de vectori liniar independenţi. Atunci $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ este bază în V.

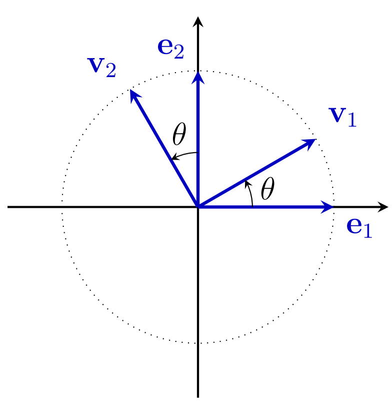
2) Fie V un spaţiu vectorial de dimensiune n şi $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ vectori liniar independenţi. Atunci $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ este o bază în V.

Demonstraţie. Exerciţiu. □

Observaţie 4.16. Dacă aplicăm rezultatul de mai sus în cazul spaţiului vectorial $\Bbbk^n$, obţinem următorul criteriu de recunoaştere a unei baze:

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \Bbbk^n \text{ bază} \iff \mathbf{v}_1, \ldots, \mathbf{v}_n \text{ liniar independenţi} \iff \det(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n) \neq 0.$$

Exerciţiu 4.17. Fie $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$ vectorii obţinuţi din $\mathbf{e}_1, \mathbf{e}_2$ în urma rotaţiei cu unghiul $\theta \in [0, \pi]$ în sens trigonometric.



1) Arătaţi că $\mathbf{v}_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ şi $\mathbf{v}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$ (indicaţie: puteţi utiliza numere complexe).

2) Verificaţi că $\{\mathbf{v}_1, \mathbf{v}_2\}$ este o bază în $\mathbb{R}^2$.

3) Cum se modifică coordonatele unui vector arbitrar $\mathbf{v} \in \mathbb{R}^2$ la schimbarea bazei? Cu alte cuvinte, dacă $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$, determinaţi $[\mathbf{v}]_{\mathcal{B}}$.

Orice mulţime de vectori liniar independenţi se poate completa la o bază

Propoziţie 4.18. Fie V un spaţiu vectorial de dimensiune n şi $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ vectori liniar independenţi, unde $r < n$. Atunci există $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \in V$ astfel încât $\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ este bază în V.

Demonstraţie. Deoarece $r < n$, vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_r$ nu formează o bază, in particular nu sunt o mulţime **maximală** de vectori liniar independenţi. Deci există $\mathbf{v}_{r+1} \in V$ astfel încât $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}$ sunt liniar independenţi. Dacă $r + 1 < n$, repetăm procedura, până când obţinem n vectori liniar independenţi. Conform Corolarului 4.15, aceştia formează o bază. □

4.4 Dimensiunea unui subspaţiu vectorial

Orice subspaţiu vectorial este de fapt un spaţiu vectorial în raport cu operaţiile induse, deci admite în particular o bază şi putem vorbi despre dimensiunea subspaţiului vectorial. Propoziţia următoare menţionează câteva proprietăţi ale (dimensiunii) unui subspaţiu vectorial:

Dimensiunea unui subspaţiu vectorial

Propoziţie 4.19. Fie V un $\mathbb{k}$-spaţiu vectorial finit dimensional.

1) Dacă $U \subseteq V$ este un subspaţiu vectorial, atunci:

   (a) $\dim_{\mathbb{k}} U \leq \dim_{\mathbb{k}} V$

   (b) $\dim_{\mathbb{k}} U = \dim_{\mathbb{k}} V \iff U = V$

2) Fie $U, W \subseteq V$ subspaţii vectoriale. Atunci:

$$\dim_{\mathbb{k}}(U + W) = \dim_{\mathbb{k}} U + \dim_{\mathbb{k}} W - \dim_{\mathbb{k}}(U \cap W) \qquad \text{(Grassmann)}$$

Demonstraţie. 1) (a) Fie $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ bază în U. În particular, vectorii $\mathbf{u}_1, \ldots, \mathbf{u}_m$ sunt liniar independenţi şi $\mathbf{u}_1, \ldots, \mathbf{u}_m \in V$, deci conform Propoziţiei 4.10, $\dim_{\mathbb{k}} U = m \leq \dim_{\mathbb{k}} V$.

(b) Fie $n = \dim_{\mathbb{k}} V$. Conform ipotezei, există o bază $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ în U cu n vectori. În particular, cei n vectori sunt liniar independenţi în V, deci conform Corolarului 4.15 formează o bază în V. Atunci $V = \mathsf{Sp}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\} = U$.

2) Să notăm $r = \dim_{\mathbb{k}} U \cap W$, $m = \dim_{\mathbb{k}} U$, $n = \dim_{\mathbb{k}} W$. Conform afirmaţiei 1).(a) de mai sus, $r \leq m$ şi $r \leq n$.

Fie acum $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ o bază în $U \cap W$. În particular, vectorii sunt liniar independenţi şi aparţin subspaţiului U, deci conform Propoziţiei4.18 îi putem extinde la o bază

$$\mathcal{B}_U = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \underbrace{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_m}_{m-r \text{ vectori}}\}$$

în U. Analog, putem completa baza din $U \cap W$ la o bază în W:

$$\mathcal{B}_W = \{\mathbf{v}_1, \ldots, \mathbf{v}_r, \underbrace{\mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+n-r}}_{n-r \text{ vectori}}\}$$

Vom arăta acum că $\mathcal{B}_U \cup \mathcal{B}_W = \{\mathbf{v}_1, \ldots, \mathbf{v}_{m+n-r}\}$ este bază în $U + W$, de unde va rezulta

$$\dim_{\Bbbk}(U + W) = |\mathcal{B}_U \cup \mathcal{B}_W| = m + n - r = \dim_{\Bbbk} U + \dim_{\Bbbk} W - \dim_{\Bbbk}(U \cap W)$$

Pasul I. Să verificăm că vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_{m+n-r}$ sunt liniar independenţi: fie $\alpha_1, \ldots, \alpha_{m+n-r} \in \Bbbk$ astfel încât $\sum_{i=1}^{m+n-r} \alpha_i \mathbf{v}_i = \mathbf{0}$. Atunci vectorul

$$\underbrace{\sum_{i=1}^{m} \alpha_i \mathbf{v}_i}_{\in U} = \underbrace{\sum_{i=m+1}^{m+n-r} (-\alpha_i) \mathbf{v}_i}_{\in W}$$

aparţine intersecţiei $U \cap W$, deci se poate scrie în mod unic ca o combinaţie liniară a vectorilor $\mathbf{v}_1, \ldots, \mathbf{v}_r$:

$$\sum_{i=1}^{m} \alpha_i \mathbf{v}_i = \sum_{i=1}^{r} \beta_i \mathbf{v}_i$$

Aceasta se mai poate scrie

$$\sum_{i=1}^{r} (\alpha_i - \beta_i) \mathbf{v}_i + \sum_{i=r+1}^{m} \alpha_i \mathbf{v}_i = \mathbf{0}$$

Dar vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_m$ formează o bază în U, în particular sunt liniar independenţi, deci $\alpha_i = \beta_i, i = \overline{1, r}$ şi $\alpha_i = 0, i = \overline{r+1, m}$.

Printr-un raţionament analog rezultă $\alpha_i = 0, i = \overline{m+1, m+n-r}$, deci combinaţia liniară nulă de la care am plecat se reduce la $\sum_{i=1}^{r} \alpha_i \mathbf{v}_i = \mathbf{0}$. Ţinând cont că $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ este bază în $U \cap W$, obţinem în final şi restul scalarilor zero.

Pasul II. $\mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_{m+n-r}\} = U + W$. Fie $\mathbf{v} \in U + W$. Atunci $\mathbf{v} = \mathbf{u} + \mathbf{w}$, unde $\mathbf{u} \in U, \mathbf{w} \in W$. În particular, există scalarii $\alpha_i, \beta_j \in \Bbbk$ astfel încât $\mathbf{u} = \sum_{i=1}^{m} \alpha_i \mathbf{v}_i$ şi $\mathbf{w} = \sum_{i \in \{1, \ldots, r, m+1, \ldots, m+n-r\}} \beta_i \mathbf{v}_i$. Atunci

$$\mathbf{v} = \sum_{i=1}^{r} (\alpha_i + \beta_i) \mathbf{v}_i + \sum_{i=r+1}^{m} \alpha_i \mathbf{v}_i + \sum_{i=m+1}^{m+n-r} \beta_i \mathbf{v}_i \in \mathsf{Sp}\{\mathbf{v}_1, \ldots, \mathbf{v}_{m+n-r}\}$$

□

Observaţie 4.20. Afirmaţia 1).(b) din Propoziţia de mai sus **nu** este neapărat adevărată în cazul spaţiilor vectoriale infinit dimensionale. De exemplu, spaţiul vectorial al polinoamelor $V = \Bbbk[X]$ are dimensiunea infinită (o bază fiind $1, X, X^2, \ldots, X^n, \ldots$), iar subspaţiul vectorial al polinoamelor pare $U = \{P \in \Bbbk[X] \mid P(X) = P(-X)\}$ are de asemenea dimensiunea infinită (o bază fiind formată cu polinoamele $1, X^2, X^4, \ldots, X^{2n}, \ldots$). Cu toate acestea, $U \neq V$.

Dimensiunea subspaţiilor asociate unei matrice

Propoziţie 4.21. Fie $A \in \mathcal{M}_{m,n}(\Bbbk)$. Atunci:

1) $\dim_\Bbbk \mathsf{Im}(A) = \mathsf{rang}(A)$

2) $\dim_\Bbbk \mathsf{Ker}(A) = n - \mathsf{rang}(A)$

Demonstraţie. 1) Fără demonstraţie.

2) Vom obţine relaţia $\dim_\Bbbk \mathsf{Ker}(A) = n - \dim_\Bbbk \mathsf{Im}(A)$ în Secţiunea 5.

□

Observaţie 4.22. În particular, rezultă $\dim_\Bbbk \mathsf{Im}(A^t) = \mathsf{rang} A$ şi $\dim_\Bbbk \mathsf{Ker}(A^t) = m - \mathsf{rang} A$.

4.5 Matricea de trecere de la o bază la altă bază

Fie V un $\Bbbk$-spaţiu vectorial de dimensiune n şi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ două baze în V. În particular, fiecare vector din $\mathcal{B}'$ se poate scrie sub forma unei combinaţii liniare (unice!) ale vectorilor din baza $\mathcal{B}$:

$$\begin{aligned}
\mathbf{w}_1 &= \alpha_{11}\ \mathbf{v}_1 + \ \ldots\ + \alpha_{n1}\ \mathbf{v}_n \\
\mathbf{w}_2 &= \alpha_{12}\ \mathbf{v}_1 + \ \ldots\ + \alpha_{n2}\ \mathbf{v}_n \\
&\ldots \\
\mathbf{w}_n &= \alpha_{1n}\ \mathbf{v}_1 + \ \ldots\ + \alpha_{nn}\ \mathbf{v}_n
\end{aligned}$$

Se obţine astfel o matrice:

$$\begin{array}{c}
\begin{array}{cccc} [\mathbf{w}_1]_\mathcal{B} & [\mathbf{w}_2]_\mathcal{B} & \ldots & [\mathbf{w}_n]_\mathcal{B} \end{array} \\
\begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}
\end{array}$$

pe care o vom numi **matricea de trecere de la baza $\mathcal{B}$ la baza $\mathcal{B}'$** şi pe care o vom nota $P_\mathcal{B}^{\mathcal{B}'}$.

Observaţie 4.23. 1) Matricea de trecere de la o bază $\mathcal{B}$ la ea însăşi este **matricea identică**

$$P_\mathcal{B}^\mathcal{B} = I_n$$

2) Matricea de trecere $P_\mathcal{B}^{\mathcal{B}'}$ de la o bază $\mathcal{B}$ la o bază $\mathcal{B}'$ este **inversabilă** şi

$$\left(P_\mathcal{B}^{\mathcal{B}'}\right)^{-1} = P_{\mathcal{B}'}^\mathcal{B}$$

Exemplu 4.24. 1) Fie $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2\}$ baza canonică din $\mathbb{R}^2$ şi $\mathcal{B}' = \{\mathbf{w}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}\}$ o altă bază (verificaţi că $\mathcal{B}'$ este într-adevăr o bază). Atunci matricea de trecere de la baza canonică la baza $\mathcal{B}'$ este $P_\mathcal{B}^{\mathcal{B}'} = \begin{pmatrix} 2 & -2 \\ -1 & 3 \end{pmatrix}$ (de ce?). Ce observaţi? Generalizaţi pentru cazul $\Bbbk^n$.

2) Să luăm acum o a treia bază $\mathcal{B}'' = \{\mathbf{u}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}\}$. Ca mai sus, avem $P_{\mathcal{B}}^{\mathcal{B}''} = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$. Scriind vectorii din $\mathcal{B}''$ în funcţie de baza $\mathcal{B}'$, obţinem:

$$\begin{aligned} \mathbf{u}_1 &= \frac{11}{4}\mathbf{w}_1 + \frac{5}{4}\mathbf{w}_2 \\ \mathbf{u}_2 &= -\frac{3}{4}\mathbf{w}_1 + \frac{1}{4}\mathbf{w}_2 \end{aligned} \tag{6}$$

de unde rezultă $P_{\mathcal{B}'}^{\mathcal{B}''} = \begin{pmatrix} \frac{11}{4} & -\frac{3}{4} \\ \frac{5}{4} & \frac{1}{4} \end{pmatrix}$.

3) Să observăm mai departe că putem scrie vectorii din $\mathcal{B}''$ în două moduri: mai întâi, sub forma combinaţiilor liniare ale vectorilor din baza canonică:

$$\begin{aligned} \mathbf{u}_1 &= 3\mathbf{e}_1 + \mathbf{e}_2 \\ \mathbf{u}_2 &= (-1)\mathbf{e}_1 + \mathbf{e}_2 \end{aligned} \tag{7}$$

Pe de altă parte, în relaţiile (6) putem exprima mai departe şi $\mathbf{w}_1, \mathbf{w}_2$ sub forma unor combinaţii liniare ale vectorilor din baza canonică:

$$\begin{aligned} \mathbf{u}_1 &= \frac{11}{4}\mathbf{w}_1 + \frac{5}{4}\mathbf{w}_2 &&= \frac{11}{4}(2\mathbf{e}_1 + (-1)\mathbf{e}_2) + \frac{5}{4}((-2)\mathbf{e}_1 + 3\mathbf{e}_2) \\ \mathbf{u}_2 &= -\frac{3}{4}\mathbf{w}_1 + \frac{1}{4}\mathbf{w}_2 &&= -\frac{3}{4}(2\mathbf{e}_1 + (-1)\mathbf{e}_2) + \frac{1}{4}((-2)\mathbf{e}_1 + 3\mathbf{e}_2) \end{aligned} \tag{8}$$

Comparând (7) şi (8), rezultă

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

adică

$$P_{\mathcal{B}}^{\mathcal{B}''} = P_{\mathcal{B}}^{\mathcal{B}'} P_{\mathcal{B}'}^{\mathcal{B}''}$$

Să observăm că raţionamentele de mai sus sunt de fapt valabile în orice spaţiu vectorial şi pot fi rezumate diagramatic astfel:

$$\begin{tikzcd} \mathcal{B} \arrow[r, "P_{\mathcal{B}}^{\mathcal{B}'}"] \arrow[rr, bend right, "P_{\mathcal{B}}^{\mathcal{B}''} = P_{\mathcal{B}}^{\mathcal{B}'} P_{\mathcal{B}'}^{\mathcal{B}''}"'] & \mathcal{B}' \arrow[r, "P_{\mathcal{B}'}^{\mathcal{B}''}"] & \mathcal{B}'' \end{tikzcd} \tag{9}$$

4.6 Schimbarea coordonatelor la schimbarea bazei

Am văzut că într-un spaţiu vectorial putem avea mai multe baze. Ce se întâmplă cu coordonatele unui vector la schimbarea bazei? Mai precis, vom studia în continuare următoarea situaţie: $\mathbf{v} \in V$ este un vector arbitrar şi $\mathcal{B}$, $\mathcal{B}'$ sunt baze în V. Reamintim că am notat cu $[\mathbf{v}]_{\mathcal{B}}$ şi $[\mathbf{v}]_{\mathcal{B}'}$ coordonatele vectorului în baza $\mathcal{B}$, respectiv $\mathcal{B}'$:

Baze în V	$\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$	$\mathcal{B}' = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$
Vector $\mathbf{v} \in V$	$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n$	$\mathbf{v} = \beta_1 \mathbf{w}_1 + \ldots + \beta_n \mathbf{w}_n$
Coordonate	$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$	$[\mathbf{v}]_{\mathcal{B}'} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$

Atunci, cu notaţiile din secţiunea precedentă, putem scrie:

$$\begin{aligned}
\mathbf{v} &= \beta_1 \mathbf{w}_1 + \ldots + \beta_n \mathbf{w}_n \\
&= \beta_1(\alpha_{11}\mathbf{v}_1 + \ldots + \alpha_{n1}\mathbf{v}_n) + \ldots + \beta_n(\alpha_{1n}\mathbf{v}_1 + \ldots + \alpha_{nn}\mathbf{v}_n) \\
&= \underbrace{(\beta_1\alpha_{11} + \ldots + \beta_n\alpha_{1n})}\mathbf{v}_1 + \ldots + \underbrace{(\beta_1\alpha_{n1} + \ldots + \beta_n\alpha_{nn})}\mathbf{v}_n
\end{aligned}$$

coordonatele vectorului $\mathbf{v}$ în raport cu baza $\mathcal{B}$

Ţinând cont de unicitatea scrierii unui vector într-o bază, rezultă următoarea formulă de schimbare a coordonatelor la schimbarea bazei:

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = P_{\mathcal{B}}^{\mathcal{B}'} [\mathbf{v}]_{\mathcal{B}'}$$

5 Aplicaţii liniare şi matrice

Matrices act. They don't just sit there.

Gilbert Strang, *Introduction to Linear Algebra*

5.1 Preliminarii

Să observăm că orice matrice $A \in \mathcal{M}_{m,n}(\mathbb{k})$ induce o funcţie

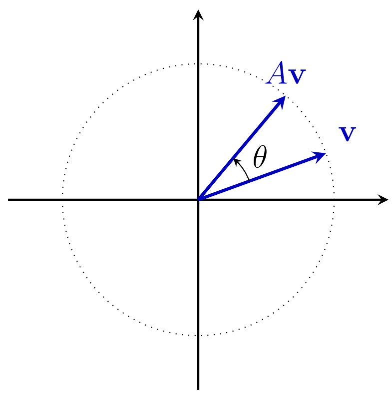
$$f : \mathbb{k}^n \longrightarrow \mathbb{k}^m \quad , \quad f(\mathbf{x}) = A\mathbf{x}$$

care verifică următoarele relaţii

$$\begin{aligned} &f(\mathbf{x}+\mathbf{x}') = f(\mathbf{x}) + f(\mathbf{x}') \\ &f(\alpha\mathbf{x}) = \alpha f(\mathbf{x}) \end{aligned}$$

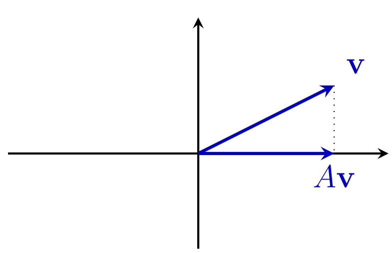
În particular, o serie de transformări geometrice familiare pot fi obţinute astfel:

1) Rotaţia în plan de unghi $\theta \in [0, 2\pi)$ în sens trigonometric:



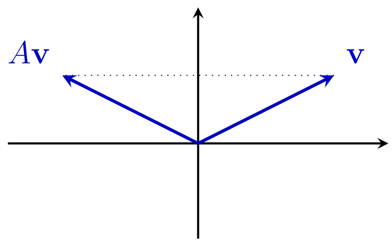
Fie $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Arătaţi că pentru orice $\mathbf{v} \in \mathbb{R}^2$, vectorul $A\mathbf{v}$ este vectorul obţinut în urma rotaţiei în sens trigonometric de unghi θ (indicaţie: puteţi utiliza numere complexe).

2) Proiecţia pe o dreaptă care trece prin origine. Vom exemplifica cu proiecţia pe axa Ox: dacă $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, atunci $A\mathbf{v} = \begin{pmatrix} x \\ 0 \end{pmatrix}$, unde $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.



Pentru cazul general: dacă $d : y = \lambda x$ este ecuaţia unei drepte care trece prin origine[13], unde $\lambda \in \mathbb{R}$, atunci verificaţi că matricea $A = \begin{pmatrix} \frac{1}{\lambda^2+1} & \frac{\lambda}{\lambda^2+1} \\ \frac{\lambda}{\lambda^2+1} & \frac{\lambda^2}{\lambda^2+1} \end{pmatrix}$ realizează proiecţia. Mai precis, dacă $\mathbf{v} \in \mathbb{R}^2$ este un vector arbitrar, atunci $\mathsf{pr}_d \mathbf{v} = A\mathbf{v}$.

3) Simetria faţă de o dreaptă care trece prin origine. Vom exemplifica cu simetria faţă de axa Oy, cu ajutorul matricei $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$:



Utilizând eventual rezultatul de la punctul precedent, determinaţi matricea pentru cazul general al simetriei faţă de o dreaptă arbitrară de ecuaţie $y = \lambda x$ care trece prin origine.

5.2 Aplicaţii liniare

Aplicaţie liniară

Definiţie 5.1. Fie U, V două $\mathbb{k}$-spaţii vectoriale. **O aplicaţie liniară** este o funcţie $f : U \to V$ cu următoarele proprietăţi:

$$\begin{aligned} f(\mathbf{x} + \mathbf{x}') &= f(\mathbf{x}) + f(\mathbf{x}') \\ f(\alpha \mathbf{x}) &= \alpha f(\mathbf{x}) \end{aligned} \tag{10}$$

pentru orice $\mathbf{x}, \mathbf{x}' \in U$ şi orice $\alpha \in \mathbb{k}$.

Observaţie 5.2. 1) Arătăţi că o funcţie $f : U \to V$ este aplicaţie liniară dacă şi numai dacă satisface

$$f(\alpha \mathbf{x} + \alpha' \mathbf{x}') = \alpha f(\mathbf{x}) + \alpha' f(\mathbf{x}')$$

pentru orice $\mathbf{x}, \mathbf{x}' \in U$ şi orice $\alpha, \alpha' \in \mathbb{k}$.

2) Orice aplicaţie liniară verifică următoarele:

$$f(\mathbf{0}) = \mathbf{0} \quad \text{şi} \quad f(-\mathbf{x}) = -f(\mathbf{x}),\ \forall\ \mathbf{x} \in U$$

Izomorfism

Definiţie 5.3. O aplicaţie liniară bijectivă $f : U \to V$ se numeşte **izomorfism**. În acest caz, spunem că spaţiile vectoriale U şi V sunt **izomorfe** şi notăm $U \cong V$.

[13] Mai puţin axa Oy.

Exemple 5.4. 1) Exemple universale: funcţia identică $\mathsf{Id}_V : V \to V$, $\mathsf{Id}_V(\mathbf{v}) = \mathbf{v}$ este aplicaţie liniară. De asemenea, funcţia constantă nulă $\mathbf{0}_{U,V} : U \to V$, $\mathbf{0}_{U,V}(\mathbf{u}) = \mathbf{0}$ este aplicaţie liniară.

2) Fie V un $\mathbb{k}$-spaţiu vectorial de dimensiune n şi $\mathcal{B}$ o bază în V. Utilizând Criteriul bazei (unicitatea reprezentării fiecărui vector în baza $\mathcal{B}$), arătaţi că funcţia

$$f : V \to \mathbb{k}^n \,,\; f(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$$

este o aplicaţie liniară.

3) Orice matrice $A \in \mathcal{M}_{m,n}(\mathbb{k})$ induce o aplicaţie liniară

$$f : \mathbb{k}^n \to \mathbb{k}^m, \quad f(\mathbf{x}) = A\mathbf{x}$$

În particular, transformările geometrice descrise în secţiunea precedentă sunt exemple de aplicaţii liniare. Observaţi că dintre acestea lipseşte una importantă: translaţia. Motivul este simplu: translaţia nu păstrează vectorul nul. Translaţia face parte dintr-o clasă mai largă de transformări, numite *transformări afine*.

4) Operatorii uzuali de derivare şi integrare sunt aplicaţii liniare (pe spaţii vectoriale convenabil alese). De exemplu, fie $\mathcal{C}^1(\mathbb{R})$ spaţiul vectorial al funcţiilor $\mathbb{R} \to \mathbb{R}$ de clasă $\mathcal{C}^1$, respectiv $\mathcal{C}^0(\mathbb{R})$ spaţiul vectorial al funcţiilor continue, cu operaţiile punctuale, ca în Exemplul 2.3.3. Atunci se observă uşor că $f : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$, $f(\varphi) = \varphi'$ şi $g : \mathcal{C}^0(\mathbb{R}) \to \mathcal{C}^1(\mathbb{R})$, $g(\varphi)(x) = \displaystyle\int_0^x \varphi(t)\mathtt{d}t$ verifică (10), deci sunt aplcaţii liniare între spaţiile vectoriale considerate.

5) Putem găsi exemple de aplicaţii liniare şi în spaţiile vectoriale de matrice, cum ar fi operaţia de transpunere sau urma unei matrice:

$$f : \mathcal{M}_{m,n}(\mathbb{k}) \to \mathcal{M}_{n,m}(\mathbb{k}) \,,\; f(X) = X^t$$

sau

$$g : \mathcal{M}_{m,n}(\mathbb{k}) \to \mathbb{k} \,,\; g(X) = \mathtt{tr}(X)$$

Să vedem acum comportamentul aplicaţiilor liniare în raport cu compunerea funcţiilor:

Propoziţie 5.5. 1) Dacă $f : U \to V$, $g : V \to W$ sunt aplicaţii liniare, atunci şi $g \circ f : U \to W$ este o aplicaţie liniară.

2) Dacă $f : U \to V$ este un izomorfism, atunci şi f^{-1} este o aplicaţie liniară, în particular un izomorfism.

Demonstraţie. Exerciţiu. □

Următoarea propoziţie ne arată că este suficient să cunoaştem acţiunea unei aplicaţii liniare pe o bază:

O aplicaţie liniară este perfect determinată de acţiunea sa pe o bază

Propoziţie 5.6. Fie U, V două $\Bbbk$-spaţii vectoriale, $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ bază în U şi $\mathbf{v}_1, \dots, \mathbf{v}_n$ vectori arbitrari din V. Atunci există o unică aplicaţie liniară $f : U \to V$ astfel încât $f(\mathbf{u}_1) = \mathbf{v}_1, \dots, f(\mathbf{u}_n) = \mathbf{v}_n$.

Demonstraţie. Existenţa. Fie $\mathbf{u} \in U$ un vector arbitrar, şi $\mathbf{u} = \alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n$ scrierea sa în baza $\mathcal{B}$. Definim $f(\mathbf{u}) = \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{u}_n$. Deoarece scalarii $\alpha_1, \dots, \alpha_n$ sunt unici pentru fiecare $\mathbf{u} \in U$, $f : U \to V$ este bine-definită. Să arătăm că f este o aplicaţie liniară: dacă $\mathbf{u}'$ este alt vector din U, şi $\mathbf{u}' = \alpha_1'\mathbf{u}_1 + \ldots + \alpha_n'\mathbf{u}_n$, atunci $\mathbf{u} + \mathbf{u}' = (\alpha_1 + \alpha_1')\mathbf{u}_1 + \ldots + (\alpha_n + \alpha_n')\mathbf{u}_n$. Din construcţia lui f, rezultă

$$\begin{aligned} f(\mathbf{u} + \mathbf{u}') &= (\alpha_1 + \alpha_1')\mathbf{v}_1 + \ldots + (\alpha_n + \alpha_n')\mathbf{v}_n \\ &= \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n + \alpha_1'\mathbf{v}_1 + \ldots + \alpha_n'\mathbf{v}_n \\ &= f(\mathbf{u}) + f(\mathbf{u}') \end{aligned}$$

Fie acum $\alpha \in \Bbbk$. Atunci $\alpha\mathbf{u} = (\alpha\alpha_1)\mathbf{u}_1 + \ldots (\alpha\alpha_n)\mathbf{u}_n$, deci

$$\begin{aligned} f(\alpha\mathbf{u}) &= (\alpha\alpha_1)\mathbf{v}_1 + \ldots + (\alpha\alpha_n)\mathbf{v}_n \\ &= \alpha f(\mathbf{u}) \end{aligned}$$

Deci f este o aplicaţie liniară.

Unicitatea. Dacă $g : U \to V$ este o altă aplicaţie liniară cu proprietatea că $g(\mathbf{u}_1) = \mathbf{v}_1, \dots, g(\mathbf{u}_n) = \mathbf{v}_n$, atunci pentru orice $\mathbf{u} \in U$, avem

$$\begin{aligned} g(\mathbf{u}) &= g(\alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n) \\ &= \alpha_1 g(\mathbf{u}_1) + \ldots + \alpha_n g(\mathbf{u}_n) \\ &= \alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n \\ &= f(\mathbf{u}) \end{aligned}$$

□

5.3 Subspaţiile vectoriale asociate unei aplicaţii liniare

Nucleul şi imaginea unei aplicaţii liniare

Definiţie 5.7. Fie $f : U \to V$ o aplicaţie liniară.

1) Mulţimea $\mathsf{Ker}(f) = \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}\}$ se numeşte **nucleul** lui f.

2) Mulţimea $\mathsf{Im}(f) = \{\mathbf{v} \in V \mid \exists\, \mathbf{u} \in U.\ f(\mathbf{u}) = \mathbf{v}\}$ se numeşte **imaginea** lui f.

Exemplu 5.8. Fie $A \in \mathcal{M}_{m,n}(\Bbbk)$ şi $f_A : \Bbbk^n \to \Bbbk^m$ aplicaţia liniară asociată, $f(\mathbf{x}) = A\mathbf{x}$. Atunci

$$\mathsf{Ker}(f) = \mathsf{Ker}(A) \quad \text{şi} \quad \mathsf{Im}(f) = \mathsf{Im}(A)$$

Cele două mulţimi introduse mai sus sunt de fapt subspaţii vectoriale. Nucleul unei aplicaţii liniare este util în studiul injectivităţii. De asemenea, cu ajutorul imaginii caracterizăm surjectivitatea (fapt cunoscut încă din liceu).

Nucleu şi imagine

Propoziţie 5.9. Fie $f : U \to V$ o aplicaţie liniară. Atunci:

1) $\mathsf{Ker}(f)$ este subspaţiu vectorial al lui U şi $\mathsf{Im}(f)$ este subspaţiu vectorial al lui V.

2) f este injectivă $\Longleftrightarrow \mathsf{Ker}(f) = \{\mathbf{0}_U\}$.

3) f este surjectivă $\Longleftrightarrow \mathsf{Im}(f) = V$.

Demonstraţie. 1) Exerciţiu.

2) Să presupunem că f este injectivă şi fie $\mathbf{v} \in \mathsf{Ker}(f)$. Atunci $f(\mathbf{v}) = \mathbf{0}$; dar $f(\mathbf{0}) = \mathbf{0}$, de unde rezultă $\mathbf{v} = \mathbf{0}$. Invers, să presupunem că $ker f = \{\mathbf{0}\}$ şi fie $\mathbf{u}_1, \mathbf{u}_2 \in U$ cu $f(\mathbf{u}_1) = f(\mathbf{u}_2)$. Atunci

$$f(\mathbf{u}_1 - \mathbf{u}_2) = f(\mathbf{u}_1 + (-1)\mathbf{u}_2) = f(\mathbf{u}_1) + (-1)f(\mathbf{u}_2) = f(\mathbf{u}_1) - f(\mathbf{u}_2) = \mathbf{0}$$

deci $\mathbf{u}_1 - \mathbf{u}_2 \in \mathsf{Ker}(f) = \{\mathbf{0}\}$, adică $\mathbf{u}_1 = \mathbf{u}_2$.

3) Rezultatul este cunoscut.

□

Corolar 5.10. Dacă $f : \mathbb{k}^n \to \mathbb{k}^m$, $f(\mathbf{x}) = A\mathbf{x}$, unde $A \in \mathcal{M}_{m,n}(\mathbb{k})$, atunci:

1) f este injectivă $\Longleftrightarrow \mathsf{Ker}(A) = \{\mathbf{0}_n\} \Longleftrightarrow \mathsf{rang}(A) = n$ (în particular, $n \leq m$).

2) f este surjectivă $\Longleftrightarrow \mathsf{Im}(A) = \mathbb{k}^m \Longleftrightarrow \mathsf{rang}(A) = m$ (în particular, $m \leq n$).

Următorul rezultat referitor la dimensiunile nucleului şi imaginii unei aplicaţii liniare deschide calea pentru caracterizarea fundamentală a spaţiilor vectoriale pe care o vom obţine în finalul acestei secţiuni:

Teorema rang-defect

Teoremă 5.11. Fie U, V $\mathbb{k}$-spaţii vectoriale finit dimensionale şi $f : U \to V$ o aplicaţie liniară. Atunci:

$$\dim_{\mathbb{k}} \mathsf{Ker}(f) + \dim_{\mathbb{k}} \mathsf{Im}(f) = \dim_{\mathbb{k}} U$$

Denumirea teoremei provine de la următorii termeni: $\dim_{\mathbb{k}} \mathsf{Ker}(f)$ se mai numeşte **defectul** aplicaţiei liniare f, iar $\dim_{\mathbb{k}} \mathsf{Im}(f)$ **rangul** lui f.

Demonstraţie. Fie $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ bază în $\mathsf{Ker}(f)$, pe care o prelungim la o bază $\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ în U. Atunci $\dim_{\mathbb{k}} \mathsf{Ker}(f) = r$ şi $\dim_{\mathbb{k}} U = n$. Vom arăta că $f(\mathbf{u}_{r+1}), \ldots, f(\mathbf{u}_n)$ formează o bază în $\mathsf{Im}(f)$. În particular, va rezulta $\dim_{\mathbb{k}} \mathsf{Im}(f) = n - r = \dim_{\mathbb{k}} U - \dim_{\mathbb{k}} \mathsf{Ker}(f)$.

Pas I. Liniar independenţa vectorilor $f(\mathbf{u}_{r+1}), \ldots, f(\mathbf{u}_n)$. Fie $\alpha_{r+1}, \ldots, \alpha_n \in \mathbb{k}$ astfel încât

$$\alpha_{r+1} f(\mathbf{u}_{r+1}) + \ldots + \alpha_n f(\mathbf{u}_n) = \mathbf{0}$$

Ţinând cont de liniaritatea funcţiei f, rezultă

$$f(\alpha_{r+1}\mathbf{u}_{r+1} + \ldots + \alpha_n\mathbf{u}_n) = \alpha_{r+1}f(\mathbf{u}_{r+1}) + \ldots + \alpha_n f(\mathbf{u}_n) = \mathbf{0}$$

deci $\alpha_{r+1}\mathbf{u}_{r+1} + \ldots + \alpha_n\mathbf{u}_n \in \mathsf{Ker}(f)$. În particular, vectorul de mai sus se poate scrie în mod unic sub forma unei combinaţii liniare a bazei din $\mathsf{Ker}(f)$:

$$\alpha_{r+1}\mathbf{u}_{r+1} + \ldots + \alpha_n\mathbf{u}_n = \beta_1\mathbf{u}_1 + \ldots + \beta_r\mathbf{u}_r$$

de unde rezultă

$$\beta_1\mathbf{u}_1 + \ldots + \beta_r\mathbf{u}_r + (-\alpha_{r+1})\mathbf{u}_{r+1} + \ldots + (-\alpha_n)\mathbf{u}_n = \mathbf{0}$$

Dar $\mathbf{u}_1, \ldots, \mathbf{u}_n$ formează o bază în U, în particular sunt liniar independenţi. Deci $\beta_1 = \ldots = \beta_r = 0$ şi $\alpha_{r+1} = \ldots = \alpha_n = 0$.

Pas II. $\mathsf{Sp}\{f(\mathbf{u}_{r+1}), \ldots, f(\mathbf{u}_n)\} = \mathsf{Im}(f)$. Fie $\mathbf{v} \in \mathsf{Im}(f)$. Atunci există $\mathbf{u} \in U$ cu $f(\mathbf{u}) = \mathbf{v}$. Scriind vectorul $\mathbf{u}$ în baza $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ sub forma $\mathbf{u} = \alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n$, rezultă

$$\begin{aligned}
\mathbf{v} &= f(\mathbf{u}) \\
&= f(\alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n) \\
&= \alpha_1\underbrace{f(\mathbf{u}_1)}_{\mathbf{0}} + \ldots + \alpha_r\underbrace{f(\mathbf{u}_r)}_{\mathbf{0}} + \alpha_{r+1}f(\mathbf{u}_{r+1}) + \ldots + \alpha_n f(\mathbf{u}_n) \\
&= \alpha_{r+1}f(\mathbf{u}_{r+1}) + \ldots + \alpha_n f(\mathbf{u}_n) \\
&\in \mathsf{Sp}\{f(\mathbf{u}_{r+1}), \ldots, f(\mathbf{u}_n)\}
\end{aligned}$$

□

Putem enunţa acum teorema fundamentală de caracterizare a spaţiilor vectoriale:

Oricare două spaţii vectoriale cu aceeaşi dimensiune sunt izomorfe

Teoremă 5.12. Fie U, V $\Bbbk$-spaţii vectoriale finit dimensionale. Atunci U şi V sunt izomorfe dacă şi numai dacă $\dim_{\Bbbk} U = \dim_{\Bbbk} V$.

Demonstraţie. "⇒" Fie $f : U \to V$ un izomorfism. Atunci $\mathsf{Ker}(f) = \{\mathbf{0}\}$ şi $\mathsf{Im}(f) = V$, deci

$$\dim_{\Bbbk} U = \dim_{\Bbbk} \mathsf{Ker}(f) = \dim_{\Bbbk} \mathsf{Im}(f) = 0 + \dim_{\Bbbk} V = \dim_{\Bbbk} V$$

"⇐" Fie $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ o bază în U şi $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ o bază în V. Conform Propoziţiei 5.6, există o unică aplicaţie liniară $f : U \to V$ astfel încât $f(\mathbf{u}_i) = \mathbf{v}_i$, $\forall i = \overline{1,n}$, şi o unică aplicaţie liniară $g : V \to U$ astfel încât $g(\mathbf{v}_i) = \mathbf{u}_i$, $\forall i = \overline{1,n}$. În particular rezultă $(g \circ f)(\mathbf{u}_i) = g(f(\mathbf{u}_i)) = g(\mathbf{v}_i) = \mathbf{u}_i$ pentru orice $i = \overline{1,n}$. Deci aplicaţia liniară $g \circ f : U \to U$ (Propoziţia 5.5) coincide cu aplicaţia identică pe toţi vectorii din baza $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Din Propoziţia 5.6 rezultă $g \circ f = \mathsf{Id}_U$. Analog, se arată că $f \circ g = \mathsf{Id}_V$. Deci f este un izomorfism de spaţii vectoriale şi $g = f^{-1}$. □

Corolar 5.13. Fie U un $\Bbbk$-spaţiu vectorial cu $\dim_{\Bbbk} U = n$. Atunci $U \cong \Bbbk^n$.

5.4 Matricea asociată unei aplicaţii liniare

Am văzut în Secţiunea precedentă că fiecărei matrice $A \in \mathcal{M}_{m,n}(\Bbbk)$ îi putem asocia o aplicaţie liniară $\Bbbk^n \to \Bbbk^m$, şi că fiecare spaţiu vectorial (finit dimensional) este izomorf cu un spaţiu vectorial de tip $\Bbbk^n$. Coroborând cele două rezultate, vom arăta în această Secţiune că fiecare aplicaţie liniară este *esenţial* determinată de o matrice. Schematic, situaţia se prezintă astfel:

$$\begin{array}{ccc}
\text{Spaţiu vectorial (vectori)} & \xrightarrow{\text{Aplicaţie liniară (funcţie)}} & \text{Spaţiu vectorial (vectori)} \\
\Big\downarrow \text{Bază} & & \Big\downarrow \text{Bază} \\
\text{Coordonate (scalari)} & \xrightarrow{\qquad\text{Matrice}\qquad} & \text{Coordonate (scalari)}
\end{array}$$

Fie deci o aplicaţie liniară $f : U \to V$ între două spaţii vectoriale finit dimensionale şi $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ o bază în U, respectiv $\mathcal{B}' = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ o bază în V. Dacă scriem vectorii $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \in V$ sub forma unor combinaţii liniare a vectorilor din baza $\mathcal{B}'$, obţinem

$$\begin{aligned}
f(\mathbf{u}_1) &= \alpha_{11}\ \mathbf{v}_1\ + \dots +\ \alpha_{m1}\ \mathbf{v}_m \\
f(\mathbf{u}_2) &= \alpha_{12}\ \mathbf{v}_1\ + \dots +\ \alpha_{m2}\ \mathbf{v}_m \\
&\ \vdots \\
f(\mathbf{u}_n) &= \alpha_{1n}\ \mathbf{v}_1\ + \dots +\ \alpha_{mn}\ \mathbf{v}_m
\end{aligned}$$

Vom nota cu $M(f)^{\mathcal{B}}_{\mathcal{B}'}$ matricea formată cu scalarii α_{ij}, $i = \overline{1,m}$, $j = \overline{1,n}$ şi o vom numi **matricea asociată aplicaţiei liniare** f în raport cu cele două baze $\mathcal{B}, \mathcal{B}'$:

$$M(f)^{\mathcal{B}}_{\mathcal{B}'} = \begin{array}{c} \begin{array}{cccc} [f(\mathbf{u}_1)]_{\mathcal{B}'} & [f(\mathbf{u}_2)]_{\mathcal{B}'} & \dots & [f(\mathbf{u}_n)]_{\mathcal{B}'} \end{array} \\ \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \end{array} \in \mathcal{M}_{m,n}(\Bbbk)$$

În cazul în care nu există pericol de confuzie, vom nota matricea de mai sus cu $M(f)$ în loc de $M(f)^{\mathcal{B}}_{\mathcal{B}'}$.

Fie acum $\mathbf{u} \in U$ un vector arbitrar; dacă reprezentăm vectorul sub forma unei combinaţii liniare a vectorilor din baza $\mathcal{B}$, $\mathbf{u} = \alpha_1\mathbf{u}_1 + \dots + \alpha_n\mathbf{u}_n$, atunci imaginea sa prin f devine

$$\begin{aligned}
f(\mathbf{u}) &= \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_n f(\mathbf{u}_n) \\
&= \alpha_1(\alpha_{11}\mathbf{v}_1 + \dots + \alpha_{m1}\mathbf{v}_m) + \dots + \alpha_n(\alpha_{1n}\mathbf{v}_1 + \dots + \alpha_{mn}\mathbf{v}_m) \\
&= \underbrace{(\alpha_1\alpha_{11} + \dots + \alpha_n\alpha_{1n})}_{\text{coordonatele vectorului}}\mathbf{v}_1 \ \underset{f(\mathbf{u})\text{ în raport cu}}{+ \ \dots \ +} \ \underbrace{(\alpha_1\alpha_{m1} + \dots + \alpha_n\alpha_{mn})}_{\text{baza } \mathcal{B}'}\mathbf{v}_m
\end{aligned}$$

Cu alte cuvinte, coordonatele vectorului $f(\mathbf{u})$ se obţin din coordonatele vectorului $\mathbf{u}$ prin multiplicarea cu $M(f)$:

$$[f(\mathbf{u})]_{\mathcal{B}'} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = M(f)[\mathbf{u}]_{\mathcal{B}}$$

Deci modulo alegerea bazelor $\mathcal{B}$ şi $\mathcal{B}'$ în domeniu, respectiv codomeniu, orice aplicaţie liniară se reduce în esenţă la multiplicarea cu o matrice.

Exemplu 5.14. 1) Dacă $f : \Bbbk^n \to \Bbbk^m$, $f(\mathbf{u}) = A\mathbf{u}$, atunci matricea asociată lui f în raport cu bazele canonice $M(f)$ este chiar A.

2) Fie $f : \Bbbk_n[X] \to \Bbbk_{n-1}[X]$ aplicaţia liniară care asociază fiecărui polinom de grad cel mult n derivata sa formală: $f(P) = P'$. Fie $\mathcal{B} = \{1, X, \ldots, X^n\}$, respectiv $\mathcal{B}' = \{1, X, \ldots, X^{n-1}\}$ bazele canonice din $\Bbbk_n[X]$ şi $\Bbbk_{n-1}[X]$. Atunci matricea asociată aplicaţiei liniare f este

$$M(f) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ldots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \ldots & n \end{pmatrix}$$

3) Să considerăm acum aplicaţia liniară $f : \mathcal{M}_2(\Bbbk) \to \mathcal{M}_2(\Bbbk)$ care asociază fiecărei matrice transpusa sa, $f(X) = X^t$. În acest exemplu, domeniul şi codomeniul coincid. Dacă luăm $\mathcal{B} = \mathcal{B}'$ baza canonică din $\mathcal{M}_2(\Bbbk)$, şi anume $\{E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$, atunci matrice asociată aplicaţiei liniare de transpunere va fi

$$M(f) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Proprietăţile matricei asociate unei aplicaţii liniare

Propoziţie 5.15. 1) Fie U şi V spaţii vectoriale finit dimensionale, $\mathcal{B}$ o bază pentru U şi $\mathcal{B}'$ o bază pentru V. Dacă f şi g sunt aplicaţii liniare $U \to V$, atunci $f = g \iff M(f)^{\mathcal{B}}_{\mathcal{B}'} = M(g)^{\mathcal{B}}_{\mathcal{B}'}$.

2) Matricea aplicaţiei nule $\mathbf{0}_{U,V} : U \to V$ este nulă în raport cu orice baze: $M(\mathbf{0}_{U,V}) = 0_{m,n}$, unde $m = \dim_{\Bbbk} V$, $n = \dim_{\Bbbk} U$.

3) Fie V un spaţiu vectorial finit dimensional şi $\mathcal{B}, \mathcal{B}'$ baze pentru V. Atunci matricea aplicaţiei identice $\mathsf{Id}_V : V \to V$ *în raport cu bazele* $\mathcal{B}'$ *şi* $\mathcal{B}$ este $P^{\mathcal{B}'}_{\mathcal{B}}$, matricea de trecere de la baza $\mathcal{B}$ la baza $\mathcal{B}'$ (atenţie la schimbarea ordinii bazelor!).

În particular, matricea aplicaţiei identice $\mathsf{Id}_V : V \to V$ în raport cu o singură bază $\mathcal{B}$ este matricea identică $M(\mathsf{Id}_V)^{\mathcal{B}}_{\mathcal{B}} = I_n$.

4) Fie U, V, W spaţii vectoriale finit dimensionale şi $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ baze pentru U, V, respectiv W. Fie $f : U \to V, g : V \to W$ aplicaţii liniare. Atunci matricea asociată aplicaţiei liniare $g \circ f$ în bazele $\mathcal{B}$ şi $\mathcal{B}''$ este

$$M(g \circ f)^{\mathcal{B}}_{\mathcal{B}''} = M(g)^{\mathcal{B}'}_{\mathcal{B}''} M(f)^{\mathcal{B}}_{\mathcal{B}'}$$

5) Fie U, V spaţii vectoriale finit dimensionale şi $\mathcal{B}$, $\mathcal{B}'$ baze pentru U, respectiv V. O aplicaţie liniară $f : U \to V$ este un **izomorfism** dacă şi numai dacă matricea asociată $M(f)$ este **inversabilă** (şi atunci $M(f^{-1}) = M(f)^{-1}$).

Demonstraţie. 1) Implicaţia "$\Rightarrow$" este evidentă.

"$\Leftarrow$" Să notăm $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Dacă $M(f)_{\mathcal{B}}^{\mathcal{B}'} = M(g)_{\mathcal{B}}^{\mathcal{B}'}$, atunci pentru orice $i = 1, n$, $[f(\mathbf{u}_i)]_{\mathcal{B}'} = [g(\mathbf{u}_i)]_{\mathcal{B}'}$, deci $f(\mathbf{u}_i) = g(\mathbf{u}_i)$ (reamintim din Observaţia 4.8 că doi vectori sunt egali dacă şi dacă au aceleaşi coordonate în raport cu o bază). Aplicând Propoziţia 5.6, rezultă $f = g$.

2) Clar.

3) Rezultă din definiţia matricei de trecere de la o bază la altă bază (Secţiunea 4.5).

4) Fie $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, $\mathcal{B}' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, $\mathcal{B}'' = \{\mathbf{w}_1, \ldots, \mathbf{w}_p\}$ baze în U,V, respectiv W. Dacă notăm

$$M(f)_{\mathcal{B}'}^{\mathcal{B}} = (\alpha_{ij})_{i=1,m;j=1,n} \qquad \text{unde} \qquad f(\mathbf{u}_j) = \sum_{i=1}^{m} \alpha_{ij}\mathbf{v}_i$$

$$M(g)_{\mathcal{B}''}^{\mathcal{B}'} = (\beta_{ki})_{k=1,p;i=1,m} \qquad \text{unde} \qquad g(\mathbf{v}_i) = \sum_{k=1}^{p} \beta_{ki}\mathbf{w}_k$$

atunci

$$\begin{aligned}(g \circ f)(\mathbf{u}_j) &= g(f(\mathbf{u}_j)) = g\Big(\sum_{i=1}^{m} \alpha_{ij}\mathbf{v}_i\Big) = \sum_{i=1}^{m} \alpha_{ij} g(\mathbf{v}_i) \\ &= \sum_{i=1}^{m} \alpha_{ij} \sum_{k=1}^{p} \beta_{ki}\mathbf{w}_k = \sum_{k=1}^{p} \underbrace{\left(\sum_{i=1}^{m} \beta_{ki}\alpha_{ij}\right)}_{\left(M(g\circ f)_{\mathcal{B}''}^{\mathcal{B}}\right)_{kj}} \mathbf{w}_k\end{aligned}$$

Rezultă deci $M(g \circ f)_{\mathcal{B}''}^{\mathcal{B}} = M(g)_{\mathcal{B}''}^{\mathcal{B}'} M(f)_{\mathcal{B}'}^{\mathcal{B}}$.

5) "$\Rightarrow$" Fie $f : U \to V$ un izomorfism. În particular, $\dim_{\Bbbk} U = \dim_{\Bbbk} V$. Să notăm cu n dimensiunea comună. Atunci $M(f)$ este o matrice pătratică. Fie f^{-1} inversa lui f; atunci din $f \circ f^{-1} = \mathsf{Id}_V$ şi $f^{-1} \circ f = \mathsf{Id}_U$ rezultă, aplicând subpunctele 3) şi 4) anterioare, că $M(f)M(f^{-1}) = I_n$ şi $M(f^{-1})M(f) = I_n$, deci $M(f)$ este inversabilă cu inversa $M(f^{-1})$.

"$\Leftarrow$" Dacă $M(f)$ este inversabilă, în particular este o matrice pătratică, deci $\dim_{\Bbbk} U = \dim_{\Bbbk} V$. Din nou, fie n valoarea comună a dimensiunii. Va fi suficient să demonstrăm că f este injectivă (echivalent, că $\mathsf{Ker}(f) = \{\mathbf{0}\}$) şi să aplicăm teorema 5.11. Fie atunci $\mathbf{u} \in \mathsf{Ker}(f)$. Atunci $f(\mathbf{u}) = \mathbf{0}$. Trecând la coordonate, obţinem $\mathbf{0}_n = [f(\mathbf{u})]_{\mathcal{B}'} = M(f)[u]_{\mathcal{B}}$. Dar $M(f)$ este inversabilă, de unde rezultă $[\mathbf{u}]_{\mathcal{B}} = \mathbf{0}_n$, adică $\mathbf{u} = \mathbf{0}_U$.

□

Observaţie 5.16. 1) Ultimul rezultat de mai sus produce un criteriu practic de decizie dacă o aplicaţie liniară este un izomorfism sau nu: este suficient să testăm dacă matricea asociată acesteia este inversabilă.

2) Cu toate acestea, demonstraţia pleacă de la baze prestabilite atât în domeniu cât şi în codomeniu; ce se întâmplă dacă schimbăm bazele? Vom răspunde în continuare la această întrebare.

   Fie U, V spaţii vectoriale finit dimensionale şi $\mathcal{B}, \mathcal{C}$ baze în U, respectiv $\mathcal{B}', \mathcal{C}'$ baze în V. În sfârşit, fie $f : U \to V$ o aplicaţie liniară. Atunci avem următoarea situaţie:

$$\begin{tikzcd} U \arrow[r, "\mathsf{Id}_U"] \arrow[rrr, bend left, "f=\mathsf{Id}_V\circ f\circ\mathsf{Id}_U"] & U \arrow[r, "f"] & V \arrow[r, "\mathsf{Id}_V"] & V \end{tikzcd}$$

$$\begin{array}{ccccccc} \mathcal{C} & & \mathcal{B} & & \mathcal{B}' & & \mathcal{C}' \\ f & = & \mathsf{Id}_V & \circ & f & \circ & \mathsf{Id}_U \end{array}$$

   Trecând la matricele asociate aplicaţiilor liniare în bazele corespunzător indicate mai sus, obţinem

$$\begin{aligned} M(f)^{\mathcal{C}}_{\mathcal{C}'} &= M(\mathsf{Id}_V)^{\mathcal{B}'}_{\mathcal{C}'} \cdot M(f)^{\mathcal{B}}_{\mathcal{B}'} \cdot M(\mathsf{Id}_U)^{\mathcal{C}}_{\mathcal{B}} \\ &= P^{\mathcal{B}'}_{\mathcal{C}'} \cdot M(f)^{\mathcal{B}}_{\mathcal{B}'} \cdot P^{\mathcal{C}}_{\mathcal{B}} \\ &= \left(P^{\mathcal{C}'}_{\mathcal{B}'}\right)^{-1} \cdot M(f)^{\mathcal{B}}_{\mathcal{B}'} \cdot P^{\mathcal{C}}_{\mathcal{B}} \end{aligned}$$

   Rezultă următoarea formulă de schimbare a matricei asociate unei aplicaţii liniare la schimbarea bazelor:

$$M(f)^{\mathcal{C}}_{\mathcal{C}'} = \left(P^{\mathcal{C}'}_{\mathcal{B}'}\right)^{-1} \cdot M(f)^{\mathcal{B}}_{\mathcal{B}'} \cdot P^{\mathcal{C}}_{\mathcal{B}} \tag{11}$$

Observaţie 5.17. Fie V un spaţiu vectorial finit dimensional şi $\mathcal{B}$ o bază pentru V. Fie $f : V \to V$ o aplicaţie liniară (un endomorfism). Pentru a simplifica notaţia, în loc de $M(f)^{\mathcal{B}}_{\mathcal{B}}$, vom nota cu $M(f)_{\mathcal{B}}$ matricea (pătratică) asociată lui f în raport cu baza $\mathcal{B}$.

Dacă $\mathcal{B}'$ este altă bază pentru V şi $P = P^{\mathcal{B}'}_{\mathcal{B}}$ matricea de trecere de la baza $\mathcal{B}$ la baza $\mathcal{B}'$, atunci aplicând rezultatul precedent (11), rezultă că matricea asociată lui f în raport cu baza $\mathcal{B}'$ este

$$M(f)_{\mathcal{B}'} = P^{-1} M(f)_{\mathcal{B}} P \tag{12}$$

Vom utiliza această formulă în următorul capitol.

6 Matrice şi endomorfisme diagonalizabile

6.1 Valori şi vectori proprii

De multe ori, pentru a evidenţia efectul pe care îl are o multiplicarea cu o matrice (pătratică) asupra vectorilor coloană, sau, mai general, efectul unei aplicaţii liniare, baza în care se lucrează (de regulă, baza canonică) nu este cea mai adecvată. Apare necesitatea determinării unei baze convenabile pentru calcule (presupunând că o asemenea bază există).

În această secţiune vom considera doar matrice pătratice (endomorfisme), urmând ca în Secţiunea următoare să tratăm şi cazul matricelor arbitrare.

Vom ilustra aplicabilitatea rezultatelor dezvoltate aici prin (cel puţin!) două exemple: algoritmul PageRank şi automatele probabilistice (lanţurile Markov). Vom vedea şi mai multe exemple după ce vom studia noţiunile de produs scalar, ortogonalitate şi proiecţie.

Matrice diagonală/diagonalizabilă

Definiţie 6.1. O matrice $A \in \mathcal{M}_n(\Bbbk)$ este

1) **diagonală** dacă $A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$

2) **diagonalizabilă** dacă există $P \in \mathcal{M}_n(\Bbbk)$ inversabilă astfel încât $A = PDP^{-1}$, unde D este o matrice diagonală.

Întrucât fiecare matrice determină o aplicaţie liniară, şi invers, fiecărui aplicaţii liniare îi corespunde o matrice (în raport cu o bază pe domeniu, respectiv o bază pe codomeniu), putem extinde definiţia de mai sus cu uşurintă şi pentru aplicaţiile liniare. Reamintim că o aplicaţie liniară pentru care domeniul şi codomeniul coincid, adică de tipul $f : V \to V$, se numeşte **endomorfism**.

Endomorfism diagonalizabil

Definiţie 6.2. Fie V un $\Bbbk$-spaţiu vectorial. Un endomorfism $f : V \to V$ este **diagonalizabil** dacă există o bază $\mathcal{B}$ în raport cu care matricea asociată $M(f)$ este diagonală.

Rezultatele pe care le vom obţine în această secţiune vor răspunde la următoarele întrebări (şi nu numai):

1) În ce condiţii o matrice $A \in \mathcal{M}_n(\Bbbk)$ este **diagonalizabilă**?

2) Dacă o matrice A este diagonalizabilă, cum determinăm matricele P şi D astfel încât $A = PDP^{-1}$? Sunt acestea unice?

3) Ce aplicaţii au matricele/endomorfismele diagonalizabile?

Valori şi vectori proprii

Definiţie 6.3. Fie $A \in \mathcal{M}_n(\mathbb{k})$ o matrice. Spunem că $\lambda \in \mathbb{k}$ este **valoare proprie** pentru matricea A dacă există un vector $\mathbf{v} \in \mathbb{k}^n, \mathbf{v} \neq \mathbf{0}_n$ astfel încât

$$A\mathbf{v} = \lambda\mathbf{v}$$

Vectorul $\mathbf{v}$ se numeşte **vector propriu** asociat valorii proprii λ.

Analog, dacă $f : V \to V$ este un endomorfism, atunci $\lambda \in \mathbb{k}$ este **valoare proprie** pentru f dacă există un vector $\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}_V$ astfel încât $f(\mathbf{v}) = \lambda\mathbf{v}$. Vom numi ca mai sus vectorul $\mathbf{v} \in V$ **vector propriu** asociat valorii proprii $\lambda \in \mathbb{k}$.

Observaţie 6.4. 1) Mulţimea valorilor proprii ale unei matrice A se numeşte **spectrul matricei** şi se notează cu $\sigma(A)$.

2) Se verifică cu uşurinţă că mulţimea $V_\lambda = \{\mathbf{v} \in \mathbb{k}^n \mid A\mathbf{v} = \lambda\mathbf{v}\}$ este $\mathbb{k}$-subspaţiu vectorial netrivial[14] în $\mathbb{k}^n$, numit **subspaţiul vectorilor proprii** corespunzători valorii proprii $\lambda \in \mathbb{k}$.[15] Analog, pentru un endomorfism $f : V \to V$, notăm cu $V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}$ subspaţiul vectorilor proprii asociaţi valorii proprii λ.

3) Fie $f : V \to V$ un endomorfism şi $\lambda \in \mathbb{k}$ o valoare proprie pentru f, cu $\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}_V$ vector propriu asociat. Să presupunem că V este un $\mathbb{k}$-spaţiu vectorial finit dimensional cu $\dim_{\mathbb{k}} V = n$ şi că $\mathcal{B}$ este o bază în V. Dacă $M(f)$ este matricea asociată lui f în raport cu această bază, atunci trecând la coordonate, din relaţia $f(\mathbf{v}) = \lambda\mathbf{v}$ rezultă $M(f)[\mathbf{v}]_\mathcal{B} = [f(\mathbf{v})]_\mathcal{B} = \lambda[\mathbf{v}]_\mathcal{B}$. Întrucât $\mathbf{v} \neq \mathbf{0}_V \iff [\mathbf{v}]_\mathcal{B} \neq \mathbf{0}_n$, rezultă că λ este valoare proprie şi pentru matricea $M(f)$ şi reciproc. Deci endomorfismul f şi matricea asociată $M(f)$ au aceleaşi valori proprii. Din acest motiv, în continuare, vom discuta în principal cazul matricelor, având în vedere că aceleaşi rezultate sunt valabile şi în cazul aplicaţiilor liniare (endomorfismelor).

Exemplu 6.5. 1) Fie $y = \lambda x$, $\lambda \neq 0$ ecuaţia unei drepte d din plan care trece prin origine, şi A matricea asociată (aplicaţiei liniare) proiecţiei pe dreapta d, ca în Secţiunea 5.1, punctul 2. Atunci orice vector $\mathbf{v} \in \mathbb{R}^2$ a cărui dreaptă suport este d, este vector propriu pentru matricea A, corespunzător valorii proprii $\lambda_1 = 1$ (deoarece $A\mathbf{v} = \mathbf{v}$), respectiv orice vector $\mathbf{w} \in \mathbb{R}^2$ ortogonal pe d este vector propriu pentru matricea A, corespunzător valorii proprii $\lambda_2 = 0$ ($A\mathbf{w} = \mathbf{0} = 0\mathbf{w}$). Mai mult chiar, subspaţiile proprii asociate V_{λ_1} şi V_{λ_2} pot fi interpretate geometric ca dreapta d, respectiv dreapta de ecuaţie $y = -\frac{1}{\lambda}x$ (dreapta care trece prin origine şi este perpendiculară pe d).

2) Fie V spaţiul vectorial al funcţiilor $\mathbb{R} \to \mathbb{R}$ de clasă $\mathcal{C}^\infty$ şi $f : V \to V$ operatorul de derivare, $F(\varphi) = \varphi'$.[16] Atunci funcţia exponenţială $\varphi : \mathbb{R} \to \mathbb{R}, \varphi(x) = \mathsf{e}^{\lambda x}$ (unde $\lambda \in \mathbb{R}$) este vector propriu pentru f, corespunzător valorii proprii λ. Puteţi găsi alţi vectori proprii?

3) În acest exemplu considerăm endomorfismul $f : \mathcal{M}_n(\mathbb{k}) \to \mathcal{M}_n(\mathbb{k})$, $f(X) = X^t$. Atunci orice matrice simetrică $X = X^t$ este vector propriu pentru f corespunzător valorii proprii $\lambda_1 = 1$. Analog, orice matrice antisimetrică $X = -X^t$ este vector propriu corespunzător valorii proprii $\lambda_2 = -1$. Există şi alte valori proprii pentru f?

[14] Adică diferit de subspaţiul nul $\{\mathbf{0}_V\}$.

[15] Atenţie, V_λ conţine şi vectorul nul, dar acesta nu este considerat a fi vector propriu!

[16] În matematică şi nu numai, termenul de **operator** se utilizează pentru funcţii care primesc la rândul lor funcţii ca argument.

În exemplele de mai sus, vectorii şi valorile proprii au fost determinate prin observaţii directe ale particularităţilor matricelor, respectiv endomorfismelor considerate. Vom stabili mai jos o metodă generală de calcul al valorilor proprii.

Determinarea valorilor proprii ale unei matrice

Propoziţie 6.6. Fie $A \in \mathcal{M}_n(\Bbbk)$. Atunci $\lambda \in \Bbbk$ este valoare proprie pentru matricea A dacă şi numai dacă $\det(A - \lambda I_n) = 0$.

Demonstraţie. $\lambda \in \Bbbk$ este valoare proprie pentru matricea A dacă şi numai dacă există $\mathbf{v} \in \Bbbk^n$, $\mathbf{v} \neq 0$, astfel încât $A\mathbf{v} = \lambda\mathbf{v}$, dacă şi numai dacă sistemul liniar şi omogen $(A - \lambda I_n)\mathbf{v} = \mathbf{0}$ admite şi soluţii nebanale, dacă şi numai dacă matricea sistemului este neinversabilă, adică $\det(A - \lambda I_n) = 0$. □

Să observăm că expresia $\det(A - \lambda I_n)$ este o funcţie polinomială în λ. Mai precis:

Polinomul caracteristic

Definiţie 6.7. Fie $A \in \mathcal{M}_n(\Bbbk)$, $A = (a_{ij})_{i,j=1,n}$. Polinomul

$$P_A(X) = \det(A - XI_n) = \begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - X & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - X \end{vmatrix} \in \Bbbk[X]$$

se numeşte **polinomul caracteristic** al matricei A.

Observaţie 6.8. Polinomul caracteristic al unei matrice $A \in \mathcal{M}_n(\Bbbk)$ este un polinom având gradul n, iar coeficientul termenului de grad maxim X^n este $(-1)^n$. Termenul liber al polinomului se determină uşor observând că acesta este $P_A(0) = \det(A)$, iar coeficientul termenului X^{n-1} este $(-1)^{n-1}\operatorname{tr}(A)$. Deci

$$P_A(X) = (-1)^n \left(X^n - \operatorname{tr}(A)X^{n-1} + \ldots + \det(A)\right)$$

În particular, pentru $n = 2$, $P_A(X) = X^2 - \operatorname{tr}(A)X + \det(A)$. Conform propoziţiei precedente, valorile proprii ale matricei A sunt exact rădăcinile polinomului său caracteristic $P_A(X)$. Determinarea rădăcinilor unui polinom depinde foarte mult de corpul $\Bbbk$ peste care se lucrează, după cum vom vedea în exemplele următoare.

Exemplu 6.9. Fie $A \in \mathcal{M}_2(\mathbb{R})$, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Atunci $P_A(X) = \begin{vmatrix} -X & -1 \\ 1 & -X \end{vmatrix} = X^2 + 1$ este un polinom fără rădăcini reale. Deci A nu are valori proprii reale, şi implicit nici vectori proprii cu componentele numere reale. Acest rezultat ar fi putut fi obţinut şi direct, plecând de la observaţia că A este matricea de rotaţie cu un unghi de $90°$ în sens trigonometric (Secţiunea 5.1, punctul 1), şi că orice vector propriu $\mathbf{v}$ este coliniar cu imaginea sa $A\mathbf{v}$ (deoarece $A\mathbf{v} = \lambda\mathbf{v}$).

Pe de altă parte, dacă considerăm A o matrice complexă, $A \in \mathcal{M}_2(\mathbb{C})$, atunci A admite valorile proprii $\pm\mathrm{i}$ (Exerciţiu: determinaţi vectorii proprii. Ce observaţi?). Am utilizat aici proprietatea fundamentală a corpului numerelor complexe de a fi algebric închis.[17]

[17] Un corp $\Bbbk$ se numeşte **algebric închis** dacă orice polinom $P \in \Bbbk[X]$ admite cel puţin o rădăcină în $\Bbbk$.

Propoziţie 6.10. Fie $A, B \in \mathcal{M}_n(\Bbbk)$ cu proprietatea că există o matrice inversabilă $P \in \mathcal{M}_n(\Bbbk)$ astfel încât $A = P^{-1}BP$ (spunem că matricele A şi B sunt **similare**). Atunci A şi B au acelaşi polinom caracteristic.

Demonstraţie. Rezultă din următoarele calcule:

$$\begin{aligned} P_A(X) &= \det(A - XI_n) &&= \det\big(P^{-1}BP - XI_n\big) \\ &= \det P^{-1}(B - XI_n)P &&= \det P^{-1} \det(B - XI_n) \det P \\ &= P_B(X) \end{aligned}$$

□

Observaţie 6.11. 1) Deci două matrice similare au acelaşi polinom caracteristic, aceleaşi valori proprii (rădacinile polinomului caracteristic), acelaşi determinant, aceeaşi urmă şi acelaşi rang. Dar reciproc? (a se vedea exerciţiul următor)

2) În Observaţia 6.4.3, am văzut că un endomorfism $f : V \to V$ admite aceleaşi valori proprii ca şi matricea asociată $M(f)$ în raport cu o bază $\mathcal{B}$. Dacă utilizăm o altă bază $\mathcal{B}'$, atunci matricea asociată aplicaţiei liniare în baza $\mathcal{B}'$ va fi, conform formulei (12),

$$M(f)_{\mathcal{B}'} = P^{-1}M(f)_{\mathcal{B}}P$$

adică $M(f)_{\mathcal{B}}$ şi $M(f)_{\mathcal{B}'}$ sunt matrice similare. Rezultatul de mai sus ne spune că indiferent în ce bază lucrăm (pentru a determina matricea asociată aplicaţiei linare f), vom obţine mereu aceleaşi valori proprii.

Exemplu 6.12. Cercetaţi dacă matricele $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ şi $B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ sunt similare.

Teorema Hamilton-Cayley

Teoremă 6.13. Orice matrice pătratică $A \in \mathcal{M}_n(\Bbbk)$ satisface relaţia

$$P_A(A) = O_n$$

Demonstraţie. Fără demonstraţie. □

6.2 Criteriul de diagonalizare

Valori proprii şi multiplicităţi

Definiţie 6.14. Fie $A \in \mathcal{M}_n(\Bbbk)$ o matrice pătratică şi $\lambda \in \Bbbk$ o valoare proprie pentru A.

Vom numi **multiplicitate algebrică**, notată $a(\lambda)$, multiplicitatea lui λ ca rădăcină a polinomului caracteristic $P_A(X)$.

De asemenea, vom numi **multiplicitate geometrică** a valorii proprii λ, şi vom nota $g(\lambda)$, dimensiunea subspaţiului vectorilor proprii V_λ.

Aşa cum în Secţiunea precedentă am văzut că putem caracteriza complet un spaţiu vectorial doar cu ajutorul unei caracteristici numerice (dimensiunea spaţiului vectorial), şi aici vom utiliza cele două tipuri de multiplicităţi introduse mai sus pentru a decide când o matrice este diagonalizabilă.

Criteriul de diagonalizare

Teoremă 6.15. Fie $A \in \mathcal{M}_n(\Bbbk)$ o matrice având polinomul caracteristic

$$P_A(X) = (-1)^n (X - \lambda_1)^{a_1} \cdot \ldots \cdot (X - \lambda_p)^{a_p} \in \Bbbk[X]$$

unde $a_1 + \ldots + a_p = n$. Atunci următoarele afirmaţii sunt echivalente:

1) Matricea A este diagonalizabilă.

2) Există o bază în $\Bbbk^n$ formată cu vectori proprii ai matricei A.

3) Pentru orice valoare proprie λ a matricei A, multiplicitatea algebrică $a(\lambda)$ şi cea geometrică $g(\lambda)$ coincid.

Demonstraţia acestei teoreme necesită o serie de rezultate adiţionale, fiecare având importanţa sa, pe care le vom verifica şi exemplifica pas cu pas. Însă înainte de aceasta, să consemnăm următoarea consecinţă imediată a Criteriului de diagonalizare, ale cărei aplicaţii le vom discuta în secţiunile următoare:

Corolar 6.16. Dacă matricea $A \in \mathcal{M}_n(\Bbbk)$ are n valori proprii distincte în $\Bbbk$, atunci este diagonalizabilă.

Demonstraţie. Exerciţiu. □

Lemă 6.17. Fie $A \in \mathcal{M}_n(\Bbbk)$ o matrice pătratică şi $\lambda_1, \ldots \lambda_p \in \Bbbk$ valori proprii **distincte** ale matricei A. Pentru fiecare $i = \overline{1,p}$, fie $\mathbf{v}_i \in V$ un vector propriu corespunzător valorii proprii λ_i. Atunci vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_p$ sunt **liniar independenţi**.

Demonstraţie. Vom demonstra prin inducţie după p. Dacă $p = 1$, atunci avem un singur vector propriu $\mathbf{v}$ corespunzător unei valori proprii $\lambda \in \Bbbk$: $A\mathbf{v} = \lambda\mathbf{v}$. În particular, din definiţie rezultă că $\mathbf{v}$ este nenul, adică liniar independent. Să presupunem acum afirmaţia adevărată pentru orice p vectori proprii şi să o verificăm pentru $p+1$ vectori proprii. Fie deci $\lambda_1, \ldots, \lambda_{p+1} \in \Bbbk$ valori proprii distincte pentru matricea A şi $\mathbf{v}_1, \ldots, \mathbf{v}_{p+1} \in V$ vectori proprii corespunzători valorilor proprii $\lambda_1, \ldots, \lambda_p$, respectiv λ_{p+1}. Fie $\alpha_1, \ldots, \alpha_{p+1} \in \Bbbk$ astfel încât

$$\alpha_1\mathbf{v}_1 + \ldots + \alpha_{p+1}\mathbf{v}_{p+1} = \mathbf{0}$$

Multiplicând la stânga cu matricea A în relaţia de mai sus şi ţinând cont de relaţiile $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, $i = \overline{1, p+1}$, rezultă

$$\alpha_1\lambda_1\mathbf{v}_1 + \ldots + \alpha_{p+1}\lambda_{p+1}\mathbf{v}_{p+1} = \mathbf{0}$$

Revenind la combinaţia liniară nulă iniţială, şi multiplicând cu λ_{p+1} în loc de matricea A, se obţine

$$\alpha_1\lambda_{p+1}\mathbf{v}_1 + \ldots + \alpha_{p+1}\lambda_{p+1}\mathbf{v}_{p+1} = \mathbf{0}$$

Scăzând termen cu termen ultimele două relaţii, avem

$$\alpha_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \ldots + \alpha_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0}$$

Dar, conform ipotezei de inducţie, vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_p$ sunt liniar independenţi. Deci $\alpha_1(\lambda_1 - \lambda_{p+1}) = 0, \ldots, \alpha_p(\lambda_p - \lambda_{p+1}) = 0$. Deoarece valorile proprii $\lambda_1, \ldots, \lambda_{p+1}$ sunt distincte, rezultă $\alpha_= 0, \ldots, \alpha_p = 0$, de unde obţinem şi $\alpha_{p+1} = 0$. Deci vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_{p+1}$ sunt liniar independenţi. □

Exemplu 6.18. Fie V spaţiul vectorial al funcţiilor $\mathbb{R} \to \mathbb{R}$ de clasă $\mathcal{C}^\infty$ şi $f : V \to V$, $f(\varphi) = \varphi'$ operatorul liniar de derivare. Fie $\lambda_1, \ldots, \lambda_p$ numerele reale **distincte**. Atunci funcţiile $\mathsf{e}^{\lambda_1 x}, \ldots, \mathsf{e}^{\lambda_p x}$ sunt vectori proprii pentru f corespunzători valorilor proprii distincte $\lambda_1, \ldots, \lambda_p$, deci sunt liniar independente.

Analog, funcţiile $1, \sin x, \sin 2x, \ldots, \sin nx$, sunt vectori proprii pentru $f \circ f$ (derivare de două ori), deci în particular sunt liniar independente.

Observaţie 6.19. În situaţia descrisă de Lema 6.17, dacă $p = n$, atunci vectorii proprii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ formează o bază în $\mathbb{k}^n$.

Lemă 6.20. Fie $\lambda_1, \ldots \lambda_p \in \mathbb{k}$ valori proprii **distincte** ale matricei A şi, pentru fiecare $i = \overline{1, p}$, fie $\mathcal{B}_i$ o bază în V_{λ_i}.
Atunci $\mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_p$ este o mulţime de vectori **liniar independenţi**.

Demonstraţie. Acesta este de fapt generalizarea rezultatului demonstrat în Lema precedentă.

Fie $g_i = \dim_{\mathbb{k}} V_{\lambda_i}$ (multiplicitatea geometrică a valorii proprii λ_i) şi $\mathcal{B}_i = \{\mathbf{v}_i^1, \ldots, \mathbf{v}_i^{g_i}\}$ o bază în V_{λ_i}, pentru fiecare $i \in \{1, \ldots, p\}$.

Să observăm că mulţimile $\mathcal{B}_i$ sunt disjuncte două câte două: dacă ar exista $\mathbf{v} \in \mathcal{B}_i \cap \mathcal{B}_j$, pentru $i \neq j$, atunci $\lambda_i \mathbf{v} = A\mathbf{v} = \lambda_j \mathbf{v}$, de unde $(\lambda_i - \lambda_j)\mathbf{v} = \mathbf{0}$. Dar $\lambda_i \neq \lambda_j$ şi $\mathbf{v} \neq 0$, contradicţie.

Fie acum $\alpha_i^j \in \mathbb{k}$ astfel încât $\sum_{i=1}^{p} \sum_{j=1}^{g_i} \alpha_i^j \mathbf{v}_i^j = \mathbf{0}$. Notând $\mathbf{w}_i = \sum_{j=1}^{g_i} \alpha_i^j \mathbf{v}_i^j$, avem $\mathbf{w}_i \in V_{\lambda_i}$ pentru orice $i = \overline{1, p}$ şi $\sum_{i=1}^{p} \mathbf{w}_i = \mathbf{0}$. Dar dacă $\mathbf{w}_i$ sunt vectori proprii (ştim deja că $\mathbf{w}_i \in V_{\lambda_i}$, aşa că informaţia "$\mathbf{w}_i$ este vector propriu" este echivalentă cu $\mathbf{w}_i \neq \mathbf{0}$), atunci conform Lemei precendente, vectorii $\mathbf{w}_1, \ldots, \mathbf{w}_p$ sunt liniar independenţi. Dar $\sum_{i=1}^{p} \mathbf{w}_i = \mathbf{0}$, contradicţie. Deci pentru orice $i = \overline{1, p}$, $\mathbf{w}_i = \mathbf{0}$.

Să ne reamintim că fiecare vector (nul) $\mathbf{w}_i$ este $\sum_{j=1}^{g_i} \alpha_i^j \mathbf{v}_i^j$, şi că $\mathcal{B}_i = \{\mathbf{v}_i^1, \ldots, \mathbf{v}_i^{g_i}\}$ este o bază în V_{λ_i}. Rezultă atunci că toţi scalariii α_i^j sunt nuli, adică vectorii din $\mathcal{B}_1 \cup \ldots \mathcal{B}_p$ sunt liniar independenţi. □

Exemplu 6.21. Să determinăm valorile proprii şi vectorii proprii (inclusiv o bază pentru fiecare dintre subspaţiile proprii) pentru matricea $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix}$.

Polinomul caracteristic este

$$P_A(X) = \begin{vmatrix} 2-X & 1 & 0 \\ 0 & 1-X & -1 \\ 0 & 2 & 4-X \end{vmatrix} = -(X-2)^2(X-3)$$

deci valorile proprii sunt $\lambda_1 = 2$ şi $\lambda_2 = 3$ cu multiplicităţile (algebrice) $a(\lambda_1) = 2$, respectiv $a(\lambda_2) = 1$.

Pentru determinarea (subspaţiilor) vectorilor proprii, să reamintim că un vector nenul $\mathbf{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ este vector propriu pentru matricea A dacă şi numai dacă $A\mathbf{v} = \lambda\mathbf{v}$, unde $\lambda \in \{\lambda_1, \lambda_2\}$, sau echivalent, dacă

$$\begin{cases} 2x + y = \lambda x \\ y - z = \lambda y \\ 2y + 4z = \lambda z \end{cases}$$

Rezolvând pe rând pentru λ_1 şi λ_2 sistemul de mai sus, obţinem:

$$\lambda_1 = 2 \quad \Rightarrow \quad V_{\lambda_1} = \{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \mid x \in \mathbb{R}\} = \mathsf{Sp}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \quad \Rightarrow \quad g(\lambda_1) = 1$$

$$\lambda_2 = 3 \quad \Rightarrow \quad V_{\lambda_2} = \{ \begin{pmatrix} y \\ y \\ -2y \end{pmatrix} \mid y \in \mathbb{R}\} = \mathsf{Sp}\left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\} \quad \Rightarrow \quad g(\lambda_2) = 1$$

Ultimul rezultat pe care îl vom demonstra se referă la caracteristicile numerice asociate fiecărei valori proprii:

Lemă 6.22. Pentru orice valoare proprie λ a matricei A, multiplicitatea geometrică este mai mică sau cel mult egală cu multiplicitatea sa algebrică.

Demonstraţie. Fie $g = \dim_{\Bbbk} V_\lambda$ multiplicitatea geometrică şi $\{\mathbf{v}_1, \ldots, \mathbf{v}_g\}$ o bază în V_λ (să ne reamintim că V_λ este un subspaţiu în $\Bbbk^n$, deci în particular $g \leq n$).

Completăm la o bază $\{\mathbf{v}_1, \ldots, \mathbf{v}_g, \ldots, \mathbf{v}_n\}$ în $\Bbbk^n$. Atunci multiplicând cu matricea A la stânga fiecare vector din bază, putem scrie

$$A\mathbf{v}_1 = \lambda\mathbf{v}_1 \qquad \ldots \qquad A\mathbf{v}_g = \lambda\mathbf{v}_g \qquad A\mathbf{v}_{g+1} = \ldots \qquad \ldots \qquad A\mathbf{v}_n = \ldots$$

sau echivalent, în formă matriceală

$$A\underbrace{\left(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_g \mid \mathbf{v}_{g+1} \mid \ldots \mid \mathbf{v}_n \right)}_{P} = \underbrace{\left(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_g \mid \mathbf{v}_{g+1} \mid \ldots \mid \mathbf{v}_n \right)}_{P} \underbrace{\left(\begin{array}{cccc|c} \lambda & 0 & \ldots & 0 & \ldots \\ 0 & \lambda & \ldots & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda & \ldots \\ \hline 0 & 0 & \ldots & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \ldots \end{array}\right)}_{B}$$

Dacă notăm cu P matricea având coloanele $\mathbf{v}_1, \ldots, \mathbf{v}_n$, atunci P este o matrice inversabilă (vectorii formează o bază), şi avem relaţia $A = P^{-1}BP$, unde B este matricea de mai sus. Cu alte cuvinte, A şi B sunt matrice similare, deci conform Propoziţiei 6.10, au acelaşi polinom caracteristic:

$$P_A(X) = P_B(X) = (\lambda - X)^g Q(X)$$

unde $Q \in \Bbbk[X]$. Rezultă că ordinul de multiplicitate al valorii proprii λ ca rădăcină a polinomului caracteristic este cel puţin g, adică $g = g(\lambda) \leq a(\lambda)$. $\square$

Demonstraţie. (Teorema 6.15) 1) $\Leftrightarrow$ 2). A este diagonalizabilă $\Longleftrightarrow$ există o matrice $P \in \mathcal{M}_n(\Bbbk)$ inversabilă astfel încât $P^{-1}AP$ să fie o matrice diagonală $\begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_n \end{pmatrix}$. Dacă notăm $P = (\mathbf{v}_1|\dots|\mathbf{v}_n)$, atunci P inversabilă este echivalent cu $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ bază în $\Bbbk^n$, şi $P^{-1}AP = \begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_n \end{pmatrix}$ este echivalent cu

$$A\,(\mathbf{v}_1|\dots|\mathbf{v}_n) = AP = PD = (\mathbf{v}_1|\dots|\mathbf{v}_n) \begin{pmatrix} \lambda_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \lambda_n \end{pmatrix} = (\lambda_1\mathbf{v}_1|\dots|\lambda_n\mathbf{v}_n)$$

Rezumând, o matrice A este diagonalizabilă dacă şi numai dacă există o bază $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ în $\Bbbk^n$ şi scalarii $\lambda_1, \dots, \lambda_n \in \Bbbk$ astfel încât

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \dots, A\mathbf{v}_n = \lambda_n\mathbf{v}_n$$

adică vectorii $\mathbf{v}_1, \dots, \mathbf{v}_n$ sunt vectori proprii pentru matricea A.

2) $\Rightarrow$ 3) Fie $\lambda_1, \dots, \lambda_p \in \Bbbk$ valorile proprii **distincte** ale matricei A, cu multiplicităţile algebrice $a_1, \dots, a_p$, respectiv multiplicităţile geometrice $g_1, \dots, g_p$.

Conform ipotezei, există o bază în $\Bbbk^n$ formată cu vectori proprii ai matricei A. Grupând vectorii proprii din această bază după valorile proprii corespunzătoare, observăm că cel mult g_1 dintre vectori se găsesc în V_{λ_1}, ..., cel mult g_p dintre vectori în V_{λ_p}. Dar baza conţine n vectori, deci

$$n \leq g_1 + \dots + g_p \leq a_1 + \dots + a_p = n$$

Dar $g_1 \leq a_1, \dots, g_p \leq a_p$, de unde obţinem rezultatul dorit:

$$g_1 = a_1, \dots, g_p = a_p$$

3) $\Rightarrow$ 2) Pentru valorile proprii $\lambda_1, \dots, \lambda_p$, fie $\mathcal{B}_1, \dots, \mathcal{B}_p$ baze în subspaţiile proprii corespunzătoare. Conform Lemei 6.20, $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_p$ este o mulţime de vectori liniar independenţi, şi întrucât bazele sunt disjuncte două câte două (a se vedea demonstraţia Lemei 6.20), $|\mathcal{B}_1 \cup \dots \cup \mathcal{B}_p| = |\mathcal{B}_1| + \dots + |\mathcal{B}_p| = g(\lambda_1) + \dots + g(\lambda_p) = a(\lambda_1) + \dots + a(\lambda_p) = n$. Rezultă că $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_p$ este o bază în $\Bbbk^n$ (formată cu vectori proprii). $\square$

6.3 Diagonalizarea matricelor 2x2

Una dintre cele mai utilizate aplicaţii ale algoritmului de diagonalizare este calculul puterilor matricelor pătratice.

În această secţiune, vom discuta cazul matricelor reale cu 2 linii şi 2 coloane. Fie $A \in \mathcal{M}_2(\mathbb{R})$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Atunci polinomul caracteristic al matricei A este

$$P_A(X) = \det(A - XI_2) = \begin{vmatrix} a - X & b \\ c & d - X \end{vmatrix} = X^2 - \operatorname{tr}(A)X + \det(A)$$

şi în funcţie de natura rădăcinilor (valorilor proprii), distingem următoarele trei cazuri:

1) $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$.
2) $\lambda_1 = \lambda_2 \in \mathbb{R}$.
3) $\lambda_1, \lambda_2 \in \mathbb{C} - \mathbb{R}$, $\lambda_2 = \overline{\lambda_1}$.

Vom analiza pe rând fiecare dintre cele trei cazuri.

1) $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$.

   În acest caz, matricea admite valori proprii reale şi simple (de multiplicitate algebrică 1). Atunci conform Corolarului 6.16, A este diagonalizabilă peste corpul numerelor reale.

   Dacă $V_{\lambda_1} = \mathsf{Sp}\{\mathbf{v}_1\}$ şi $V_{\lambda_2} = \mathsf{Sp}\{\mathbf{v}_2\}$, atunci din relaţiile

   $$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \qquad \text{şi} \qquad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$$

   rezultă

   $$A\underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P} = (\lambda_1\mathbf{v}_1|\lambda_2\mathbf{v}_2) = \underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P}\underbrace{\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}}_{D}$$

   sau, echivalent, $A = PDP^{-1}$. În acest caz, $A^n = PD^nP^{-1} = P\begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}P^{-1}$.

2) $\lambda_1 = \lambda_2 \in \mathbb{R}$. Vom nota simplu valoarea proprie cu λ.

   În această situaţie distingem două cazuri:

   (a) Dacă $\dim_{\mathbb{R}} V_\lambda = 2$, atunci multiplicităţile algebrice şi geometrice coincid, deci A este diagonalizabilă. Punând $V_\lambda = \mathsf{Sp}\{\mathbf{v}_1, \mathbf{v}_2\}$, rezultă ca mai sus $A = PDP^{-1}$, unde $P = (\mathbf{v}_1|\mathbf{v}_2)$ şi $D = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \lambda I_2$. Atunci $A = PDP^{-1} = \lambda I_2$.

   (b) Dacă $\dim_{\mathbb{R}} V_\lambda = 1 < a(\lambda) = 2$, atunci A nu este diagonalizabilă. În această situaţie, vom proceda în felul următor:

   - Să observăm că există cel puţin un vector $\mathbf{v}_2 \in \mathbb{R}^2$ (vom justifica ulterior notaţia) astfel încât vectorul $(A - \lambda I_2)\mathbf{v}_2$ să fie nenul: dacă prin absurd $(A - \lambda I_2)\mathbf{v}_2 = \mathbf{0}$ pentru orice $\mathbf{v}_2 \in \mathbb{R}^2$, atunci $\mathsf{Ker}(A - \lambda I_2) = \mathbb{R}^2$. Dar $\mathsf{Ker}(A - \lambda I_2) = V_\lambda$ este un subspaţiu de dimensiune 1. Deci presupunerea făcută este falsă.
   - Notăm $\mathbf{v}_1 = (A - \lambda I_2)\mathbf{v}_2$. Atunci $\mathbf{v}_1 \neq \mathbf{0}$ şi $(A - \lambda I_2)\mathbf{v}_1 = (A - \lambda I_2)^2\mathbf{v}_2 = \mathbf{0}$. Am ţinut cont aici de teorema Hamilton-Cayley, conform căreia $(A - \lambda I_2)^2 = 0_2$. Rezultă $A\mathbf{v}_1 = \lambda\mathbf{v}_1$. De asemenea, avem şi $A\mathbf{v}_2 = (A - \lambda I_2)\mathbf{v}_2 + \lambda I_2\mathbf{v}_2 = (A - \lambda I_2)\mathbf{v}_2 + \lambda\mathbf{v}_2$.
   - Vectorii $\mathbf{v}_1$ şi $\mathbf{v}_2$ sunt liniar independenţi, deoarece $\mathbf{v}_1 \in V_\lambda$ şi $\mathbf{v}_2 \notin V_\lambda$. În particular, matricea $P = (\mathbf{v}_1|\mathbf{v}_2)$ este inversabilă.
   - Relaţiile $A\mathbf{v}_1 = \lambda\mathbf{v}_1$şi $A\mathbf{v}_2 = (A - \lambda I_2)\mathbf{v}_2 + \lambda\mathbf{v}_2$ se mai pot scrie

     $$A\underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P} = \underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P}\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

     Dacă notăm şi $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$[18] atunci $A = PJP^{-1}$.

[18] Acest tip de matrice se mai numeşte şi matrice Jordan.

În particular, $A^n = PJ^nP^{-1}$ şi prin inducţie se verifică că $J^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$.

3) $\lambda_1, \lambda_2 \in \mathbb{C} - \mathbb{R}$, $\lambda_2 = \overline{\lambda_1}$. Să notăm mai simplu valorile proprii $\lambda = \alpha_{\mathsf{i}}\beta$ şi $\overline{\lambda} = \alpha - \mathsf{i}\beta$, $\beta \neq 0$. Întrucât valorile proprii sunt simple (au multiplicitate algebrică 1) şi complexe, rezultă că matricea A este diagonalizabilă peste corpul numerelor complexe, conform Corolarului 6.16. Deci reluând raţionamentul de la punctul 1), putem diagonaliza matricea dar de data aceasta cu $P, D \in \mathcal{M}_2(\mathbb{C})$.

Deoarece vom utiliza rezultatele obţinute mai sus în următorul capitol (în rezolvarea sistemelor de ecuaţii diferenţiale liniare), ar fi de preferat să obţinem pentru matricea A o descompunere alternativă cu matrice reale şi nu cu matrice complexe.

Pentru început, să observăm că dacă $\mathbf{v} \in V_\lambda \subseteq \mathbb{C}^2$ este un vector propriu cu componente complexe pentru valoarea proprie λ, atunci din relaţia $A\mathbf{v} = \lambda\mathbf{v}$, prin conjugare complexă, rezultă $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$, unde $\overline{\mathbf{v}}$ este vectorul care are drept componente conjugatele complexe ale componentelor vectorului $\mathbf{v}$. Deci pentru o matrice reală, la valori proprii complexe conjugate $\overline{\lambda}$ corespund, prin extensie, vectori proprii conjugaţi $\overline{\mathbf{v}} \in V_{\overline{\lambda}}$.

Să notăm acum $\mathbf{v} = \mathbf{v}_1 + \mathsf{i}\mathbf{v}_2$, cu $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$. Atunci relaţia $A\mathbf{v} = \lambda\mathbf{v}$ se mai scrie

$$A(\mathbf{v}_1 + \mathsf{i}\mathbf{v}_2) = (\alpha + \mathsf{i}\beta)(\mathbf{v}_1 + \mathsf{i}\mathbf{v}_2)$$

de unde, identificând părţile reale şi imaginare, rezultă

$$A\mathbf{v}_1 = \alpha\mathbf{v}_1 - \beta\mathbf{v}_2 \qquad \text{şi} \qquad A\mathbf{v}_2 = \beta\mathbf{v}_1 + \alpha\mathbf{v}_2$$

Putem scrie relaţiile de mai sus în forma matriceală

$$A\underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P} = \underbrace{(\mathbf{v}_1|\mathbf{v}_2)}_{P}\underbrace{\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}}_{Q}$$

Exerciţiu 6.23. Arătaţi că vectorii $\mathbf{v}_1, \mathbf{v}_2$ sunt liniar independenţi.

Rezultă că matricea $P = (\mathbf{v}_1|\mathbf{v}_2)$ este inversabilă, şi dacă notăm $Q = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, atunci $A = PQP^{-1}$. Să observăm în particular că dacă reprezentăm $\lambda = \alpha + \mathsf{i}\beta$ sub formă trigonometrică $\lambda = r(\cos\theta + \mathsf{i}\sin\theta)$, atunci $Q = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = r\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ este, modulo factorul r, o matrice de rotaţie (în sens invers trigonometric, cu un unghi θ în jurul originii).[19] Atunci $A^n = PQ^nP^{-1} = Pr^n\begin{pmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{pmatrix}P^{-1}$.

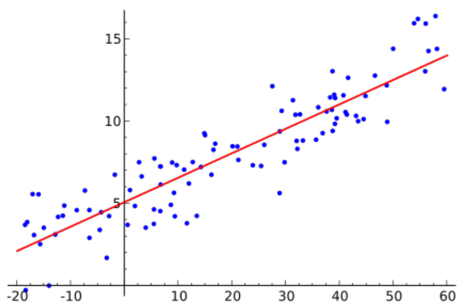
[19]Vom vedea în Secţiunea următoare că matricea Q este o matrice ortogonală (o matrice cu proprietatea că $Q^tQ = I_2$), ceea ce va justifică şi notaţia.

7 Spaţii euclidiene

Conceptele din algebra liniară întâlnite în cadrul acestui curs au fost ilustrate, majoritatea, cu exemple din geometrie. Totuşi, poate aţi remarcat că una dintre principalele noţiuni geometrice, şi anume perpendicularitatea, a lipsit din exemplele discutate până acum. Motivaţia este simplă: conceptele precedente de algebră liniară fiind insuficiente pentru a discuta despre perpendicularitate, vom avea nevoie de instrumente noi: produs scalar, ortogonalitate, proiecţie.

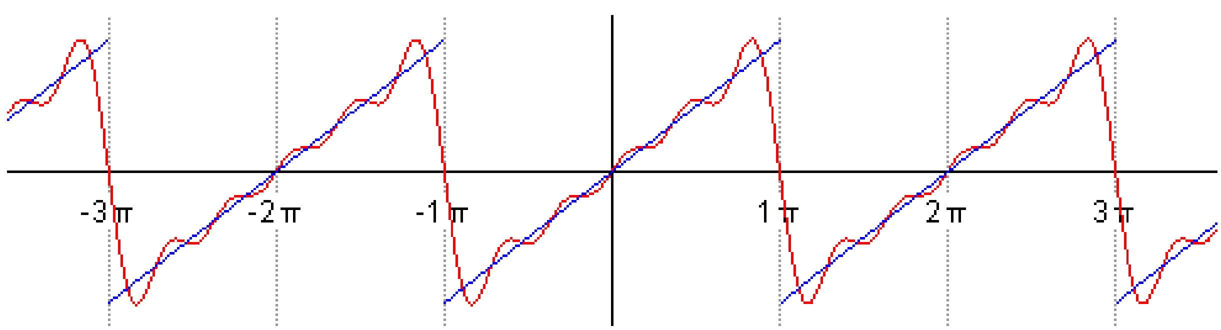
Pe de altă parte, introducerea unor noţiuni doar pentru ilustrarea conceptelor geometrice nu este o justificare suficientă, în particular pentru Ştiinţa Calculatoarelor. Vom da mai jos două aplicaţii: dreapta de regresie (poate cel mai simplu model de învăţare supravegheată) şi descompunerea Fourier (teoria şi prelucrarea semnalelor).

Exemplu 7.1. 1) Dreapta de regresie: **aproximarea** liniară a unei mulţimi de puncte.[20]



2) Aproximarea funcţiilor periodice cu polinoame trigonometrice (analiza Fourier a semnalelor). Să considerăm de exemplu semnalul periodic discontinuu $f(t) = \frac{t}{2}$, $t \in (-\pi, \pi)$, care se poate aproxima printr-un şir de semnale trigonometrice periodice şi *continue*:[21]

$$f(t) \approx \sin t - \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t - \frac{1}{4}\sin 4t + \frac{1}{5}\sin 5t - \dots$$

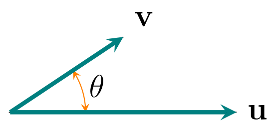


Matematic, această aproximare se realizează prin determinarea coordonatelor semnalului f în baza **ortonormată** formată cu funcţiile trigonometrice, şi *reprezentarea* semnalului în această bază (mai general, prin *proiecţia* semnalului pe spaţiul generat de funcţiile trigonometrice).

[20] Metode Numerice (an I, sem. 2), Inteligenţă Artificială (an IV), Învăţare Automată (an IV).
[21] Matematici speciale (an I, sem 2), Procesarea Semnalelor (an IV).

7.1 Produs scalar

În geometrie, aţi întâlnit deja noţiuni ca modulul unui vector $|\mathbf{u}|$, unghiul θ dintre (dreptele suport a) doi vectori $\mathbf{u}$ şi $\mathbf{v}$ ($0 \leq \theta \leq \pi$), sau produsul scalar a doi vectori $\langle \mathbf{u}, \mathbf{v} \rangle = |\mathbf{u}||\mathbf{v}| \cos \theta$.



Cum putem generaliza acestea pentru un spaţiu vectorial oarecare? Întrucât vom lucra cu noţiuni cum ar fi lungimea unui vector, avem nevoie de conceptul de pozitivitate, valabil doar pentru numerele reale. Ca atare, vom lua $\Bbbk = \mathbb{R}$ în acest capitol (toate spaţiile vectoriale vor fi reale).[22]

Spaţiu euclidian

Definiţie 7.2. Fie V un $\mathbb{R}$-spaţiu vectorial. Un **produs scalar** pe V este o funcţie

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R} \qquad \mathbf{u}, \mathbf{v} \in V \mapsto \langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{R}$$

astfel încât

1) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2) $\langle \mathbf{u}_1 + \mathbf{u}_2, \mathbf{v} \rangle = \langle \mathbf{u}_1, \mathbf{v} \rangle + \langle \mathbf{u}_2, \mathbf{v} \rangle$
3) $\langle \lambda \mathbf{u}, \mathbf{v} \rangle = \lambda \langle \mathbf{u}, \mathbf{v} \rangle$
4) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$
5) $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Longleftrightarrow \mathbf{u} = \mathbf{0}$

Un spaţiu vectorial real finit dimensional V înzestrat cu un produs scalar $\langle\ ,\ \rangle$ se numeşte **spaţiu euclidian** $(V, \langle\ ,\ \rangle)$.

Exemplu 7.3. 1) Pentru $V = \mathbb{R}^n$, următoarea corespondenţă determină un produs scalar, numit **produsul scalar canonic** pe $\mathbb{R}^n$: pentru $\mathbf{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$, notăm $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + \ldots + x_n y_n$. Să observăm că acesta se mai poate scrie matricial astfel:

$$\langle \mathbf{u}, \mathbf{v} \rangle \quad = \quad x_1 y_1 + \ldots + x_n y_n \quad = \quad \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad = \quad \mathbf{u}^t \mathbf{v}$$

2) În cazul $V = \mathbb{R}_n[X]$, utilizând izomorfismul $\mathbb{R}_n[X] \cong \mathbb{R}^{n+1}$ şi produsul scalar anterior, obţinem următoarea formulă pentru produsul scalar a două polinoame având gradul cel mult n:

$$\langle a_0 + a_1 X + \ldots + a_n X^n, b_0 + b_1 X + \ldots + b_n X^n \rangle = a_0 b_0 + \ldots + a_n b_n$$

[22]Se poate lucra şi cu spaţii vectoriale complexe, înlocuind proprietatea 1) din Definiţia 7.2 cu $\langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle}$.

3) De asemenea, utilizând izomorfismul $\mathcal{M}_{m,n}(\mathbb{R}) \cong \mathbb{R}^{mn}$, rezultă următorul produs scalar pentru matrice

$$A, B \in \mathcal{M}_{m,n}(\mathbb{R}), A = (a_{ij})_{i,j=\overline{1,n}}, B = (b_{ij})_{i,j=\overline{1,n}} \implies \langle A, B\rangle = \sum_{i,j=1}^{n} a_{ij}b_{ij}$$

care se mai poate scrie în formă condensată

$$\langle A, B\rangle = \mathsf{tr}(A^t B)$$

4) Pe spaţiul vectorial $V = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ continuă}\}$, formula

$$\langle f, g\rangle = \int_a^b f(t)g(t)\mathsf{d}t$$

defineşte un produs scalar, unde integrala este calculată pe un interval $[a, b]$ (cel mai des utilizate intervale fiind $[0, 1]$, $[-1, 1]$, $[0, 2\pi]$, $[-\pi, \pi]$).

5) Fie acum ℓ^2 spaţiul vectorial al şirurilor reale de pătrat sumabil (Exemplul 3.5.3c). Arătaţi că formula

$$\langle (a_n)_{n\geq 0}, (b_n)_{n\geq 0}\rangle = \sum_{n\geq 0} |a_n b_n|$$

este bine definită (seria converge) şi că determină un produs scalar pe ℓ^2.

Exerciţiu 7.4. Verificaţi pentru fiecare din exemplele de mai sus definiţia produsului scalar.

Norma unui vector într-un spaţiu euclidian

Definiţie 7.5. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian. **Norma** unui vector $\mathbf{v} \in V$ este numărul real pozitiv

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

Să observăm că definiţia de mai sus are sens, întrucât $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, $\forall \mathbf{v} \in V$.

Exemplu 7.6. Plecând de la produsul scalar canonic pe $\mathbb{R}^n$ se obţine **norma euclidiană** a unui vector $\mathbf{v} \in \mathbb{R}^n$ prin formula

$$\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \implies \|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} = \sqrt{\mathbf{v}^t\mathbf{v}} = \sqrt{x_1^2 + \ldots + x_n^2}$$

Proprietăţile produsului scalar şi ale normei

Propoziţie 7.7. Într-un spaţiu euclidian arbitrar $(V, \langle\ ,\ \rangle)$, produsul scalar şi norma verifică următoarele relaţii:

1) $\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$

2) $\|\mathbf{v}\| \geq 0$ şi $(\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0})$

3) $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ (*inegalitatea Cauchy-Schwarz*)

4) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ (*inegalitatea triunghiului*)

5) $\langle \mathbf{u}, \mathbf{v} \rangle = 0 \iff \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ (*teorema lui Pitagora*)

6) $2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2 = \|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{u} + \mathbf{v}\|^2$ (*identitatea paralelogramului*)

Demonstraţie. Vom demonstra proprietăţile 3),4) şi 5), restul lăsându-le ca exerciţiu.

3) Fie $\mathbf{u}, \mathbf{v} \in V$ şi $\lambda \in \mathbb{R}$ (reamintim că un spaţiu euclidian este în particular un $\mathbb{R}$-spaţiu vectorial). Atunci

$$0 \leq \langle \mathbf{u} + \lambda \mathbf{v}, \mathbf{u} + \lambda \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\lambda \langle \mathbf{u}, \mathbf{v} \rangle + \lambda^2 \langle \mathbf{v}, \mathbf{v} \rangle$$

Cum inegalitatea de mai sus are loc pentru orice $\lambda \in \mathbb{R}$, rezultă că trinomul de gradul II în λ admite discriminantul negativ sau nul:

$$0 \geq \Delta = 4\langle \mathbf{u}, \mathbf{v} \rangle - 4\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle$$

de unde rezultă inegalitatea dorită.

4) Luând $\lambda = 1$ în produsul scalar $\langle \mathbf{u} + \lambda \mathbf{v}, \mathbf{u} + \lambda \mathbf{v} \rangle$, se obţine ca mai sus

$$\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \tag{13}$$

Utilizând şi inegalitatea Cauchy-Schwarz, rezultă

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$$

deci $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

5) Rezultă din relaţia (13). □

Observaţie 7.8. Din inegalitatea Cauchy-Schwarz rezultă $\left| \dfrac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \right| \leq 1$, pentru orice $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$. Prin convenţie, numim **unghiul** $\sphericalangle(\mathbf{u}, \mathbf{v})$ vectorilor $\mathbf{u}$ şi $\mathbf{v}$ acel unghi $t \in [0, \pi]$ pentru care

$$\cos t = \left| \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} \right|$$

De exemplu, pentru $V = \mathbb{R}^2$, noţiunea de mai sus coincide cu cea studiată în liceu.

Exerciţiu 7.9. Nu orice normă provine dintr-un produs scalar. Mai precis, un spaţiu vectorial real V se numeşte **normat** dacă există o funcţie numită normă $\| \ \| : V \to \mathbb{R}_+$ care satisface proprietăţile 1, 2 şi 4 din Propoziţia 7.7. Arătaţi că pentru un spaţiu vectorial normat $(V, \| \ \|)$ există un produs scalar $\langle \ , \ \rangle$ care induce induce norma $\| \ \|$ dacă şi numai dacă are loc identitatea paralelogramului 6, şi că în acest caz, formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4} \left(\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2 \right) \tag{14}$$

defineşte produsul scalar căutat pe V (deci putem exprima atât norma în funcţie de produsul scalar cât şi invers, produsul scalar în funcţie de normă, şi că cele două corespondenţe sunt inverse una celeilalte).

Un exemplu de norma neindusă de un produs scalar pe $\mathbb{R}^n$ este norma $\|\mathbf{u}\|_\infty = \max_{i=\overline{1,n}} |u_i|$ (verificaţi că aceasta este o normă, adică proprietăţile 1, 2 şi 4 din Propoziţia 7.7, şi că $\| \ \|_\infty$ nu verifică identitatea paralelogramului, deci nu provine dintr-un produs scalar).

Exemplu 7.10 (Norme de matrice). 1) Pe $\mathcal{M}_{m,n}(\mathbb{R})$, norma indusă de produsul scalar din Exemplul 7.3.3 este $\|A\|_F = \sqrt{\mathsf{tr}(A^tA)}$. Aceasta se mai numeşte şi **norma Frobenius** a unei matrice.

2) (De la $\|\mathbf{v}\|$ la $\|A\mathbf{v}\|$) Multiplicând un vector $\mathbf{v} \in \mathbb{R}^n$ cu o matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$, norma $\|A\mathbf{v}\|$ poate creşte sau poate scădea. Formula

$$\|A\|_2 = \sup_{\mathbf{v}\neq\mathbf{0}} \frac{\|A\mathbf{v}\|}{\|\mathbf{v}\|}$$

indică cel mai mare factor de creştere posibil la multiplicarea cu matricea A, unde pentru $\mathbf{v} \in \mathbb{R}^n$ şi $A\mathbf{v} \in \mathbb{R}^m$ s-a considerat norma euclidiană. În particular, $\|A\mathbf{v}\| \leq \|A\|_2\|\mathbf{v}\|$.

Arătaţi că $\| \ \|_2$ este într-adevăr o normă, numită **norma operator**, adică verifică proprietăţile 1, 2 şi 4 din Propoziţia 7.7.

Verificaţi că norma operator se poate determina şi cu formula (utilă în calcule)

$$\|A\|_2 = \sup_{\|\mathbf{v}\|=1} \|A\mathbf{v}\|$$

Norma operator a matricelor este utilă mai ales în detectarea variaţiilor care pot interveni la rezolvarea sistemelor liniare de către maşinile de calcul (erori de aproximare, de virgulă mobilă, etc.). De exemplu, să considerăm sistemele

$$\begin{cases} x+y &= 2 \\ x+1.001y &= 2 \end{cases} \qquad \begin{cases} x+y &= 2 \\ x+1.001y &= 2.001 \end{cases}$$

cu soluţiile $x = 2, y = 0$, respectiv $x = y = 1$. O variaţie minoră a termenilor liberi a condus la o diferenţă semnificativă a soluţiilor. În asemenea cazuri, matricea sistemului se numeşte **prost condiţionată**. Un alt exemplu tipic în acest sens este rezolvarea sistemelor prin metoda lui Gauss:

$$\begin{cases} 0.001x+y &= 1 \\ x+y &= 2 \end{cases}$$

Sistemul de mai sus are soluţia $x = \dfrac{1000}{999} \approx 1$, $y = \dfrac{998}{999} \approx 1$. Scăzând din doua ecuaţie prima ecuaţie înmulţită cu 1000 conduce la

$$\begin{cases} 0.001x+y &= 1 \\ -999y &= -998 \end{cases}$$

de unde prin rotunjirea ca 3 zecimale exacte $\frac{998}{999} \approx 1$ rezultă

$$\begin{cases} 0.001x+y &= 1 \\ y &= 1 \end{cases}$$

cu soluţia $x = 0, y = 1$ incorectă. Deci mici variaţii în cadrul procesului de calcul pot conduce la erori majore în determinarea soluţiei.

Numărul de condiţionare al unei matrice inversabile se defineşte ca fiind $\|A\|_2 \|A^{-1}\|_2$. Cu cât numărul de condiţionare al unei matrice este mai mare, cu atât erorile obţinute în cadrul procesului de rezolvare al sistemelor liniare asociate vor fi mai mari.

Determinaţi numărul de condiţionare pentru matricele $\begin{pmatrix} 1 & 1 \\ 1 & 1.001 \end{pmatrix}$ şi $\begin{pmatrix} 0.001 & 1 \\ 1 & 1 \end{pmatrix}$ care apar în sistemele de mai sus.

Exerciţiu 7.11 (Înmulţirea matricelor şi norma). 1) Spre deosebire de vectori în general, matricele se pot înmulţi: pentru $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{R})$, $B = (b_{jk}) \in \mathcal{M}_{n,p}(\mathbb{R})$, produsul acestora este o matrice $AB \in \mathcal{M}_{m,p}(\mathbb{R})$ cu $(AB)_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Arătaţi că atât norma Frobenius cât şi norma $\| \ \|_2$ sunt *submultiplicative*, adică satisfac relaţia

$$\|AB\| \leq \|A\| \|B\|, \forall A \in \mathcal{M}_{m,n}(\mathbb{R}), B \in \mathcal{M}_{n,p}(\mathbb{R})$$

(indicaţie: utilizaţi inegalitatea Cauchy-Schwarz).

2) (Norma inversei unei matrice de normă mică este mare) Dacă $A \in \mathcal{M}_n(\mathbb{R})$ este o matrice inversabilă cu $\|A\|_F \leq \epsilon$, atunci $\|A^{-1}\|_F \geq \frac{\sqrt{n}}{\epsilon}$.

Distanţa dintre doi vectori într-un spaţiu euclidian

Definiţie 7.12. Fie $(V, \langle \ , \ \rangle)$ un spaţiu euclidian. **Distanţa** dintre doi vectori $\mathbf{u}, \mathbf{v} \in V$ este $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.

Exemplu 7.13. În exemplul standard al vectorilor coloană, plecând de la produsul scalar canonic pe $\mathbb{R}^n$, obţinem **distanţa euclidiană**

$$\mathbf{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \mathbf{v} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \implies d(\mathbf{u}, \mathbf{v}) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$$

Proprietăţile distanţei

Propoziţie 7.14. Fie $(V, \langle \ , \ \rangle)$ un spaţiu euclidian. Atunci distanţa verifică următoarele proprietăţi:

1) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

2) $d(\mathbf{u}, \mathbf{v}) = 0 \iff \mathbf{u} = \mathbf{v}$

3) $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ (*inegalitatea triunghiului*)

Demonstraţie. Rezultă imediat din proprietăţile normei (Propoziţia 7.7). □

Observaţie 7.15. Nu orice distanţă este indusă de o normă/de un produs scalar. De exemplu, pentru distanţa **discretă** (care măsoară dacă doi vectori coincid sau nu)

$$d(\mathbf{u},\mathbf{v}) = \begin{cases} 1 & ,\mathbf{u} \neq \mathbf{v} \\ 0 & ,\mathbf{u} = \mathbf{v} \end{cases}$$

nu există nici un produs scalar $\langle\ ,\ \rangle$ pe $\mathbb{R}^n$ pentru care $d(\mathbf{u},\mathbf{v}) = \sqrt{\langle \mathbf{u}-\mathbf{v}, \mathbf{u}-\mathbf{v}\rangle}$ şi nici o normă $\|\ \|$ astfel încât $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$.

Ce puteţi spune despre distanţa **taxicab** (numită şi sau distanţa **Manhattan**) pe $\mathbb{R}^n$, dată de formula $d(\mathbf{u},\mathbf{v}) = \sum_{i=1}^{n} |u_i - v_i|$?[23]

7.2 Ortogonalitate şi proiecţie

Vectori ortogonali

Definiţie 7.16. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $\mathbf{u},\mathbf{v} \in V$. Dacă

$$\langle \mathbf{u},\mathbf{v}\rangle = 0$$

spunem că vectorii $\mathbf{u},\mathbf{v}$ sunt **ortogonali** în raport cu produsul scalar $\langle\ ,\ \rangle$ şi notăm $\mathbf{u} \perp \mathbf{v}$. Dacă în plus $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$, spunem că vectorii sunt **ortonormaţi**.

Să observăm că vectorul nul este ortogonal pe orice alt vector; că dacă $\mathbf{u} \perp \mathbf{v}$ atunci şi $\mathbf{v} \perp \mathbf{u}$ (relaţia de ortogonalitate este simetrică); şi că dacă $\mathbf{u} \perp \mathbf{v}, \mathbf{u} \perp \mathbf{w}$, atunci $\mathbf{u}$ este ortogonal (sau perpendicular)[10] pe orice combinaţie liniară $\alpha\mathbf{v} + \beta\mathbf{w}$, $\forall\ \alpha,\beta \in \mathbb{R}$.

Exemplu 7.17. 1) În $\mathbb{R}^n$, vectorii $\mathbf{e}_1,\ldots,\mathbf{e}_n$ din baza canonică sunt ortogonali doi câte doi în raport cu produsul scalar canonic, şi au norma egală cu unitatea, deci sunt şi ortonormaţi.

2) În spaţiul vectorial

$$V = \{ f : \mathbb{R} \to \mathbb{R} \mid f \text{continuă şi periodică cu perioada } 2\pi \}$$

oricare două dintre funcţiile $1, \sin t, \cos t, \sin 2t, \cos 2t, \ldots, \sin nt, \cos nt, \ldots,$ sunt ortogonale în raport cu produsul scalar

$$\langle f, g\rangle = \int_0^{2\pi} f(t)g(t)\mathsf{d}t$$

Exerciţiu 7.18. (Matrice ortogonale) Pentru un alt exemplu, verificaţi că matricele Pauli (Exemplul 2.13.3), împreună cu matricea identică I_2, formează o bază ortogonală în $\mathcal{M}_2(\mathbb{C})$. Atenţie: în acest caz se lucrează peste corpul numerelor complexe, cu produsul scalar $\langle A, B\rangle = \mathsf{tr}(A^t\overline{B})$, unde $A, B \in \mathcal{M}_2(\mathbb{C})$, şi $\overline{B} = (\overline{b_{ij}})_{i,j}$ este matricea obţinută prin conjugarea complexă a componentelor matricei B.

[23] Mai multe detalii despre distanţa taxicab în setul 6-7 de exerciţii.

[10] Noţiunea de perpedicularitate provine din limba latină şi este uzual utilizată în geometrie, în timp ce noţiunea de ortogonalitate, originară din limba greacă, apare mai ales în algebra liniară, chiar şi în situaţii în care aparent nu avem nici o intuiţie geometrică.

Exerciţiu 7.19. (Polinoamele ortogonale - polinoamele Cebîşev) Formulele cunoscute $\cos t = \cos t, \cos 2t = 2\cos^2 t - 1, \cos 3t = 4\cos^3 t - 3\cos t$, etc. sugerează o dependenţă polinomială de forma $\cos nt = T_n(\cos t)$.

1) Să se verifice că $T_1(X) = X$, $T_2(X) = 2X^2 - 1$, $T_3(X) = 4X^3 - 3$ şi că $T_{n+1}(X) = 2XT_n(X) - T_{n-1}(X)$ (indicaţie: utilizaţi relaţia $\cos(n+1)t + \cos(n-1)t = 2\cos t \cos nt$). În particular, rezultă că T_n este într-adevăr un polinom. Mai mult, deducem că $\mathsf{grad} T_n = n$ şi că termenul de grad maxim are coeficientul 2^{n-1}. Prin convenţie, punem $T_0(X) = 1$.

   Polinoamele $(T_n)_{n\geq 0}$ sunt cunoscute drept **polinoamele Cebîşev**. Datorită proprietăţilor acestora, au o largă aplicabilitate în prelucrarea semnalelor (filtre) şi în algoritmi numerici de învăţare şi interpolare.[24]

2) Să se arate că formula $\langle P, Q\rangle = \displaystyle\int_{-1}^{1} P(x)Q(x)\frac{1}{\sqrt{1-x^2}}\mathsf{d}x$ determină un produs scalar pe spaţiul vectorial al polinoamelor $\mathbb{R}[X]$.

3) Să se verifice că polinoamele Cebîşev $(T_n)_{n\geq 0}$ sunt ortogonale în raport cu produsul scalar de mai sus (indicaţie: efectuaţi schimbarea de variabilă $x = \cos t$). Mai precis, arătaţi că

$$\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\mathsf{d}x = \begin{cases} 0 & m \neq n \\ \frac{\pi}{2} & m = n \neq 0 \\ \pi & m = n = 0 \end{cases}$$

4) Utilizând serii de puteri, verificaţi că

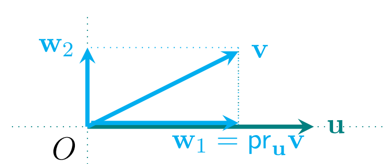
$$\sum_{n\geq 0} T_n(x)t^n = \frac{1 - tx}{1 - 2tx + t^2}$$

   pentru $|x| < 1$. Pe ce domeniu converge seria de puteri de mai sus? Funcţia $f(t) = \dfrac{1-tx}{1-2tx+t^2}$ se numeşte **funcţia generatoare** a polinoamelor Cebîşev.

Proiecţia pe un vector

Propoziţie 7.20. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} \neq \mathbf{0}$. Atunci există şi sunt unici vectorii $\mathbf{w}_1, \mathbf{w}_2 \in V$ astfel încât

$$\mathbf{v} = \underset{\text{vector coliniar cu } \mathbf{u}}{\mathbf{w}_1} + \underset{\text{vector ortogonal pe } \mathbf{u}}{\mathbf{w}_2}$$



Vom numi vectorul $\mathbf{w}_1$ **proiecţia** lui $\mathbf{v}$ pe $\mathbf{u}$ şi îl vom nota $\mathsf{pr}_{\mathbf{u}}\mathbf{v}$.

[24] Metode numerice (sem II, anul I).

Demonstraţie. Unicitatea: dacă $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1' + \mathbf{w}_2'$ cu $\mathbf{w}_1, \mathbf{w}_1'$ coliniari cu $\mathbf{u}$, respectiv $\mathbf{w}_2, \mathbf{w}_2'$ ortogonali pe $\mathbf{u}$, atunci $\mathbf{w}_1 - \mathbf{w}_1' = \mathbf{w}_2' - \mathbf{w}_2$. Din $\mathbf{w}_1, \mathbf{w}_1'$ coliniari cu $\mathbf{u}$ rezultă că există $\alpha \in \mathbb{R}$ astfel încât $\mathbf{w}_1 - \mathbf{w}_1' = \alpha\mathbf{u}$. Atunci

$$\|\mathbf{w}_1 - \mathbf{w}_1'\|^2 = \langle \mathbf{w}_1 - \mathbf{w}_1', \mathbf{w}_1 - \mathbf{w}_1' \rangle = \langle \lambda\mathbf{u}, \mathbf{w}_2' - \mathbf{w}_2 \rangle = \lambda(\underbrace{\langle \mathbf{u}, \mathbf{w}_2' \rangle}_{0} - \underbrace{\langle \mathbf{u}, \mathbf{w}_2 \rangle}_{0}) = 0$$

de unde rezultă $\mathbf{w}_1 = \mathbf{w}_1'$. Cum $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1' + \mathbf{w}_2'$, obţinem şi $\mathbf{w}_2 = \mathbf{w}_2'$.

Existenţa: Căutăm un vector $\mathbf{w}_1$ coliniar cu $\mathbf{u}$ (deci de forma $\mathbf{w} = \alpha\mathbf{u}$, unde $\alpha \in \mathbb{R}$) astfel încât $\mathbf{v} - \mathbf{w}_1 \perp \mathbf{u}$. Din condiţia de ortogonalitate rezultă

$$0 = \langle \mathbf{v} - \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v} - \alpha\mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \alpha \langle \mathbf{u}, \mathbf{u} \rangle$$

deci $\alpha = \dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$ (observaţi că $\langle \mathbf{u}, \mathbf{u} \rangle \neq 0$ deoarece $\mathbf{u} \neq \mathbf{0}$). am obţinut deci descompunerea dorită

$$\mathbf{v} = \underbrace{\left(\frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} \right)}_{\text{vector coliniar cu } \mathbf{u}} + \underbrace{\left(\mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} \right)}_{\text{vector ortogonal pe } \mathbf{u}}$$

□

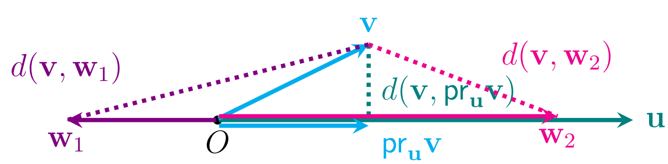
Din demonstraţia precedentă, să reţinem următoarele consecinţe:

Corolar 7.21. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} \neq 0$. Atunci:

1) (**Formula proiecţiei**) Proiecţia vectorului $\mathbf{v}$ pe vectorul $\mathbf{u}$ este $\mathsf{pr}_{\mathbf{u}}\mathbf{v} = \dfrac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$.

2) Proiecţia vectorului $\mathbf{v}$ pe vectorul $\mathbf{u}$ este acel vector care **minimizează** distanţa dintre $\mathbf{v}$ şi vectorii coliniari cu $\mathbf{u}$:

$$d(\mathbf{v}, \mathsf{pr}_{\mathbf{u}}\mathbf{v}) = \min_{\mathbf{w} \in \mathsf{Sp}\{\mathbf{u}\}} d(\mathbf{v}, \mathbf{w})$$



Demonstraţie. 1) Clar.

2) Fie $\mathbf{w} \in \mathsf{Sp}\{\mathbf{u}\}$ un vector arbitrar. Atunci

$$\begin{aligned} d(\mathbf{v}, \mathbf{w})^2 &= \|\mathbf{v} - \mathbf{w}\|^2 = \| \underbrace{(\mathbf{v} - \mathsf{pr}_{\mathbf{u}}\mathbf{v})}_{\text{vector ortogonal pe } \mathbf{u}} + \underbrace{(\mathsf{pr}_{\mathbf{u}}\mathbf{v} - \mathbf{w})}_{\text{vector coliniar cu } \mathbf{u}} \|^2 \\ &\overset{\text{Thm. Pitagora}}{=} \|\mathbf{v} - \mathsf{pr}_{\mathbf{u}}\mathbf{v}\|^2 + \|\mathsf{pr}_{\mathbf{u}}\mathbf{v} - \mathbf{w}\|^2 \geq \|\mathbf{v} - \mathsf{pr}_{\mathbf{u}}\mathbf{v}\|^2 \\ &= d(\mathbf{v}, \mathsf{pr}_{\mathbf{u}}\mathbf{v})^2 \end{aligned}$$

□

Ortogonalitate implică liniar independenţă

Propoziţie 7.22. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian. Dacă vectorii $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ sunt nenuli şi ortogonali doi câte doi, atunci sunt şi liniar independenţi.

Demonstraţie. Fie $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ astfel încât $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0}$. Înmulţind scalar relaţia cu $\mathbf{v}_1$, rezultă

$$\alpha_1\underbrace{\langle\mathbf{v}_1, \mathbf{v}_1\rangle}_{\neq 0} + \alpha_2\underbrace{\langle\mathbf{v}_2, \mathbf{v}_1\rangle}_{0} + \ldots + \alpha_n\underbrace{\langle\mathbf{v}_n, \mathbf{v}_1\rangle}_{0} = \langle\mathbf{0}, \mathbf{v}_1\rangle = 0$$

de unde deducem $\alpha_1 = 0$. Analog, $\alpha_2 = \ldots = \alpha_n = 0$. □

Baze ortogonale şi baze ortonormate

Definiţie 7.23. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian cu $\dim_{\mathbb{R}} V = n$.

O **bază ortogonală** în V este o mulţime $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de vectori nenuli şi ortogonali doi câte doi.

Dacă în plus $\|\mathbf{v}_1\| = \ldots = \|\mathbf{v}_n\| = 1$, spunem că baza $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ este o **bază ortonormată**.

Să observăm că datorită propoziţiei precedente, vectorii dintr-o bază ortogonală/ortonormată sunt liniar independenţi. Deoarece numărul acestora coincide cu $\dim_{\mathbb{R}} V$, formează într-adevăr o bază.

Exemplu 7.24. Baza canonică $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ din $\mathbb{R}^n$ este o bază ortonormată în raport cu produsul scalar canonic. Mai mult, reprezentarea unui vector $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ în raport cu această bază este

$$\mathbf{v} = \langle\mathbf{v}, \mathbf{e}_1\rangle\mathbf{e}_1 + \ldots + \langle\mathbf{v}, \mathbf{e}_n\rangle\mathbf{e}_n$$

Vom vedea că relaţia de mai sus este aplicabilă oricărui spaţiu euclidian şi oricărei baze ortonormate. Cu alte cuvinte, determinarea coordonatelor unui vector în raport cu o bază ortonormată se poate realiza foarte uşor din punct de vedere computaţional – calculând doar n produse scalare (comparaţi numărul de operaţii necesare în cazul unei baze ortonormate faţă de cazul unei baze arbitrare).

Reprezentarea vectorilor într-o bază ortogonală (ortonormată)

Propoziţie 7.25. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian cu baza **ortogonală** $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Atunci orice vector $\mathbf{v} \in V$ se scrie

$$\begin{array}{ccccccc} \mathbf{v} & = & \mathsf{pr}_{\mathbf{v}_1}\mathbf{v} & + & \ldots & + & \mathsf{pr}_{\mathbf{v}_n}\mathbf{v} \\ & = & \dfrac{\langle\mathbf{v}, \mathbf{v}_1\rangle}{\langle\mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1 & + & \ldots & + & \dfrac{\langle\mathbf{v}, \mathbf{v}_n\rangle}{\langle\mathbf{v}_n, \mathbf{v}_n\rangle}\mathbf{v}_n \end{array}$$

Deci $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \frac{\langle\mathbf{v},\mathbf{v}_1\rangle}{\langle\mathbf{v}_1,\mathbf{v}_1\rangle} \\ \vdots \\ \frac{\langle\mathbf{v},\mathbf{v}_n\rangle}{\langle\mathbf{v}_n,\mathbf{v}_n\rangle} \end{pmatrix}$ (**coordonatele** unui vector $\mathbf{v} \in V$ în raport cu o **bază ortogonală**).

În particular, dacă $\mathcal{B}$ este o bază **ortonormată**, atunci

$$\mathbf{v} \quad = \quad \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \ldots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n \quad \text{şi} \quad [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \langle \mathbf{v}, \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{v}_n \rangle \end{pmatrix} \tag{15}$$

Demonstraţie. Fie $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ astfel încât $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots \alpha_n \mathbf{v}_n$. Înmulţind scalar această relaţie cu $\mathbf{v}_1$, obţinem

$$\langle \mathbf{v}, \mathbf{v}_1 \rangle = \alpha_1 \underbrace{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}_{\neq 0} + \alpha_2 \underbrace{\langle \mathbf{v}_2, \mathbf{v}_1 \rangle}_{0} + \ldots + \alpha_n \underbrace{\langle \mathbf{v}_n, \mathbf{v}_1 \rangle}_{0} = \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle$$

Deci $\alpha_1 = \dfrac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}$. Procedând analog, rezultă $\alpha_i = \dfrac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}$ pentru orice $i \in \{1, \ldots, n\}$. Dacă în plus baza este ortonormată, vom ţine cont de $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = \|\mathbf{v}_i\|^2 = 1$. ☐

Corolar 7.26. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ o bază ortonormată.

Fie vectorii $\mathbf{u}, \mathbf{v} \in V$ având coordonatele $[\mathbf{u}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ şi $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ în raport cu baza $\mathcal{B}$.

Atunci:

1) $\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + \ldots + x_n y_n$
2) $\|\mathbf{u}\| = \sqrt{x_1^2 + \ldots + x_n^2}$
3) $d(\mathbf{u}, \mathbf{v}) := \|\mathbf{u} - \mathbf{v}\| = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$

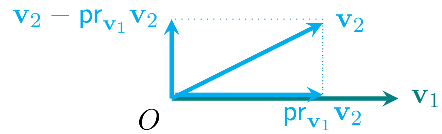
Deci cu ajutorul unei baze ortonormate, orice produs scalar pe un spaţiu euclidian abstract V se reduce la produsul scalar canonic pe $\mathbb{R}^n$, unde $n = \dim_{\mathbb{R}} V$.

Demonstraţie. Se ţine cont de relaţiile $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ şi de $\mathbf{u} = x_1 \mathbf{v}_1 + \ldots + x_n \mathbf{v}_n$, $\mathbf{v} = y_1 \mathbf{v}_1 + \ldots + y_n \mathbf{v}_n$. ☐

Am văzut mai sus utilitatea unei baze ortogonale/ortonormate. Vom demonstra mai jos existenţa unei asemenea baze. Mai mult chiar, vom determina un algoritm prin care se poate produce o bază ortogonală sau ortonormată pornind de la o bază arbitrară dintr-un spaţiu euclidian.

Pentru a înţelege mai bine, să considerăm cazul unui spaţiu vectorial V de dimensiune 2. Fie $\{\mathbf{v}_1, \mathbf{v}_2\}$ o bază în V, în particular $\mathbf{v}_1$ şi $\mathbf{v}_2$ sunt liniar independenţi (necoliniari). Dacă $\mathbf{v}_1 \perp \mathbf{v}_2$, atunci vectorii sunt ortogonali şi am terminat. În caz contrar, pornim cu primul vector din bază $\mathbf{v}_1$ şi căutăm un al doilea vector, ortogonal pe primul: Propoziţia 7.20 ne spune că vectorul $\mathbf{v}_2 - \mathsf{pr}_{\mathbf{v}_1} \mathbf{v}_2$

are proprietatea dorită:



Ideea descrisă mai sus se poate adapta pentru orice spaţiu euclidian:

Construcţia unei baze ortogonale (algoritmul Gram-Schmidt de ortogonalizare)

Teoremă 7.27. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian cu $\dim_{\mathbb{R}} V = n$ şi $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ o bază arbitrară în V. Considerăm următorii vectori şi următoarele subspaţii:

$$\begin{array}{ll}
\mathbf{u}_1 = \mathbf{v}_1 & U_1 = \mathsf{Sp}\,\{\mathbf{v}_1\} \\
\mathbf{u}_2 = \mathbf{v}_2 - \mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_2 & U_2 = \mathsf{Sp}\,\{\mathbf{v}_1, \mathbf{v}_2\} \\
\mathbf{u}_3 = \mathbf{v}_3 - (\mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_3 + \mathsf{pr}_{\mathbf{u}_2}\mathbf{v}_3) & U_3 = \mathsf{Sp}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \\
\vdots & \vdots \\
\mathbf{u}_n = \mathbf{v}_n - (\mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_n + \ldots + \mathsf{pr}_{\mathbf{u}_{n-1}}\mathbf{v}_n) & U_n = \mathsf{Sp}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}
\end{array}$$

Atunci $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ formează o bază **ortogonală** în U_k, $\forall\ k = \overline{1, n}$ (deci $\dim_{\mathbb{R}} U_k = k$).
În particular, $U_n = V$ şi $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ este o bază **ortogonală** în V.

Demonstraţie. Vom utiliza inducţia matematică după $k \in \{1, \ldots, n\}$. Pentru $k = 1$ nu avem ce demonstra. Să presupunem acum că $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ formează o bază **ortogonală** în U_k. Să verificăm ortogonalitatea vectorilor $\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}$. De fapt, trebuie doar să verificăm că $\langle \mathbf{u}_j, \mathbf{u}_{k+1}\rangle = 0$, pentru orice $j \in \{1, \ldots, k\}$, restul relaţiilor de ortogonalitate rezultând din ipoteza de inducţie. Pentru $j = 1$, avem

$$\begin{aligned}
\langle \mathbf{u}_1, \mathbf{u}_{k+1}\rangle &= \langle \mathbf{u}_1, \mathbf{v}_{k+1} - (\mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_{k+1} + \ldots + \mathsf{pr}_{\mathbf{u}_k}\mathbf{v}_{k+1})\rangle \\
&= \langle \mathbf{u}_1, \mathbf{v}_{k+1}\rangle - \big(\langle \mathbf{u}_1, \mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_{k+1}\rangle + \ldots + \langle \mathbf{u}_1, \mathsf{pr}_{\mathbf{u}_k}\mathbf{v}_{k+1}\rangle\big) \\
&= \langle \mathbf{u}_1, \mathbf{v}_{k+1}\rangle - \left(\frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\langle \mathbf{u}_1, \mathbf{u}_1\rangle + \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_2\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\underbrace{\langle \mathbf{u}_1, \mathbf{u}_2\rangle}_{0} + \ldots + \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_k\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\underbrace{\langle \mathbf{u}_1, \mathbf{u}_k\rangle}_{0}\right) \\
&= \langle \mathbf{u}_1, \mathbf{v}_{k+1}\rangle - \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_1\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\langle \mathbf{u}_1, \mathbf{u}_1\rangle \quad = \quad 0
\end{aligned}$$

Analog, rezultă $\langle \mathbf{u}_2, \mathbf{u}_{k+1}\rangle = \ldots = \langle \mathbf{u}_k, \mathbf{u}_{k+1}\rangle = 0$.

Să verificăm acum că $U_{k+1} = \mathsf{Sp}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\}$. Avem

$$\begin{aligned}
\mathsf{Sp}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\} &= \mathsf{Sp}\{\ \underbrace{\mathbf{u}_1, \ldots, \mathbf{u}_k}_{U_k = \mathsf{Sp}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}}\ , \mathbf{v}_{k+1} - \underbrace{(\mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_{k+1} + \ldots + \mathsf{pr}_{\mathbf{u}_k}\mathbf{v}_{k+1})}_{\in U_k}\ \} \\
&= \mathsf{Sp}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_{k+1}\} \\
&= \mathsf{Sp}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}\} \quad = \quad U_{k+1}
\end{aligned}$$

Mai sus, am ţinut cont de ipoteza de inducţie $U_k = \mathsf{Sp}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} = \mathsf{Sp}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

Rezultă deci că pentru orice $k \in \{1, \ldots, n\}$, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ formează o bază **ortogonală** în U_k. $\square$

Corolar 7.28. În condiţiile teoremei precedente, $\{\mathbf{q}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1, \ldots, \mathbf{q}_n = \frac{1}{\|\mathbf{u}_n\|}\mathbf{u}_n\}$ este o bază ortonormată în V.

Exemplu 7.29. Să determinăm o bază ortonormată pentru subspaţiul $U \subseteq \mathbb{R}^4$ generat de vectorii $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 3 \\ 0 \\ -3 \\ 3 \end{pmatrix}$ în raport cu produsul scalar canonic. Avem

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ -2 \end{pmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \mathsf{pr}_{\mathbf{u}_1}\mathbf{v}_3 - \mathsf{pr}_{\mathbf{u}_2}\mathbf{v}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Am obţinut astfel $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ o bază ortogonală pentru $U = \mathsf{Sp}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Împărţind fiecare vector la norma sa, producem şi o bază ortonormată pentru U:

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{u}_1\|}\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \mathbf{q}_2 = \frac{1}{\|\mathbf{u}_2\|}\mathbf{u}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ -2 \end{pmatrix}, \mathbf{q}_3 = \frac{1}{\|\mathbf{u}_3\|}\mathbf{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

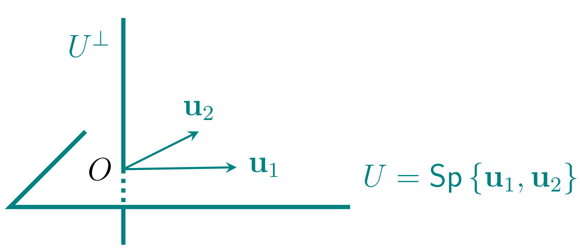
Complementul ortogonal al unui subspaţiu

Definiţie 7.30. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $U \subseteq V$ un subspaţiu vectorial. Mulţimea

$$U^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0,\ \forall\ \mathbf{u} \in U\}$$

se numeşte **complementul ortogonal** al subspaţiului vectorial U.

Observaţie 7.31. Dacă $U = \mathsf{Sp}\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, atunci $U^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{u}_i \rangle = 0, \forall\, i = \overline{1, m}\}$



Complementul ortogonal al unui subspaţiu

Propoziţie 7.32. Fie $(V, \langle\ ,\ \rangle)$ un spaţiu euclidian şi $U \subseteq V$ un subspaţiu vectorial. Atunci:

1) $U^\perp$ este subspaţiu vectorial în V.

2) $U \cap U^\perp = \{\mathbf{0}\}$

3) $(U^\perp)^\perp = U$

Demonstraţie. 1) Exerciţiu.

2) Fie $\mathbf{v} \in U \cap U^\perp$. Atunci $\langle \underbrace{\mathbf{v}}_{\in U}, \underbrace{\mathbf{v}}_{\in U^\perp} \rangle = 0$, de unde obţinem $\mathbf{v} = \mathbf{0}$.

3) Vom demonstra doar incluziunea $U \subseteq (U^\perp)^\perp$, cealaltă incluziune necesitând un rezultat adiţional. Vom reveni mai târziu asupra acesteia. Fie deci $\mathbf{u} \in U$. Atunci pentru orice $\mathbf{v} \in U^\perp$, $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, deci $\mathbf{u} \in (U^\perp)^\perp$.

□

Exemplu 7.33. Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$. Atunci

$$\mathsf{Ker}(A)^\perp = \mathsf{Im}(A^t) \quad \text{şi} \quad \mathsf{Im}(A)^\perp = \mathsf{Ker}(A^t)$$

Exerciţiu 7.34. Fie $V = \mathcal{M}_n(\mathbb{R})$ spaţiul euclidian al matricelor pătratice, cu produsul scalar introdus în Exemplul 7.3.3, şi $\mathsf{Sym}_n(\mathbb{R})$ şi $\mathsf{AntiSym}_n(\mathbb{R})$ subspaţiile vectoriale ale matricelor simetrice, respectiv antisimetrice (Exemplul 3.5.2). Am văzut în Exemplul 3.10 că $\mathcal{M}_n(\mathbb{R}) = \mathsf{Sym}_n(\mathbb{R}) \oplus \mathsf{AntiSym}_n(\mathbb{R})$. Arătaţi acum că orice matrice simetrică este ortogonală pe orice matrice antisimetrică, şi utilizaţi acest rezultat pentru a deduce că

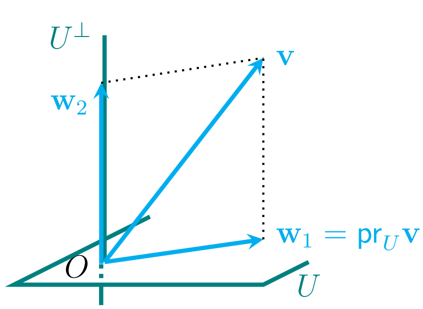
$$\mathsf{Sym}_n(\mathbb{R})^\perp = \mathsf{AntiSym}_n(\mathbb{R}) \qquad \text{şi} \qquad \mathsf{AntiSym}_n(\mathbb{R})^\perp = \mathsf{Sym}_n(\mathbb{R})$$

Proiecţia unui vector pe un subspaţiu

Propoziţie 7.35. Fie $(V, \langle , \rangle)$ un spaţiu euclidian, $U \subseteq V$ un subspaţiu nenul şi $\mathbf{v} \in V$ un vector arbitrar. Atunci **există şi sunt unici** vectorii $\mathbf{w}_1 \in U$, $\mathbf{w}_2 \in U^\perp$ astfel încât:

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$$

Vectorul $\mathbf{w}_1$ se numeşte **proiecţia** lui $\mathbf{v}$ pe subspaţiul U şi se notează $\mathsf{pr}_U \mathbf{v} \in U$.



Demonstraţie. Unicitatea: rezultă din $U \cap U^{\perp} = \{\mathbf{0}\}$.

Existenţa: căutăm un vector $\mathbf{w}_1 \in U$ astfel încât $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1 \in U^{\perp}$. Fie $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ o bază ortogonală pentru U. Reprezentăm vectorul $\mathbf{w}_1$ în baza **ortogonală** $\mathcal{B}$: $\mathbf{w}_1 = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_m \mathbf{u}_m$. Punem condiţia $\mathbf{w}_2 \in U^{\perp}$, echivalent $\langle \mathbf{v} - \mathbf{w}_1, \mathbf{u}_1 \rangle = \ldots = \langle \mathbf{v} - \mathbf{w}_1, \mathbf{u}_m \rangle = 0$. Din $\langle \mathbf{v} - \mathbf{w}_1, \mathbf{u}_1 \rangle = 0$ rezultă, ţinând cont de ortogonalitatea vectorilor $\mathbf{u}_1, ..., \mathbf{u}_m$, că $\alpha_1 = \dfrac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}$. Analog, $\alpha_2 = \dfrac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}, \ldots, \alpha_m = \dfrac{\langle \mathbf{v}, \mathbf{u}_m \rangle}{\langle \mathbf{u}_m, \mathbf{u}_m \rangle}$. Deci am găsit $\mathbf{w}_1 = \left(\frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \ldots + \frac{\langle \mathbf{v}, \mathbf{u}_m \rangle}{\langle \mathbf{u}_m, \mathbf{u}_m \rangle} \mathbf{u}_m \right) \in U$ şi $\mathbf{w}_2 = \mathbf{v} - \mathbf{w}_1 \in U^{\perp}$. □

Corolar 7.36. Fie $(V, \langle , \rangle)$ un spaţiu euclidian şi $U \subseteq V$ un subspaţiu. Dacă $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ este o bază **ortogonală** în U, atunci

$$\begin{aligned} \mathsf{pr}_U \mathbf{v} &= \mathsf{pr}_{\mathbf{u}_1} \mathbf{v} + \ldots + \mathsf{pr}_{\mathbf{u}_m} \mathbf{v} \\ &= \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \ldots + \frac{\langle \mathbf{v}, \mathbf{u}_m \rangle}{\langle \mathbf{u}_m, \mathbf{u}_m \rangle} \mathbf{u}_m \end{aligned}$$

şi pentru orice $\mathbf{v} \in V$ are loc relaţia:

$$d(\mathbf{v}, \mathsf{pr}_U \mathbf{v}) = \min\{d(\mathbf{v}, \mathbf{u}) \mid \mathbf{u} \in U\}$$

Demonstraţie. Exerciţiu (similar demonstraţiei Corolarului 7.21). □

Putem acum încheia demonstraţia Propoziţiei 7.32. Fie $\mathbf{u} \in (U^{\perp})^{\perp}$. Atunci există şi sunt unici vectorii $\mathbf{w}_1 \in U, \mathbf{w}_2 \in U^{\perp}$ astfel încât $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$. Atunci

$$0 = \langle \mathbf{u}, \mathbf{w}_2 \rangle = \langle \mathbf{w}_1 + \mathbf{w}_2, \mathbf{w}_2 \rangle = \langle \mathbf{w}_1, \mathbf{w}_2 \rangle + \langle \mathbf{w}_2, \mathbf{w}_2 \rangle = \langle \mathbf{w}_2, \mathbf{w}_2 \rangle$$

Rezultă $\mathbf{w}_2 = \mathbf{0}$, deci $\mathbf{u} = \mathbf{w}_1 \in U$.

Exerciţiu 7.37. Arătaţi că dacă $W_1, W_2 \subseteq V$ sunt subspaţii ale unui spaţiu euclidian $(V, \langle\ ,\ \rangle)$, atunci:

1) $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$.
2) $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$.

7.3 Matrice ortogonale şi descompunerea QR a unei matrice

Am văzut în secţiunile anterioare că existenţa unei baze ortonormate este o proprietatea extrem de utilă în studiul unui (sub)spaţiu euclidian. În particular, ne vom uita la subspaţii de tipul $\mathsf{Im}(A)$.

Matrice ortogonală

Definiţie 7.38. O matrice $Q \in \mathcal{M}_{m,n}(\mathbb{R})$ se numeşte **ortogonală** dacă $Q^tQ = I_n$.

Cel mai simplu exemplu de matrice ortogonală este, evident, matricea identică. Însă definiţia de mai sus nu presupune că o matrice ortogonală este neapărat şi pătratică. Pentru a găsi exemple relevante, a înţelege mai bine matricele ortogonale şi legătura acestora cu subiectul acestui capitol, vom demonstra următorul rezultat:

Proprietăţile matricelor ortogonale

Propoziţie 7.39. Fie $Q \in \mathcal{M}_{m,n}(\mathbb{R})$, $Q = \big(\mathbf{q}_1 \mid \ldots \mid \mathbf{q}_n\big)$. Atunci următoarele afirmaţii sunt echivalente:

1) Q este o matrice ortogonală.
2) $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ este o mulţime de vectori ortonormaţi (în raport cu produsul scalar canonic pe $\mathbb{R}^m$)
3) $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$, $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
4) $\|Q\mathbf{x}\| = \|\mathbf{x}\|$, $\forall\, \mathbf{x} \in \mathbb{R}^n$

Demonstraţie. Vom demonstra 1) $\Rightarrow$ 3) $\Rightarrow$ 2) $\Rightarrow$ 1) şi 3) $\iff$ 4).

1) $\Rightarrow$ 3) Avem $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^t(Q\mathbf{y}) = \mathbf{x}^tQ^tQ\mathbf{y} = \mathbf{x}^t\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$.

3) $\Rightarrow$ 2) Deoarece $Q\mathbf{e}_i = \mathbf{q}_i$, unde $\mathbf{e}_1, \ldots, \mathbf{e}_n$ este baza canonică din $\mathbb{R}^n$, avem $\langle \mathbf{q}_i, \mathbf{q}_j\rangle = \langle Q\mathbf{e}_i, Q\mathbf{e}_j\rangle = \langle \mathbf{e}_i, \mathbf{e}_j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$, deci vectorii $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ sunt ortonormaţi.

2) $\Rightarrow$ 1) Componentele matricei Q^tQ sunt $(Q^tQ)_{ij} = \mathbf{q}_i^t\mathbf{q}_j$, de unde ţinând cont de ortonormalitatea vectorilor $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ rezultă afirmaţia dorită.

3) $\Rightarrow$ 4) Este suficient să luăm $\mathbf{y} = \mathbf{x}$ în relaţia 3).

4) $\Rightarrow$ 3) Se ţine cont de formula de calcul a produsului scalar cu ajutorul normei 14. $\square$

În particular, să reţinem din cele demonstrate mai sus că prin înmulţirea cu o matrice ortogonală, se păstrează

1) Produsul scalar: $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$
2) Norma: $\|Q\mathbf{x}\| = \|\mathbf{x}\|$
3) Distanţa dintre vectori: $d(Q\mathbf{x}, Q\mathbf{y}) = d(\mathbf{x}, \mathbf{y})$

4) Unghiul dintre vectori: $\sphericalangle(Q\mathbf{x}, Q\mathbf{y}) = \sphericalangle(\mathbf{x}, \mathbf{y})$ (a se vedea şi Observaţia 7.8)

Corolar 7.40 $(m = n)$**.** Fie $Q \in \mathcal{M}_n(\mathbb{R})$, $Q = \big(\mathbf{q}_1 \mid \ldots \mid \mathbf{q}_n\big)$. Atunci Q este ortogonală $\iff$ $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ bază ortonormată în $\mathbb{R}^n$ $\iff$ Q este inversabilă şi $Q^{-1} = Q^t$

Vom reformula mai jos algoritmul de ortogonalizare Gram-Schmidt în varianta matriceală. Mai precis, putem factoriza **orice matrice ale cărei coloane sunt liniar independente** sub forma unui produs dintre o matrice **ortogonală** şi şi o matrice **inversabilă superior triunghiulară**.

Descompunerea QR a unei matrice

Propoziţie 7.41. Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$ o matrice cu $\mathsf{rang}(A) = n$ (coloanele matricei A sunt vectori liniar independenţi). Atunci există o matrice ortogonală $Q \in \mathcal{M}_{m,n}(\mathbb{R})$ şi o matrice inversabilă superior triunghiulară $R \in \mathcal{M}_n(\mathbb{R})$ astfel încât $A = QR$.

Demonstraţie. Fie $A = (\mathbf{v}_1 | \ldots | \mathbf{v}_n)$. Aplicând algoritmul Gram-Schmidt coloanelor matricei A, obţinem vectorii ortonormaţi $\mathbf{q}_1, \ldots, \mathbf{q}_n$ cu proprietăţile

$$\begin{array}{lcl}
\mathsf{Sp}\{\mathbf{v}_1\} & = & \mathsf{Sp}\{\mathbf{q}_1\} \\
\mathsf{Sp}\{\mathbf{v}_1, \mathbf{v}_2\} & = & \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2\} \\
\ldots & & \ldots \\
\mathsf{Sp}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} & = & \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}
\end{array}$$

În particular, pentru orice $k \in \{1, \ldots, n\}$, vectorii $\mathbf{q}_1, \ldots, \mathbf{q}_k$ formează o bază ortonormată în subspaţiul $\mathsf{Sp}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$, şi pentru fiecare vector $\mathbf{v}_k$ avem relaţiile $\mathbf{v}_k \in \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$ şi $\mathbf{v}_k \notin \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{k-1}\}$. Rezultă că putem scrie, utilizând formula (15)

$$\begin{array}{lcl}
\mathbf{v}_1 & = & \langle \mathbf{v}_1, \mathbf{q}_1 \rangle \mathbf{q}_1 \\
\mathbf{v}_2 & = & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{v}_2, \mathbf{q}_2 \rangle \mathbf{q}_2 \\
\ldots & & \\
\mathbf{v}_n & = & \langle \mathbf{v}_n, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{v}_2, \mathbf{q}_2 \rangle \mathbf{q}_2 + \ldots + \langle \mathbf{v}_n, \mathbf{q}_n \rangle \mathbf{q}_n
\end{array}$$

Rescriind matriceal relaţiile de mai sus, obţinem descompunerea dorită:

$$\underbrace{(\mathbf{v}_1 | \mathbf{v}_2 | \ldots | \mathbf{v}_n)}_{A} = \underbrace{(\mathbf{q}_1 | \mathbf{q}_2 | \ldots | \mathbf{q}_n)}_{Q} \underbrace{\begin{pmatrix} \langle \mathbf{v}_1, \mathbf{q}_1 \rangle & \langle \mathbf{v}_2, \mathbf{q}_1 \rangle & \ldots & \langle \mathbf{v}_n, \mathbf{q}_1 \rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2 \rangle & \ldots & \langle \mathbf{v}_n, \mathbf{q}_2 \rangle \\ \ldots & & & \\ 0 & 0 & \ldots & \langle \mathbf{v}_n, \mathbf{q}_n \rangle \end{pmatrix}}_{R}$$

Matricea Q este ortogonală, deoarece vectorii $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ sunt ortonormaţi, iar R este superior triunghiulară prin construcţie, şi inversabilă deoarece fiecare dintre produsele scalare $\langle \mathbf{v}_k, \mathbf{q}_k \rangle$ este nenul ($\mathbf{v}_k \in \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$ şi $\mathbf{v}_k \notin \mathsf{Sp}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_{k-1}\}$).

$\square$

Observaţii 7.42. 1) Din demonstraţia de mai sus, rezultă că $\mathsf{Im}(A) = \mathsf{Im}(Q)$ şi că vectorii coloane ale matricei Q formează o bază ortonormată în $\mathsf{Im}(A)$. În particular, proiecţia unui vector $\mathbf{v} \in \mathbb{R}^m$ pe subspaţiul $\mathsf{Im}(A)$ este

$$\mathsf{pr}_{\mathsf{Im}(A)}\mathbf{v} = \sum_{i=1} \langle \mathbf{q}_i, \mathbf{v}\rangle \mathbf{q}_i = \sum_{i=1} (\mathbf{q}_i^t \mathbf{v})\mathbf{q}_i = (\mathbf{q}_1 | \dots | \mathbf{q}_n) \begin{pmatrix} \mathbf{q}_1^t \mathbf{v} \\ \hline \vdots \\ \hline \mathbf{q}_n^t \mathbf{v} \end{pmatrix} = (\mathbf{q}_1 | \dots | \mathbf{q}_n) \begin{pmatrix} \mathbf{q}_1^t \\ \hline \vdots \\ \hline \mathbf{q}_n^t \end{pmatrix} \mathbf{v} = QQ^t\mathbf{v}$$

Să reţinem deci formula proiecţiei utilizând o matrice ortogonală (ale cărei coloane formează o bază ortonormată în subspaţiul respectiv)

$$\mathsf{pr}_{\mathsf{Im}(A)}\mathbf{v} = QQ^t\mathbf{v} \tag{16}$$

valabilă pentru orice matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ cu $\mathsf{rang}(A) = n$, unde $A = QR$ este descompunerea QR a matricei A.[25]

2) Matricea R poate fi aleasă astfel încât componentele sale de pe diagonală să fie pozitive: dacă $r_{ii} < 0$, atunci înlocuim $\mathbf{q}_i$ cu $-\mathbf{q}_i$ şi r_{ii} cu $-r_{ii}$.

Exemplu 7.43. Să reluăm Exemplul 7.29. Fie A matricea având coloanele

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 3 \\ 0 \\ -3 \\ 3 \end{pmatrix}$$

adică

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

În urma aplicării algoritmului Gram-Schmidt s-au obţinut vectorii ortonormaţi

$$\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \mathbf{q}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ -2 \end{pmatrix}, \mathbf{q}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

(în particular, de aici rezultă şi că $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sunt liniar independenţi, adică $\mathsf{rank}(A) = 3$).

Atunci $A = QR$, unde

$$Q = (\mathbf{q}_1 | \mathbf{q}_2 | \mathbf{q}_3) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

şi

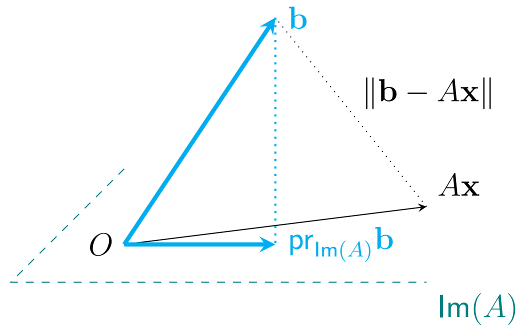
$$R = \begin{pmatrix} \langle \mathbf{v}_1, \mathbf{q}_1\rangle & \langle \mathbf{v}_2, \mathbf{q}_1\rangle & \langle \mathbf{v}_3, \mathbf{q}_1\rangle \\ 0 & \langle \mathbf{v}_2, \mathbf{q}_2\rangle & \langle \mathbf{v}_3, \mathbf{q}_2\rangle \\ 0 & 0 & \langle \mathbf{v}_3, \mathbf{q}_3\rangle \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 3\sqrt{2} \\ 0 & \sqrt{6} & -\sqrt{6} \\ 0 & 0 & \sqrt{3} \end{pmatrix}$$

[25] De fapt formula este valabilă pentru orice matrice A şi orice matrice $Q = (\mathbf{q}_1 | \dots | \mathbf{q}_r)$ ale cărei coloane formează o **bază ortonormată** pentru spaţiul coloanelor matricei A, adică pentru $\mathsf{Im}(A)$.

7.4 Aproximarea soluţiei unui sistem liniar incompatibil

În cazul în care rezolvarea unui sistem de ecuaţii liniare provine din modelarea unei situaţii fizice, există posibilitatea ca datorită erorilor (de rotunjire, de zgomot, etc.) apărute în urma culegerii datelor numerice, ca problema să nu aibă soluţie. În aceste condiţii se caută o "aproximare" care să simuleze totuşi soluţia non-existentă. În această secţiune vom da o definiţie matematică a noţiunii de "aproximare" şi vom vedea cum se poate determina.

Fie A o matrice din $\mathcal{M}_{m,n}(\mathbb{R})$ şi $\mathbf{b}$ un vector real cu m componente, astfel încât $\mathbf{b} \notin \mathsf{Im}(A)$. Echivalent, sistemul $A\mathbf{x} = \mathbf{b}$ este incompatibil.



În aceste condiţii, căutăm $\mathbf{x} \in \mathbb{R}^n$ astfel încât eroarea $d(\mathbf{b}, A\mathbf{x}) = \|\mathbf{b} - A\mathbf{x}\|$ să fie minimă. Dar conform Corolarului 7.36, $\min\limits_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\| = \left\|\mathbf{b} - \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b}\right\|$.

Pseudosoluţia unui sistem incompatibil

Definiţie 7.44. Soluţia sistemului (**compatibil**) $A\mathbf{x}^* = \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b}$ se numeşte **pseudosoluţia** sistemului (incompatibil) $A\mathbf{x} = \mathbf{b}$ în sensul **metodei celor mai mici pătrate (mcmmp)**.

Să notăm $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $A = \left(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n\right)$ şi $\mathbf{b} \in \mathbb{R}^m$. Pentru determinarea pseudosoluţiei sistemului $A\mathbf{x} = \mathbf{b}$, primul pas este aflarea proiecţiei vectorului $\mathbf{b}$ pe $\mathsf{Im}(A)$şi apoi rezolvarea sistemului $A\mathbf{x}^* = \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b}$. Prezentăm mai jos două metode:[26]

Metoda I

1) Determinăm o bază **ortonormată** $\{\mathbf{q}_1, \ldots, \mathbf{q}_r\}$ în $\mathsf{Im}(A)$ (aplicând algoritmul Gram-Schmidt coloanelor matricei A), unde $r = \mathsf{rang}(A)$.

2) Calculăm $\mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b} = \mathsf{pr}_{\mathbf{q}_1}\mathbf{b} + \ldots + \mathsf{pr}_{\mathbf{q}_r}\mathbf{b} = QQ^t\mathbf{b}$, conform formulei (16), unde $Q = (\mathbf{q}_1 | \ldots | \mathbf{q}_r)$.

3) Rezolvăm sistemul compatibil $A\mathbf{x}^* = \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b}$.

Observaţii 7.45. Observaţie: În cazul în care $\mathsf{rang}A = n$, metoda descrisă mai sus se reduce la 1) determinarea descompunerii QR a matricei, şi 2)-3) rezolvarea sistemului $QR\mathbf{x}^* = QQ^t\mathbf{b}$. Multiplicând la stânga cu Q^t şi ţinând cont că Q este ortogonală, rezultă sistemul triunghiular

$$R\mathbf{x}^* = Q^t\mathbf{b}$$

[26] Cu precizarea că deşi a doua metodă este de regulă mai atractivă la prima vedere, este mult mai instabilă din punct de vedere al erorilor de calcul apărute de exemplu la rotunjiri, şi de aceea în calculele efectuate cu ajutorul maşinilor de calcul, prima metodă este cea implementată şi utilizată.

de unde obţinem că pseudosoluţia este

$$\mathbf{x}^* = R^{-1}Q^t\mathbf{b}$$

Metoda II

1) Această metodă pleacă de la observaţia că un vector $\mathbf{x}^* \in \mathbb{R}^n$ este pseudosoluţie a sistemului $A\mathbf{x} = \mathbf{b}$ dacă şi numai dacă este soluţie a sistemului

$$A^tA\mathbf{x}^* = A^t\mathbf{b}$$

Într-adevăr, $\mathbf{x}^*$ satisface $A\mathbf{x}^* = \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b}$ dacă şi numai dacă $\mathbf{b} - A\mathbf{x}^* = \mathbf{b} - \mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b} \in \mathsf{Im}(A)^\perp = \mathsf{Ker}(A^t)$, deci dacă şi numai dacă $A^t(\mathbf{b} - A\mathbf{x}^*) = \mathbf{0}$. Echivalent, $A^tA\mathbf{x}^* = A^t\mathbf{b}$.

2) O a doua observaţie este că sistemul

$$A^tA\mathbf{x}^* = A^t\mathbf{b}$$

are soluţie pentru orice $\mathbf{b} \in \mathbb{R}^m$, şi soluţia este unică dacă şi numai dacă $\mathsf{rang}A = n \leq m$ (coloanele matricei A sunt liniar independente).

Demonstraţie. Pasul I. $\mathsf{Ker}(A^tA) = \mathsf{Ker}(A)$. Dacă $A\mathbf{v} = \mathbf{0}$, atunci $A^tA\mathbf{v} = \mathbf{0}$. Reciproc, dacă $A^tA\mathbf{v} = \mathbf{0}$, atunci $\|A\mathbf{v}\|^2 = \langle A\mathbf{v}, A\mathbf{v}\rangle = \mathbf{v}^tA^tA\mathbf{v} = 0$, de unde rezultă $A\mathbf{v} = \mathbf{0}$.

Pasul II. $\mathsf{Im}(A^tA) = \mathsf{Im}(A^t)$. Dacă $\mathbf{v} \in \mathsf{Im}(A^tA)$, atunci există $\mathbf{u} \in \mathbb{R}^n$ astfel încât $\mathbf{v} = A^tA\mathbf{u} = A^t(A\mathbf{u}) \in \mathsf{Im}(A^t)$. Deci $\mathsf{Im}(A^tA) \subseteq \mathsf{Im}(A^t)$. Dar $\dim_{\mathbb{R}} \mathsf{Im}(A^tA) = n - \dim_{\mathbb{R}} \mathsf{Ker}(A^tA) = n - \dim_{\mathbb{R}} \mathsf{Ker}(A) = \mathsf{rang}A = \mathsf{rang}A^t = \dim_{\mathbb{R}} \mathsf{Im}(A^t)$. Rezultă $\mathsf{Im}(A^tA) = \mathsf{Im}(A^t)$.

Pasul III. Matricea A^tA este inversabilă $\iff \mathsf{Ker}(A^tA) = \{\mathbf{0}\} \iff \mathsf{Ker}(A) = \{\mathbf{0}\} \iff \mathsf{rang}A = n \leq m$. □

3) În particular, dacă $\mathsf{rang}A = n$, atunci proiecţia vectorului $\mathbf{b}$ pe subspaţiul $\mathsf{Im}(A)$ va fi

$$\mathsf{pr}_{\mathsf{Im}(A)}\mathbf{b} = A\mathbf{x}^* = A(A^tA)^{-1}A^t\mathbf{b}$$

Observaţi că formula de mai sus permite determinarea proiecţiei pe subspaţiul $\mathsf{Im}(A)$ fără a cunoaşte o bază ortogonală/ortonormată a acestui subspaţiu.

Exerciţiu 7.46. Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$ o matrice cu proprietatea că $\mathsf{rang}A = n \leq m$ şi să notăm cu $P \in \mathcal{M}_m(\mathbb{R})$ matricea $A(A^tA)^{-1}A^t$. Arătaţi următoarele:

1) $P^2 = P$ (spunem că P este idempotentă).

2) Aplicaţia liniară $f : \mathbb{R}^m \to \mathbb{R}^m$ care asociază fiecărui vector $\mathbf{v} \in \mathbb{R}^m$ proiecţia sa pe subspaţiul $\mathsf{Im}(A)$ este indusă de matricea P, adică $f(\mathbf{v}) = P\mathbf{v}$. În particular, verificaţi că $\mathsf{Im}(f) = \mathsf{Im}(A)$ şi că $\mathsf{Ker}(f) = \mathsf{Im}(A)^\perp = \mathsf{Ker}(A^t)$.

3) Deoarece **orice** subspaţiu $U \subseteq \mathbb{R}^m$ poate fi scris sub forma $\mathsf{Im}(A)$, unde A este matricea ale cărei coloane formează o bază pentru U, rezultatul anterior permite calculul proiecţiei pe orice subspaţiu fără a mai recurge la o bază ortogonală/ortonormată a subspaţiului respectiv.

4) Dacă scriem descompunerea QR a matricei, $A = QR$, atunci verificaţi că $P = QQ^t$, adică regăsim formula (16) din Observaţia 7.42.

Exemplu 7.47. Să considerăm sistemul incompatibil

$$\begin{cases} 2x & = -1 \\ -x + y & = 0 \\ 2y & = 1 \end{cases}$$

Matricea sistemului este $A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix}$, iar vectorul termenilor liberi $\mathbf{b} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. Avem

$$A^t A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \qquad A^t\mathbf{b} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

deci pseudosoluţia sistemului iniţial este soluţie a sistemului

$$\begin{cases} 5x^* - y^* = -2 \\ -x^* + 5y^* = 2 \end{cases}$$

Se obţine $x^* = -\frac{1}{3}, y^* = \frac{1}{3}$. Exerciţiu: Determinaţi pseudosoluţia şi cu Metoda I.

7.5 Aplicaţie: dreapta de regresie

Fie $(x_1, y_1), \ldots, (x_n, y_n)$ puncte distincte în plan, necoliniare. Căutăm o dreaptă de ecuaţie $y = \alpha + \beta x$ care să treacă "cât mai aproape" de punctele date, în sensul minimizării erorii $\sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$ (metoda celor mai mici pătrate).

$$\begin{array}{c}\text{Sistemul} \\ \text{(incompatibil)}\end{array} \begin{cases} y_1 = \alpha + \beta x_1 \\ y_2 = \alpha + \beta x_2 \\ \ldots \\ y_n = \alpha + \beta x_n \end{cases} \quad \text{este echivalent cu} \quad \underbrace{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}}_{\text{matricea } A} \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\text{necunoscute}} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\text{vectorul } \mathbf{b}}$$

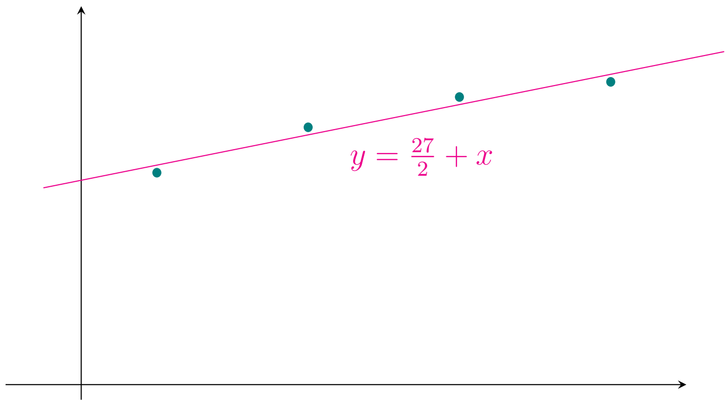
Algebric, aceasta revine la determinarea pseudosoluţiei sistemului:

$$A^t A \begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = A^t \mathbf{b}$$

Dreapta $y = \alpha^* + \beta^* x$ se numeşte **dreapta de regresie** asociată punctelor $(x_1, y_1), \ldots, (x_n, y_n)$.

Exemplu 7.48. Determinaţi dreapta de regresie asociată punctelor

$$(1, 14), (3, 17), (5, 19), (7, 20)$$



$$A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 5 \\ 1 & 7 \end{pmatrix}$$

$$\mathbf{b} = \begin{pmatrix} 14 \\ 17 \\ 19 \\ 20 \end{pmatrix}$$

$$A^t A = \begin{pmatrix} 4 & 16 \\ 16 & 84 \end{pmatrix} \quad A^t \mathbf{b} = \begin{pmatrix} 70 \\ 300 \end{pmatrix} \implies \begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \frac{1}{80} \begin{pmatrix} 84 & -16 \\ -16 & 4 \end{pmatrix} \begin{pmatrix} 70 \\ 300 \end{pmatrix} = \begin{pmatrix} \frac{27}{2} \\ 1 \end{pmatrix}$$

Deci ecuaţia dreptei căutate este $y = \frac{27}{2} + x$.

8 Matrice simetrice

8.1 Valori şi vectori proprii pentru matrice simetrice; descompunere spectrală

Reamintim că o matrice $A \in \mathcal{M}_n(\mathbb{R})$ se numeşte **simetrică** dacă $A = A^t$.

Lemă 8.1. O matrice simetrică are toate valorile proprii **reale**.

Demonstraţie. Fie $\lambda \in \mathbb{C}$ o valoare proprie pentru matricea A şi $\mathbf{v} \in \mathbb{C}^n$ un vector propriu. Atunci $A\mathbf{v} = \lambda\mathbf{v}$, de unde prin conjugare complexă rezultă $A\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$. Putem scrie $\lambda\overline{\mathbf{v}}^t\mathbf{v} = \overline{\mathbf{v}}^t A\mathbf{v} = \overline{\mathbf{v}}^t A^t\mathbf{v} = (A\overline{\mathbf{v}})^t\mathbf{v} = \overline{\lambda}\overline{\mathbf{v}}^t\mathbf{v}$. Dar dacă $\mathbf{v} = (v_i)_{i=\overline{1,n}}$, atunci $\overline{\mathbf{v}}^t\mathbf{v} = \sum_{i=1}^n |v_i|^2 \neq 0$ (pentru că $\mathbf{v} \neq \mathbf{0}$), deci $\lambda = \overline{\lambda}$, adică $\lambda \in \mathbb{R}$. □

Lemă 8.2. Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice simetrică.
Dacă $\lambda_1, \lambda_2 \in \mathbb{R}$ sunt valori proprii **distincte** ale matricei A şi $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ sunt vectori proprii asociaţi valorilor proprii λ_1, λ_2, atunci $\mathbf{v}_1, \mathbf{v}_2$ sunt vectori **ortogonali**.

Demonstraţie. Calculăm $\lambda_1\langle\mathbf{v}_1, \mathbf{v}_2\rangle = \lambda_1\mathbf{v}_1^t\mathbf{v}_2 = (A\mathbf{v}_1)^t\mathbf{v}_2 = \mathbf{v}_1^t A\mathbf{v}_2 = \mathbf{v}_1^t\lambda_2\mathbf{v}_2 = \lambda_2\langle\mathbf{v}_1, \mathbf{v}_2\rangle$, de unde $(\lambda_1 - \lambda_2)\langle\mathbf{v}_1, \mathbf{v}_2\rangle = 0$. Dar $\lambda_1 \neq \lambda_2$, deci $\langle\mathbf{v}_1, \mathbf{v}_2\rangle = 0$, adică vectorii sunt ortogonali. □

Teorema de descompunere spectrală

Teoremă 8.3. Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice. Atunci A este **simetrică** dacă şi numai dacă există o matrice **ortogonală** $Q \in \mathcal{M}_n(\mathbb{R})$ şi o matrice **diagonală** $D \in \mathcal{M}_n(\mathbb{R})$ astfel încât

$$A = QDQ^t$$

În acest caz, mai spunem că A este **orto-diagonalizabilă**.

Observaţie 8.4.

1) Dacă notăm $Q = \big(\mathbf{q}_1 \mid \ldots \mid \mathbf{q}_n\big)$, atunci $\mathbf{q}_1, \ldots, \mathbf{q}_n$ formează o bază ortonormată în $\mathbb{R}^n$.

2) Fie şi $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$, atunci relaţia $A = QDQ^t$ se mai poate scrie $AQ = QD$, adică

$$A\mathbf{q}_1 = \lambda_1\mathbf{q}_1, \ldots, A\mathbf{q}_n = \lambda_n\mathbf{q}_n$$

Rezultă că fiecare λ_i, unde $i = \overline{1,n}$, este valoare proprie pentru A, şi că vectorii $\mathbf{q}_i$ sunt vectori proprii pentru matricea A (formează o bază ortonormată); în particular reţinem că orice matrice simetrică este diagonalizabilă (şi are toate valorile proprii reale).

3) Relaţia $A = QDQ^t$ se mai poate scrie

$$A = \big(\mathbf{q}_1 \mid \ldots \mid \mathbf{q}_n\big)\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\begin{pmatrix} \mathbf{q}_1^t \\ \hline \vdots \\ \hline \mathbf{q}_n^t \end{pmatrix} = \lambda_1\mathbf{q}_1\mathbf{q}_1^t + \ldots + \lambda_n\mathbf{q}_n\mathbf{q}_n^t$$

Să observăm că matricele de forma $\mathbf{q}_i\mathbf{q}_i^t$ sunt matrice de proiecţie pe vectorii proprii $\mathbf{q}_i$ (a se vedea Exerciţiul 7.46), deci formula de mai sus produce o descompunere a matricei A sub forma unei combinaţii liniare de proiecţii pe subspaţiile proprii, în care scalarii sunt exact valorile proprii (de unde şi denumirea de **teorema de descompunere spectrală**).[27]

Demonstraţie. Implicaţia "$\Leftarrow$" este imediată: dacă $A = QDQ^t$, atunci $A^t = (QDQ^t)^t = (Q^t)^t D^t Q^t = QDQ^t = A$.

Vom demonstra implicaţia "$\Rightarrow$" prin inducţie după $n \in \mathbb{N}^*$. Pentru $n = 1$, o matrice (simetrică) se poate identifica de fapt cu un număr real $A = (a)$, şi orto-diagonalizarea revine la a scrie produsul de matrice cu o linie şi cu o coloană $A = (1)(a)(1)^t$. Să presupunem acum că orice matrice simetrică cu n linii şi n coloane este orto-diagonalizabilă. Fie $A \in \mathcal{M}_{n+1}(\mathbb{R})$ o matrice simetrică.

Fie $\lambda_1 \in \mathbb{R}$ o valoare proprie pentru matricea A şi $\mathbf{v}_1 \in V_{\lambda_1}$ un vector propriu corespunzător cu $\|\mathbf{v}_1\| = 1$. Prelungim $\mathbf{v}_1$ la o bază ortonormată $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n+1}$ pentru $\mathbb{R}^{n+1}$. Atunci

$$A\underbrace{(\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_{n+1})}_{Q_1} = \underbrace{(\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_{n+1})}_{Q_1}\underbrace{\left(\begin{array}{c|cc} \lambda_1 & \ldots & \ldots \\ \hline 0 & & \\ \vdots & & A_1 \\ 0 & & \end{array}\right)}_{B}$$

unde $A_1 \in \mathcal{M}_n(\mathbb{R})$. Dacă notăm cu Q_1 matricea ortogonală $(\mathbf{v}_1|\mathbf{v}_2|\ldots|\mathbf{v}_{n+1})$ şi cu B matricea de mai sus formată cu blocurile λ_1, respectiv A_1, atunci $A = Q_1 B Q_1^t$. Dar A este simetrică; rezultă $B^t = (Q_1^t A Q_1)^t = Q_1^t A Q_1 = B$, deci şi B este simetrică. În particular, deducem că

$$B = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right)$$

şi că $A_1^t = A_1$. Putem aplica ipoteza de inducţie matricei A_1: există o matrice ortogonală $Q_2 \in \mathcal{M}_n(\mathbb{R})$ şi o matrice diagonală $D_2 \in \mathcal{M}_n(\mathbb{R})$ astfel încât $A_1 = Q_1 D_2 Q_1^t$.

Fie acum

$$Q = Q_1 \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right) \quad \text{şi} \quad D = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right)$$

Atunci Q este o matrice ortogonală (este un produs de matrice ortogonale) şi D este o matrice diagonală. Avem

$$AQ = AQ_1 \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right)$$

[27] Reamintim că mulţimea valorilor proprii ale matricei se mai numeşte şi spectrul matricei.

$$
\begin{aligned}
&= Q_1 \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right) \\
&= Q_1 \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A_1Q_2 & \\ 0 & & & \end{array}\right) = Q_1 \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2D_2 & \\ 0 & & & \end{array}\right) \\
&= Q_1 \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} \lambda_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & D_2 & \\ 0 & & & \end{array}\right) = QD
\end{aligned}
$$

deci $A = QDQ^t$ şi demonstraţia este încheiată.

□

Exemplu 8.5. Să considerăm matricea simetrică $A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}$. Polinomul caracteristic este

$$P_A(X) = \det(A - XI_3) = \begin{vmatrix} 13-X & -4 & 2 \\ -4 & 13-X & -2 \\ 2 & -2 & 10-X \end{vmatrix} = -(X-9)^2(X-18)$$

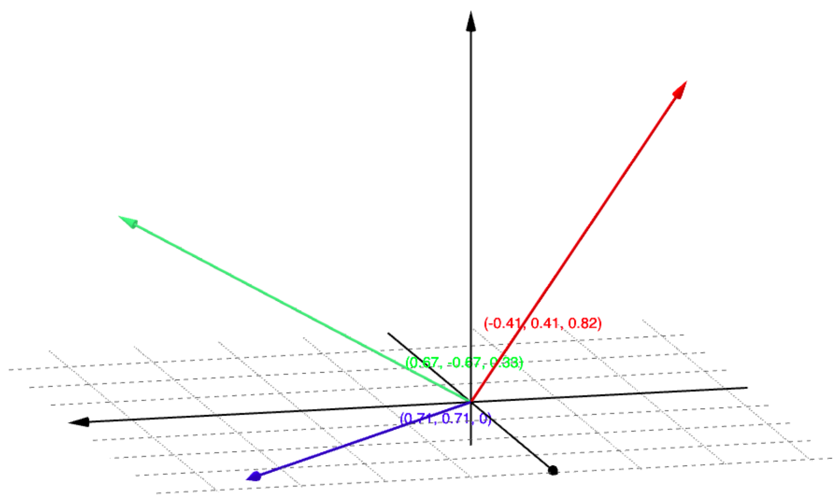
cu rădăcinile (valorile proprii) $\lambda_1 = 9$ (radăcină dublă) şi $\lambda_2 = 18$.

Pentru $\lambda_1 = 9$, vectorii proprii sunt soluţii ale sistemului

$$\begin{cases} 13x - 4y + 2z &= 9x \\ -4x + 13y - 2z &= 9y \\ 2x - 2y + 10z &= 9z \end{cases}$$

Obţinem $V_{\lambda_1} = \{ \begin{pmatrix} x \\ x+z \\ z \end{pmatrix} \mid x, z \in \mathbb{R}\} = \mathsf{Sp}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$. Întrucât urmărim să determinăm câte o bază ortonormată în fiecare subspaţiu propriu, aplicăm algoritmul Gram-Schmidt vectorilor $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Obţinem vectorii ortonormaţi $\mathbf{q}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{q}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$. Putem proceda în continuare analog pentru λ_2: determinăm V_{λ_2} şi găsim o bază ortonormată pentru V_{λ_2}. Pe de altă parte, ştim că A este diagonalizabilă (deci $\dim_{\mathbb{R}} V_{\lambda_2} = 1$) şi că vectorii proprii corespunzători valorilor proprii distincte sunt ortogonali. În particular, de aici rezultă că este suficient să determinăm un singur vector $\mathbf{q}_3 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de normă egală cu unitatea şi ortogonal pe $\mathbf{q}_1$ şi $\mathbf{q}_2$. Condiţiile

$x+y=0, -x+y+2z=0$ conduc la $x=-y=2z$, care coroborate cu $x^2+y^2+z^2=1$ produc vectorul $\mathbf{q}_3 = \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$ (sau opusul acestuia). Am astfel obţinut deci matricea ortogonală $Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{2}{3} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}$ şi matricea diagonală $D = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 18 \end{pmatrix}$ care verifică $A = QDQ^t$. În particular, matricea A se scrie ca o combinaţie liniară $A = 9\mathbf{q}_1\mathbf{q}_1^t + 9\mathbf{q}_2\mathbf{q}_2^t + 18\mathbf{q}_3\mathbf{q}_3^t$ a proiecţiilor ortogonale pe direcţiile vectoriilor proprii (în care coeficienţii sunt valorile proprii), reprezentaţi în figura de mai jos:



8.2 Forme pătratice. Conice şi cuadrice

În acestă secţiune vom discuta câteva aplicaţii ale (orto-diagonalizării) matricelor simetrice în analiza matematică şi în geometrie.

Formă pătratică

Definiţie 8.6. Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice simetrică. Vom numi **formă pătratică** asociată matricei A funcţia

$$q : \mathbb{R}^n \to \mathbb{R}\ ,\ q(\mathbf{v}) = \mathbf{v}^t A \mathbf{v}$$

Dacă notăm $\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ şi $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, atunci forma pătratică q se poate scrie

$$q(\mathbf{v}) = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

(de unde şi denumirea de formă pătratică: q conţine toate monoamele posibile de gradul 2 formate cu $x_1, \ldots, x_n$).

Exemplu 8.7. Pentru $A = I_n$, forma pătratică asociată este $q(\mathbf{v}) = \mathbf{v}^t\mathbf{v} = x_1^2 + \ldots + x_n^2$

Extremele formelor pătratice

Propoziţie 8.8. Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice simetrică şi $q : \mathbb{R}^n \to \mathbb{R}$, $q(\mathbf{v}) = \mathbf{v}^t A\mathbf{v}$ forma pătratică asociată. Dacă $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ este o bază ortonormată pentru $\mathbb{R}^n$ formată cu vectori proprii ai matricei A, corespunzători valorilor proprii $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ (în această ordine), atunci:

1) $\max_{\|\mathbf{v}\|=1} q(\mathbf{v}) = q(\mathbf{v}_1) = \lambda_1$.

2) $\min_{\|\mathbf{v}\|=1} \mathbf{v}^t A\mathbf{v} = q(\mathbf{v}_n) = \lambda_n$.

Demonstraţie. Reamintim că orice matrice simetrică este orto-diagonalizabilă (Teorema 8.1), în particular avem $A = QDQ^t$, cu $Q = (\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n)$ matrice ortogonală şi $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$ matricea diagonală a valorilor proprii. Atunci $q(\mathbf{v}) = \mathbf{v}^t A\mathbf{v} = \mathbf{v}^t QDQ^t\mathbf{v}$; fie $\mathbf{w} = Q^t\mathbf{v}$ (observaţi că $\mathbf{w}$ este vectorul coordonatelor vectorului $\mathbf{v}$ în baza ortonormată $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$). Dacă notăm $\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, atunci $q(\mathbf{v}) = \mathbf{v}^t QDQ^{-1}\mathbf{v} = \mathbf{w}^t D\mathbf{w} = \lambda_1 w_1^2 + \ldots + \lambda_n w_n^2$, deci $q(\mathbf{v}) \leq \lambda_1(w_1^2 + \ldots + w_n^2) = \lambda_1$. În relaţia de mai sus am ţinut cont că $\|\mathbf{w}\| = \|Q^t\mathbf{v}\| = \|\mathbf{v}\| = 1$, deoarece Q (şi implicit Q^t) este o matrice ortogonală, deci păstrează produsul scalar, norma şi distanţa.

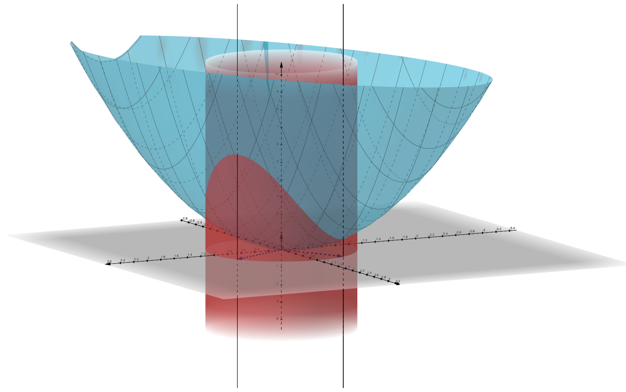
Rezultă că pentru orice vector $\mathbf{v}$ cu $\|\mathbf{v}\| = 1$, are loc inegalitatea $q(\mathbf{v}) \leq \lambda_1$. Pentru inegalitatea inversă, fie $\mathbf{v} = \mathbf{v}_1$. Atunci $q(\mathbf{v}_1) = \mathbf{v}_1^t A\mathbf{v} = \mathbf{v}_1^t\lambda_1\mathbf{v}_1 = \lambda_1$ (reamintim că $\|\mathbf{v}_1\| = 1$). Deci $\max_{\|\mathbf{v}\|=1} q(\mathbf{v}) = q(\mathbf{v}_1) = \lambda_1$. Analog rezultă că $\min_{\|\mathbf{v}\|=1} \mathbf{v}^t A\mathbf{v} = q(\mathbf{v}_n) = \lambda_n$. $\square$

Exemplu 8.9. Să se determine extremele formei pătratice $q\begin{pmatrix} x \\ y \end{pmatrix} = 5x^2 + 4xy + 2y^2$ pe cercul unitate. Matricea asociată este $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$, cu valorile proprii $\lambda_1 = 6 > \lambda_2 = 1$ şi vectorii proprii ortonormaţi $\mathbf{q}_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $\mathbf{q}_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}$. Rezultă

$$\max_{\|\mathbf{v}\|=1} q(\mathbf{v}) = q(\mathbf{q}_1) = \lambda_1 = 6, \ \min_{\|\mathbf{v}\|=1} q(\mathbf{v}) = q(\mathbf{q}_2) = \lambda_2 = 1$$

Interpretare geometrică: $z = 5x^2 + 4xy + 2y^2$ este ecuaţia unui paraboloid eliptic, iar $x^2 + y^2 = 1$ este ecuaţia unui cilindru circular drept.

Problema de optimizare a formei pătratice $q\begin{pmatrix} x \\ y \end{pmatrix} = 5x^2 + 4xy + 2y^2$ pe cercul unitate determină punctele de intersecţie ale celor două suprafeţe pentru care z este maxim, respectiv minim.



Formă pătratică pozitiv/negativ definită

Definiţie 8.10. O formă pătratică $q : \mathbb{R}^n \to \mathbb{R}$, $q(\mathbf{v}) = \mathbf{v}^t A\mathbf{v}$ este **pozitiv definită** dacă $q(\mathbf{v}) > 0$, $\forall\ \mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq \mathbf{0}$.
O formă pătratică este **negativ definită** dacă $q(\mathbf{v}) < 0$, $\forall\ \mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq \mathbf{0}$.
O formă pătratică este **nedefinită** dacă nu este nici pozitiv definită nici negativ definită.

Exemplu 8.11. 1) $q\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2x_1^2 + 5x_2^2 + x_3^2$ este pozitiv definită

2) $q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -x_1^2 + 4x_1x_2 - 5x_2^2$ este negativ definită

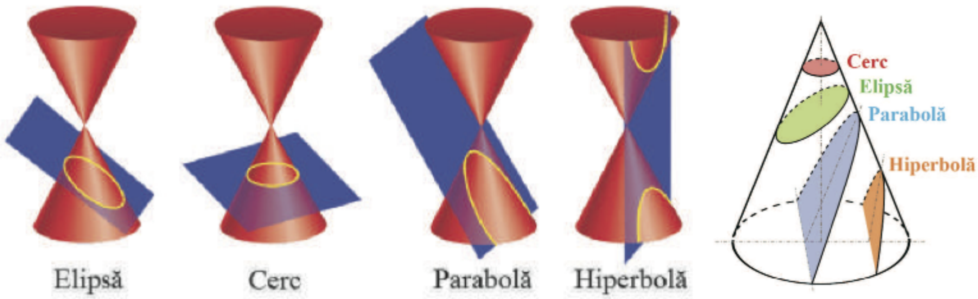
3) $q\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 - x_3^2$ este nedefinită

Propoziţie 8.12. O formă pătratică $q : \mathbb{R}^n \to \mathbb{R}, q(\mathbf{v}) = \mathbf{v}^t A\mathbf{v}$ este **pozitiv definită** (respectiv **negativ definită**) dacă şi numai dacă toate valorile proprii ale matricei A sunt **strict pozitive** (respectiv **strict negative**).

Demonstraţie. Fie QDQ^t orto-diagonalizarea matricei A, unde $D = \mathsf{diag}(\lambda_1, \ldots, \lambda_n)$. Atunci $q(\mathbf{v}) = \mathbf{v}^t QDQ^t\mathbf{v} = (Q^t\mathbf{v})^t D(Q^t\mathbf{v})$. Să observăm că putem extrage valorile proprii ale matricei A din forma pătratică, astfel: $q(Q\mathbf{e}_i) = (Q^tQ\mathbf{e}_i)^t D(Q^tQ\mathbf{e}_i) = \mathbf{e}_i^t D\mathbf{e}_i = \lambda_i$, pentru orice $i = \overline{1, n}$. Atunci ipoteza q pozitiv definită implică $0 < q(\mathbf{e}_i) = \lambda_i$. Reciproc, dacă toţi $\lambda_i > 0$, atunci $q(\mathbf{v}) = (Q^t\mathbf{v})^t D(Q^t\mathbf{v}) = \sum_{i=1}^n \lambda_i x_i'^2 \geq 0$, unde am notat $Q^t\mathbf{v} = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$. Cazul $q(\mathbf{v}) = 0$ conduce la $x_1' = \ldots = x_n' = 0$, de unde $Q^t\mathbf{v} = 0$. Dar Q este inversabilă (cu inversa Q^t), deci $\mathbf{v} = 0$. $\square$

Ca aplicaţii ale formelor pătratice, menţionăm în Analiza Matematică studiul extremelor funcţiilor de mai multe variabile cu ajutorul diferenţialei de ordinul 2, iar în Geometrie, studiul conicelor şi cuadricelor. Primul subiect este acoperit la cursul de Analiză Matematică. Vom discuta mai jos câteva aspecte legate de geometria conicelor şi cuadricelor în care intervin matricele simetrice şi algoritmul de orto-diagonalizare.

Conice. Conicele sunt curbe plane, ce au fost studiate încă din Antichitate[28]. Informal, o conică se poate obţine ca intersecţie a unui plan cu un con circular drept. Matematic, o **conică** este o



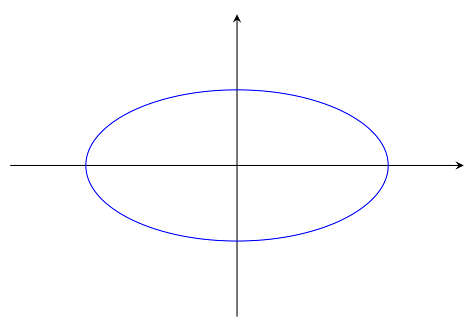
curbă plană Γ având ecuaţia

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \tag{17}$$

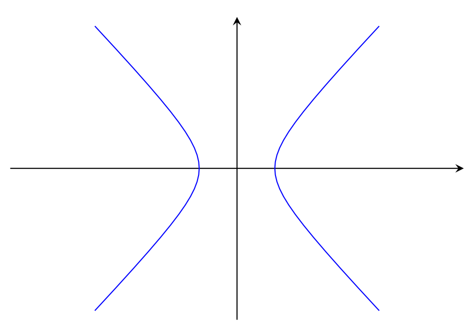
unde $a, b, \ldots \in \mathbb{R}$, cu $a^2 + b^2 + c^2 \neq 0$.

Exemplu 8.13. Prezentăm mai jos fără demonstraţie clasificarea conicelor după forma canonică:[29]

1) Elipsă: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.



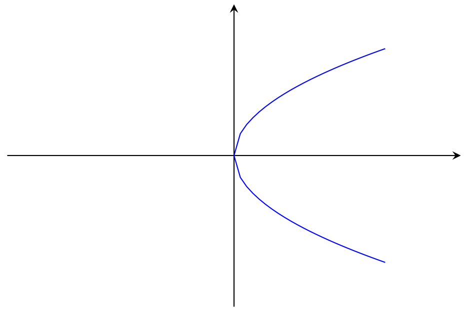
2) Hiperbolă: $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.



[28] Apollonius din Perga, sec. III î. H.

[29] Conicele din acest exemplu se mai numesc şi **conice nedegenerate**, spre deosebire de conicele degenerate (pereche de drepte paralele/concurente/confundate, punct, mulţimea vidă).

3) Parabolă: $y^2 = 2px$



Pentru a înţelege cum putem obţine şi recunoaşte conicele de mai sus, să observăm că ecuaţia 17 se mai poate scrie

$$\begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} a & b \\ b & c \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix} + \underbrace{\begin{pmatrix} d & e \end{pmatrix}}_{\mathbf{b}} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

Dacă ortodiagonalizăm matricea A: $A = QDQ^t$, unde $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, atunci obţinem

$$\begin{pmatrix} x & y \end{pmatrix} QDQ^t \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

În relaţia

$$\begin{pmatrix} x & y \end{pmatrix} QDQ^t \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

notăm $\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^t \begin{pmatrix} x \\ y \end{pmatrix}$; atunci ecuaţia conicei devine

$$\begin{pmatrix} x' & y' \end{pmatrix} D \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} Q \begin{pmatrix} x' \\ y' \end{pmatrix} + f = 0$$

Echivalent,

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0$$

Grupăm în pătrate perfecte:

$$\lambda_1(\underbrace{x' + \ldots}_{x''})^2 + \lambda_2(\underbrace{y' + \ldots}_{y''})^2 = \ldots$$

şi obţinem astfel

$$\lambda_1 x''^2 + \lambda_2 y''^2 = \ldots$$

ecuaţia conicei în forma canonică.

Exemplu 8.14. Să considerăm conica de ecuaţie

$$x^2 + xy + y^2 + x - y = 0$$

Matricial, aceasta devine

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

Orto-diagonalizăm matricea asociată formei pătratice

$$\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}}_{Q^t}$$

şi notăm $\begin{pmatrix} x' \\ y' \end{pmatrix} = Q^t \begin{pmatrix} x \\ y \end{pmatrix}$; atunci ecuaţia conicei devine

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} 3/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = 0$$

Echivalent

$$\frac{3}{2}x'^2 + \frac{1}{2}y'^2 - \sqrt{2}y' = 0$$

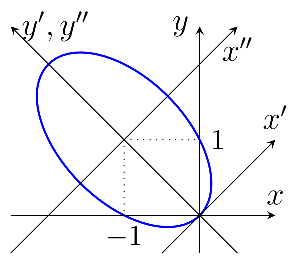
Grupăm în pătrate perfecte:

$$\frac{3}{2}(\underbrace{x'}_{x''})^2 + \frac{1}{2}(\underbrace{y' - \sqrt{2}}_{y''})^2 = 1$$

Obţinem în final o elipsă

$$\frac{3}{2}x''^2 + \frac{1}{2}y''^2 = 1$$

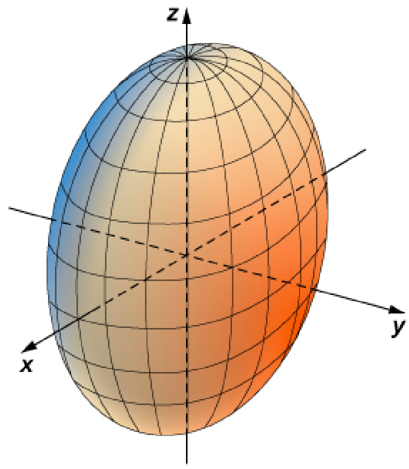
reprezentată grafic mai jos:



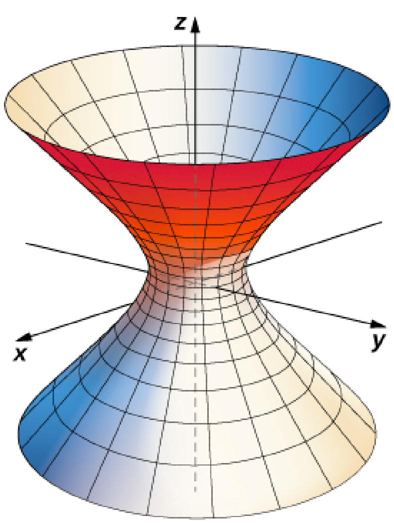
Cuadrice. Din punct de vedere algebric, cuadricele generalizează conicele în spaţiul tridimensional, în sensul că sunt determinate, ca şi acestea din urmă, de o ecuaţie de gradul 2, de data aceasta în x, y, z. Însă spre deosebire de conice, cuadricele sunt suprafeţe, nu curbe. Ecuaţia generală a unei cuadrice este

$$\begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \mathbf{b} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + c = 0$$

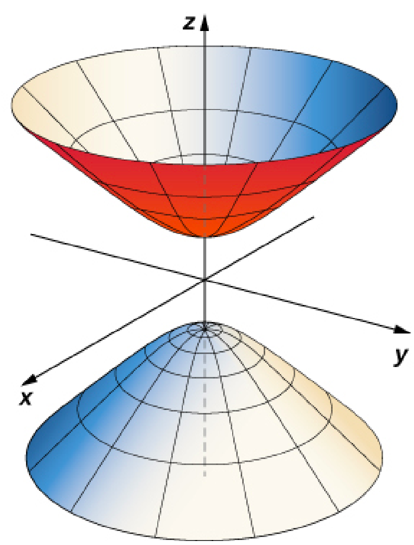
unde $A \in \mathcal{M}_3(\mathbb{R})$ este o matrice simetrică, $\mathbf{b} \in \mathbb{R}^3$, $c \in \mathbb{R}$. Putem aplica aceeaşi metodă de reducere a ecuaţiei unei cuadrice (se orto-diagonalizează matricea A, se schimbă baza/coordonatele şi se realizează în final o translaţie etc.) pentru a o aduce la forma canonică. În figurile din paginile următoare sunt prezentate principalele cuadrice.



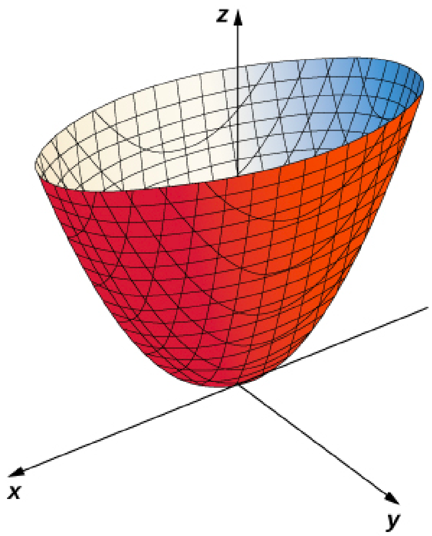
Elipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$



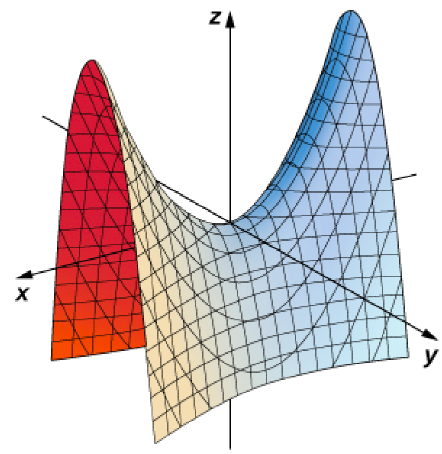
Hiperboloid cu o pânză $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$



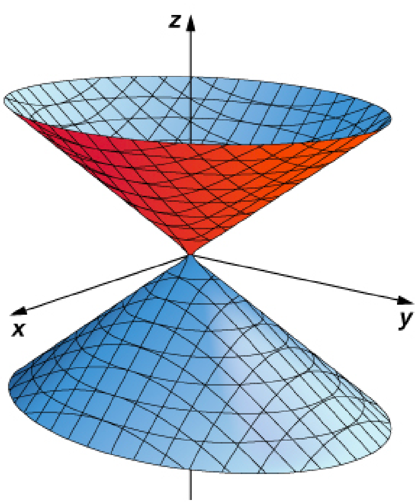
Hiperboloid cu două pânze $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$



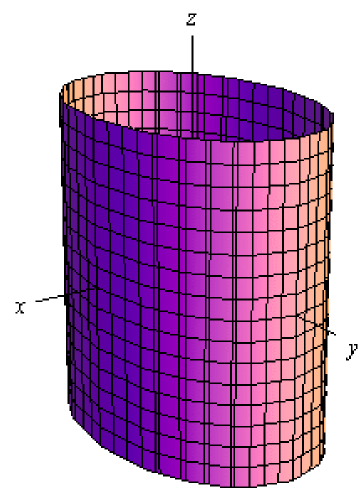
Paraboloid eliptic $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z}{c}$



Paraboloid eliptic $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}$



Con eliptic $\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$



Cilindru eliptic $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

8.3 Descompunerea valorilor singulare (SVD)

Toate rezultatele din algebra liniară obţinute până în acest moment vor fi aplicate acum pentru a obţine o nouă factorizare pentru matrice. Dacă în factorizările precedente am impus diverse restricţii (matrice diagonalizabilă, sau având coloanele liniar independente, sau simetrică), de data aceasta nu există restricţii, rezultatul fiind aplicabil matricelor arbitrare.

Descompunerea valorilor singulare

Teoremă 8.15. Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$ o matrice arbitrară. Atunci există matricile **ortogonale** $U \in \mathcal{M}_m(\mathbb{R})$, $V \in \mathcal{M}_n(\mathbb{R})$ şi scalarii (unici!) $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ (**valorile singulare** ale matricei A) astfel încât

$$A = U \cdot \underbrace{\left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)}_{\Sigma} \cdot V^t$$

În particular, matricea A se poate scrie $A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^t + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^t$, unde $U = \big(\mathbf{u}_1 \mid \ldots \mid \mathbf{u}_m\big)$ şi $V = \big(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n\big)$.

Factorizarea de mai sus se numeşte **descompunerea valorilor singulare** (SVD).[30]

Demonstraţie. (Algoritm pentru determinarea SVD $A = U\Sigma V^t$)

Pasul I: Construcţia matricei V. Matricea $A^tA \in \mathcal{M}_n(\mathbb{R})$ este o matrice simetrică, deci se poate ortodiagonaliza

$$A^tA = \big(\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n\big) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_1^t \\ \hline \vdots \\ \hline \mathbf{v}_n^t \end{pmatrix}$$

În particular, numerotăm valorile proprii astfel încât să fie ordonate descrescător

$$\lambda_1 \geq \ldots \geq \lambda_n$$

Reamintim că vectorii proprii $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sunt **ortonormaţi**. Vom nota cu $V = (\mathbf{v}_1 \mid \ldots \mid \mathbf{v}_n)$ matricea ortogonală ale cărei coloane sunt vectorii proprii ai matricei A^tA de mai sus.

Pasul II: Construcţia matricei Σ. Ştim că $\lambda_i \in \mathbb{R}$, $\forall\, i = \overline{1,n}$, deoarece sunt valorile proprii ale matricei simetrice A^tA. De asemenea, că $A^tA\mathbf{v}_i = \lambda_i \mathbf{v}_i$, de unde rezultă $\|A\mathbf{v}_i\|^2 = \langle A\mathbf{v}_i, A\mathbf{v}_i\rangle = \mathbf{v}_i^t A^t A \mathbf{v}_i = \lambda_i \mathbf{v}_i^t \mathbf{v}_i = \lambda_i \|\mathbf{v}_i\|^2 = \lambda_i$. Rezultă că valorile proprii ale matricei A^tA sunt strict pozitive sau nule:

$$\lambda_1 \geq \ldots \geq \lambda_r > 0 = \lambda_{r+1} = \ldots = \lambda_n$$

De asemenea, reamintim că multiplicitatea (algebrică = geometrică[31] a) valorii proprii nule este dimensiunea nucleului matricei. Rezultă $n - r = \dim_{\mathbb{R}} \mathsf{Ker}(A^tA) = \dim_{\mathbb{R}} \mathsf{Ker}(A) = n - \mathsf{rang}(A)$, deci $r = \mathsf{rang}(A)$.

[30] *Singular Value Decomposition*, în limba engleză.
[31] Pentru că A^tA este diagonalizabilă, fiind simetrică.

Vom numi **valorile singulare** ale matricei A numerele reale pozitive $\sigma_i = \sqrt{\lambda_i}$, $i = \overline{1,r}$ şi vom nota

$$\Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0}_{r,n-r} \\ & & \sigma_r & \\ \hline & \mathbf{0}_{m-r,r} & & \mathbf{0}_{m-r,n-r} \end{array}\right)$$

Pasul III: Construcţia matricei U. Pentru orice $i, j = \overline{1,n}$, avem

$$\langle A\mathbf{v}_i, A\mathbf{v}_j\rangle = \mathbf{v}_i^t A^t A\mathbf{v}_j = \lambda_j \mathbf{v}_i^t \mathbf{v}_j = \lambda_j \langle \mathbf{v}_i, \mathbf{v}_j\rangle = \begin{cases} \lambda_i, & i = j \\ 0 & i \neq j \end{cases}$$

deci vectorii $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$ sunt ortogonali şi $\|A\mathbf{v}_i\|^2 = \lambda_i$, $i = \overline{1,n}$. Rezultă că vectorii $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$, $i = \overline{1,r}$, sunt ortonormaţi (şi $A\mathbf{v}_{r+1} = \ldots = A\mathbf{v}_n = \mathbf{0}$).

Dacă $r < m$, **completăm** $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ la o bază ortonormată $\mathbf{u}_1, \ldots, \mathbf{u}_m$ în $\mathbb{R}^m$. Vom nota

$$U = \big(\mathbf{u}_1 \mid \ldots \mid \mathbf{u}_r \mid \mathbf{u}_{r+1} \mid \ldots \mid \mathbf{u}_m\big)$$

Atunci

$$A\underbrace{(\mathbf{v}_1|\ldots|\mathbf{v}_r|\mathbf{v}_{r+1}|\ldots|\mathbf{v}_n)}_{V} = \underbrace{(\mathbf{u}_1|\ldots|\mathbf{u}_r|\mathbf{u}_{r+1}|\ldots|\mathbf{u}_m)}_{U}\underbrace{\left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)}_{\Sigma}$$

Ţinând cont că V este o matrice ortogonală şi pătratică, relaţia de mai sus se poate scrie

$$A = U\Sigma V^t$$

producând astfel factorizarea dorită. □

Descompunerea valorilor singulare are avantajul de a produce toate informaţiile referitoare la matricea respectivă, în particular rangul, baze ortonormate şi dimensiunile celor 4 subspaţii asociate (nuclee, imagini), ortodiagonalizări ale matricelor simetrice asociate, etc.

Corolar 8.16. Fie $A \in \mathcal{M}_{m,n}(\mathbb{R})$ cu descompunerea singulară $A = U\Sigma V^t$. Atunci:

1) $\mathsf{rang}(A) = r$

2) $\{\mathbf{u}_1, \ldots \mathbf{u}_r\}$ bază ortonormată în $\mathsf{Im}(A)$.

3) $\{\mathbf{u}_{r+1}, \ldots \mathbf{u}_m\}$ bază ortonormată în $\mathsf{Ker}(A^t)$.

4) $\{\mathbf{v}_{r+1}, \ldots \mathbf{v}_n\}$ bază ortonormată în $\mathsf{Ker}(A)$.

5) $\{\mathbf{v}_1, \ldots \mathbf{v}_r\}$ bază ortonormată în $\mathsf{Im}(A^t)$.

6) Descompunerea valorilor singulare pentru A^t este $V\Sigma^t U$.

7) Au loc ortodiagonalizările simultane

$$A^tA = V \begin{array}{c} \begin{array}{cc} \quad r \quad & \quad n-r \end{array} \\ \left(\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & 0 \\ & & \lambda_r & \\ \hline & 0 & & 0 \end{array}\right) \end{array} V^t\,,\; AA^t = U \begin{array}{c} \begin{array}{cc} \quad r \quad & \quad m-r \end{array} \\ \left(\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & 0 \\ & & \lambda_r & \\ \hline & 0 & & 0 \end{array}\right) \end{array} U^t$$

Demonstrație. 1) Am văzut deja în Demonstrația teoremei precedente că $\mathsf{rang}(A) = r$.

2) Deoarece vectorii ortonormați $\mathbf{u}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$, $i = \overline{1,r}$ se găsesc în imaginea matricei A, iar $\dim_{\mathbb{R}} \mathsf{Im}(A) = \mathsf{rang}(A) = r$, rezultă că formează o bază ortonormată în $\mathsf{Im}(A)$.

3) Se obține din $\mathsf{Im}(A)^{\perp} = \mathsf{Ker}(A^t)$.

4) Rezultă din $A\mathbf{v}_{r+1} = \ldots = A\mathbf{v}_n$ și din $\dim_{\mathbb{R}} \mathsf{Ker}(A) = n - r$.

5) Se utilizează relația $\mathsf{Ker}(A)^{\perp} = \mathsf{Im}(A^t)$.

6) Evident.

7) Prima orto-diagonalizare este cea utilizată în producerea descompunerii valorilor singulare, iar a doua rezultă din $AA^t = U\Sigma V^t(U\Sigma V^t)^t = U\Sigma V^tV\Sigma^tU^t = U\Sigma\Sigma^tU^t$.

□

Exemplu 8.17. Fie $A = \begin{pmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{pmatrix}$. Să determinăm descompunerea valorilor singulare pentru matricea A. Calculăm

$$A^tA = \begin{pmatrix} 20 & -10 \\ -10 & 5 \end{pmatrix}$$

cu valorile proprii $\lambda_1 = 25 \geq \lambda_2 = 0$ și vectorii proprii ortonormați

$$\mathbf{v}_1 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}$$

Deci

$$V = (\mathbf{v}_1|\mathbf{v}_2) = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

Avem o singura valoare singulară $\sigma_1 = \sqrt{\lambda_1} = 5$. Rezultă

$$\Sigma = \begin{pmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Calculăm

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}$$

Prelungim la o bază ortonormată $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ în $\mathbb{R}^3$: de exemplu, vectorii

$$\mathbf{u}_1, \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sunt liniar independenţi, iar după aplicarea algoritmului Gram-Schmidt, producem vectorii ortonormaţi

$$\mathbf{u}_1, \mathbf{u}_2 = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \\ 0 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Deci

$$U = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

şi

$$A = \underbrace{\begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}}_{V^t} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^t$$

Observaţie 8.18. (Interpretarea geometrică a SVD)

Pentru o mai bună intuiţie, vom discuta aici cum se poate interpreta geometric descompunerea valorilor singulare. Vom considera cazul unei matrice $A \in \mathcal{M}_2(\mathbb{R})$, de exemplu

$$A = \begin{pmatrix} -2 & 11 \\ -10 & 5 \end{pmatrix}$$

Procedând ca în exemplul precedent, obţinem descompunerea valorilor singulare a matricei A:

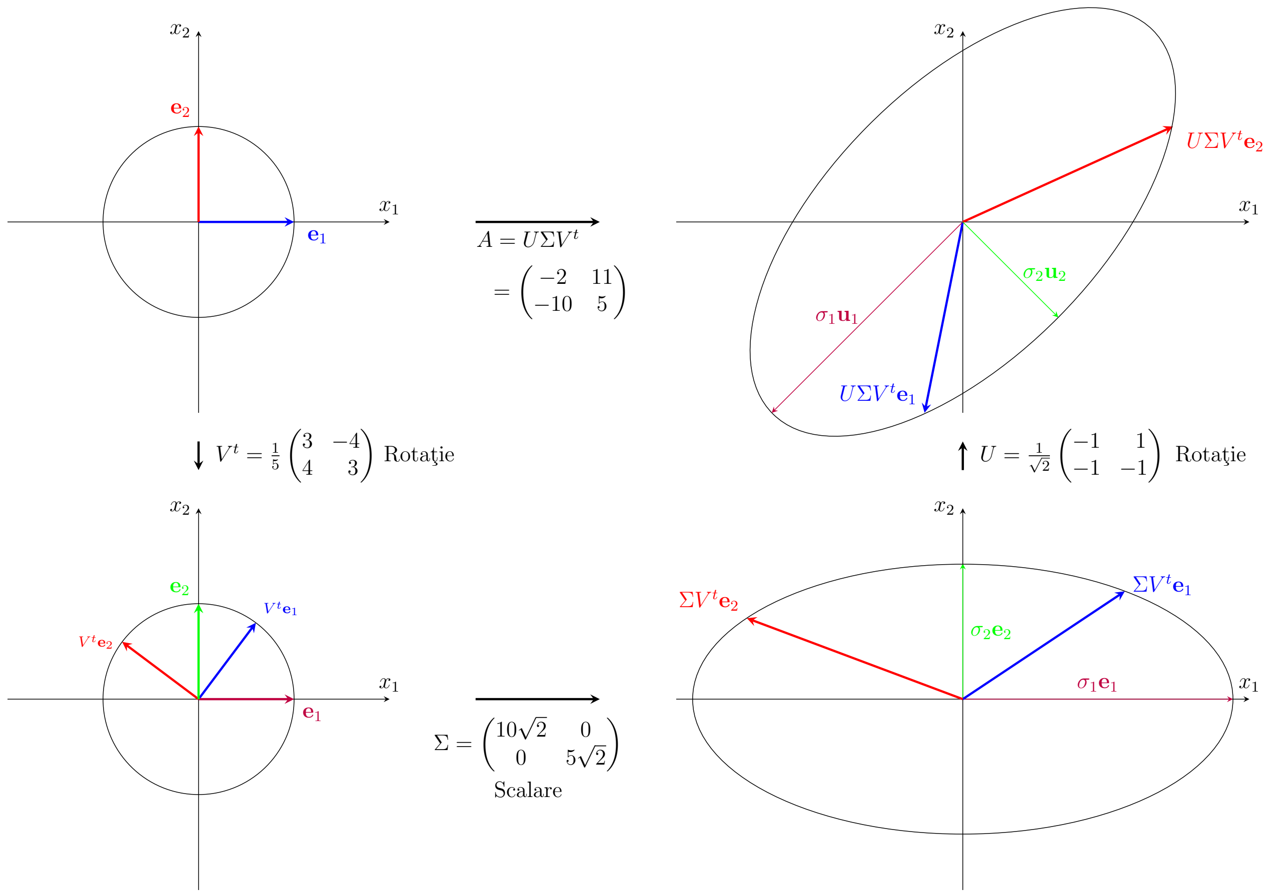
$$A = U\Sigma V^t \ , \quad \text{unde } U = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} \ , \ \Sigma = \begin{pmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{pmatrix} \ , \ V = \frac{1}{5} \begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}$$

Atunci aplicaţia liniară $\mathbb{R}^2 \to \mathbb{R}^2$ indusă de multiplicarea cu matricea A poate fi descrisă geometric ca o succesiune de trei transformări ale planului, după cum urmează:

1) O rotaţie indusă de matricea ortogonală V^t (dacă $\det V = 1$, altfel V^t corespunde unei rotaţii urmate de o simetrie).

2) O scalare a axelor datorată matricei diagonale Σ.

3) O nouă rotaţie indusă de matricea ortogonală U (analog, dacă $\det U = 1$).



Aplicaţie: analiza componentelor principale (PCA)[32] şi recunoaşterea facială. Analiza componentelor principale este o metodă eficientă de compresie a datelor, prin extragerea caracteristicilor principale ale acestora şi proiectarea datelor pe un spaţiu de dimensiune mai mică (spaţiul caracteristicilor principale). Vom ilustra metoda în cazul particular al algoritmului de **recunoaştere facială**, dezvoltat de M. Turk şi A. Pentland în lucrarea *Eigenfaces for Recognition* (Journal of Cognitive Neuroscience 1991). Imaginile de mai jos provin din acest articol.

Pasul I. Antrenare.

1) Input: un set (de antrenare) format din n imagini faciale (training faces), transformate din matrice în vectori $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^m$ (Figura 8)

2) Se determină media imaginilor (mean face) $\boldsymbol{\mu} = \frac{1}{n}(\mathbf{x}_1 + \dots + \mathbf{x}_n)$ (Figura 9)

3) Se centrează informaţiile şi se reprezintă matriceal

$$X = (\mathbf{x}_1 - \boldsymbol{\mu} \mid \dots \mid \mathbf{x}_n - \boldsymbol{\mu}) \in \mathcal{M}_{m,n}(\mathbb{R})$$

4) Se determină descompunerea SVD a matricei "feţelor" centrate: $X = U\Sigma V^t$

5) Se obţine "spaţiul feţelor" $\mathsf{Im}(X)$ cu componentele principale (eigenfaces) $U_r = (\mathbf{u}_1 \mid \dots \mid \mathbf{u}_r)$ (Figura 10), unde $r = \mathsf{rang}(X)$.

[32] *Principal Component Analysis*, în limba engleză.

Figura 8: Setul de antrenare

Figura 9: Media imaginilor

6) Se determină lungimea proiecţiei (ponderea) fiecărui vector (centrat) din spaţiul feţelor $\mathbf{x}_i - \boldsymbol{\mu}$, unde $i = \overline{1,n}$, pe fiecare componentă principală $\mathbf{u}_k$ ($k = \overline{1,r}$)

$$U^t X = \left(\mathbf{u}_k^t(\mathbf{x}_i - \boldsymbol{\mu})\right)_{k=\overline{1,r}; i=\overline{1,n}} \in \mathcal{M}_{r,n}(\mathbb{R})$$

Pasul II: Recunoaştere.

1) Fiecare imagine nouă $\mathbf{x} \in \mathbb{R}^m$ (Figura 11) se centrează $\mathbf{x} \mapsto \mathbf{x} - \boldsymbol{\mu}$.

2) Se determină ponderea proiecţiei noii imagini centrate pe fiecare dintre componentele principale (Figura 12)

$$U^t(\mathbf{x} - \boldsymbol{\mu}) = \left(\mathbf{u}_1^t(\mathbf{x} - \boldsymbol{\mu}) \mid \dots \mid \mathbf{u}_r^t(\mathbf{x} - \boldsymbol{\mu})\right) \in \mathbb{R}^r$$

3) Se caută distanţa minimă în spaţiul feţelor dintre proiecţia imaginii centrate $\mathbf{x} - \boldsymbol{\mu}$ şi (proiecţiile) imaginilor din baza de date pentru recunoaştere (best match)

$$\min_{i=\overline{1,n}} \sqrt{\sum_k |u_k^t(\mathbf{x} - \boldsymbol{\mu}) - u_k^t(\mathbf{x}_i - \boldsymbol{\mu})|^2}$$

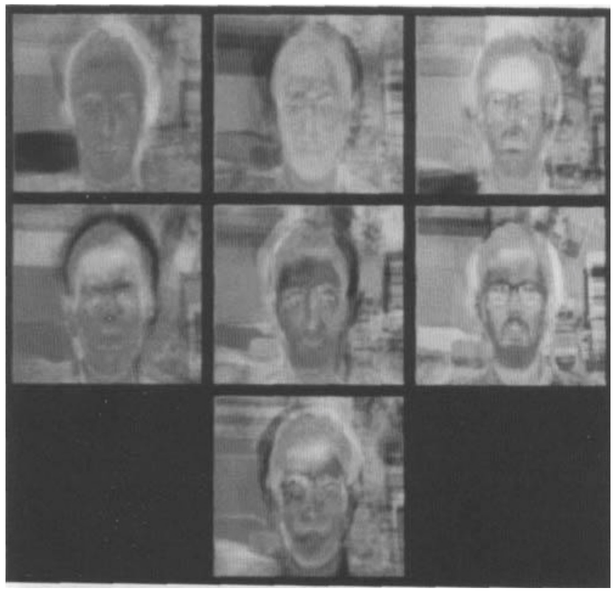
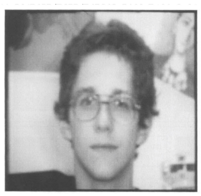


Figura 10: Spaţiul feţelor



Imagine nouă

Proiecţia noii imagini pe spaţiul feţelor

De regulă, se stabileşte şi un nivel admis al erorii. Dacă acesta este depăşit, se consideră că noua faţă nu este recunoscută şi se adaugă setului de antrenare, repetând algoritmul de mai sus pentru generarea unui nou set de componente principale.

Exerciţiu 8.19. Citiţi articolul menţionat mai sus pentru a afla cum s-a finalizat algoritmul de recunoaştere a imaginii din Figura 11.

Capitolul II

Ecuaţii diferenţiale liniare

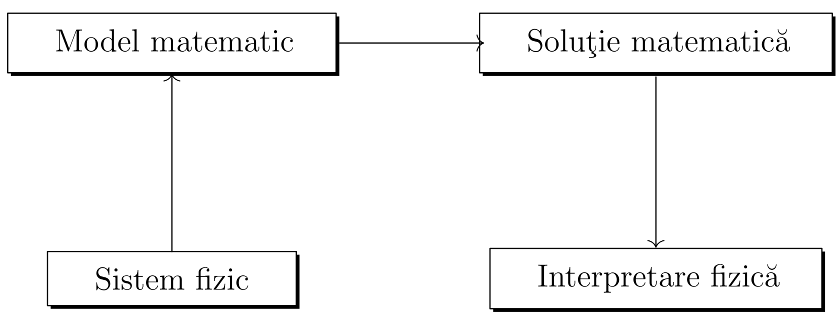
Among all of the mathematical disciplines the theory of differential equations is the most important ... It furnishes the explanation of all those elementary manifestations of nature which involve time.

Sophus Lie (1842-1899)

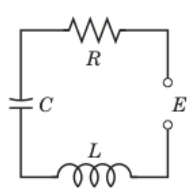
1 Ecuaţii diferenţiale de ordinul I

1.1 Noţiuni introductive

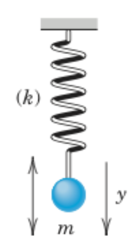
Din punct de vedere practic, ecuaţiile diferenţiale reprezintă o etapă a paradigmei *modelare, rezolvare, interpretare* aplicabilă oricărui sistem dinamic a cărui evoluţie se defăşoară în timp continuu:



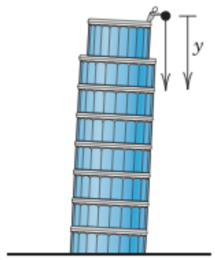
Menţionăm mai jos câteva exemple binecunoscute:



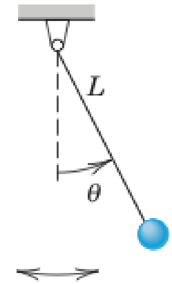
Intensitatea curentului I
într-un circuit RLC
$LI' + RI + \frac{1}{C}\int_0^t I\mathrm{d}t = E$



Resortul elastic
$my'' + ky = 0$



Căderea liberă
$y'' = g$



Oscilaţia pendulului
$L\theta'' + g \sin\theta = 0$

Definiţie 1.1. Fie $F : I \times D \to \mathbb{R}^n$ o funcţie continuă, unde $I \subseteq \mathbb{R}$ este un interval deschis şi $D \subseteq \mathbb{R}^n$ o mulţime deschisă.
O **ecuaţie diferenţială** de ordinul I are forma (explicită)

$$\mathbf{x}' = F(t, \mathbf{x}) \tag{18}$$

unde

- t este variabila independentă;
- $\mathbf{x}$ este funcţia necunoscută şi $\mathbf{x}'$ este derivata sa.

O **soluţie** a ecuaţiei (18) este o funcţie derivabilă $\mathbf{x} : J \to D \subseteq \mathbb{R}^n$, unde J este un interval deschis cu $J \subseteq I$, care satisface

$$\mathbf{x}'(t) = F(t, \mathbf{x}(t)), \forall t \in J$$

Problema Cauchy pentru ecuaţia diferenţială (18)

$$\begin{cases} \mathbf{x}' &= F(t, \mathbf{x}) \\ \mathbf{x}(t_0) &= \mathbf{x}_0 \end{cases}$$

constă în determinarea unei soluţii $\mathbf{x} = \mathbf{x}(t)$ care verifică **condiţia iniţială** $\mathbf{x}(t_0) = \mathbf{x}_0$ (unde $t_0 \in I$, $\mathbf{x}_0 \in D$).

În teoria ecuaţiilor diferenţiale, se întâlnesc probleme ca:

1) Existenţa şi unicitatea soluţiilor.
2) Prelungirea soluţiilor pe un domeniu dat.
3) Stabilitatea soluţiilor.
4) Determinarea soluţiilor cu anumite proprietăţi (de exemplu, problema Cauchy).

Referitor la existenţa şi unicitatea soluţiilor, menţionăm următorul rezultat (fără demonstraţie):[33]

[33] Pentru mai multe detalii se poate consulta [7].

Teorema Cauchy-Lipschitz de existenţă şi unicitate a soluţiei problemei Cauchy

Teoremă 1.2. Fie $F : I \times D \to \mathbb{R}^n$ o funcţie continuă, cu proprietatea că există o constantă pozitivă $L > 0$ astfel încât pentru orice $(t_0, x_0) \in I \times D$ există $\delta, \delta' > 0$ cu proprietatea că pentru orice $|t - t_0| < \delta, \|x_1 - x_0\| < \delta', \|x_2 - x_0\| < \delta'$, are loc relaţia

$$\|f(t, x_1) - f(t, x_2)\| \leq K\|x_1 - x_2\|$$

Atunci pentru orice $t_0 \in I$, $\mathbf{x}_0 \in D$, există un interval deschis $J \subseteq I$ cu $t_0 \in J$ astfel încât problema Cauchy

$$\begin{cases} \mathbf{x}' & = & F(t, \mathbf{x}) \\ \mathbf{x}(t_0) & = & \mathbf{x}_0 \end{cases}$$

admite soluţie unică pe intervalul J.

În particular, orice funcţie F continuă pentru care $\dfrac{\partial F}{\partial x}$ există şi este continuă (de exemplu, orice funcţie de clasă $\mathcal{C}^1$) satisface condiţiile teoremei.

Soluţia se obţine prin metoda aproximaţiilor succesive (Picard): $\mathbf{x}(t) = \lim\limits_{n\to\infty} \mathbf{x}_n(t)$, unde şirul de funcţii $\mathbf{x}_n(t)$ este

$$\begin{cases} \mathbf{x}_0(t) & = & \mathbf{x}_0 \\ \mathbf{x}_{n+1}(t) & = & \mathbf{x}_0 & + & \displaystyle\int_{t_0}^{t} F(s, \mathbf{x}_n(s))\mathsf{d}s\,, \ n \geq 0 \end{cases}$$

1.2 Ecuaţi diferenţiale liniare de ordinul I

În continuare, vom studia doar un sigur tip de ecuaţii diferenţiale de ordinul I, şi anume cele liniare. O ecuaţie diferenţială $x' = F(t, x)$ se numeşte **liniară** dacă funcţia $F(t, x)$ este liniară în x (dar nu neapărat şi în variabila t). Pentru o mai bună intuiţie, vom începe cu cel mai simplu caz, urmând să creştem gradual complexitatea ecuaţiei diferenţiale.

1) **Ecuaţia diferenţială** $x' = \lambda x \quad (\lambda \in \mathbb{R})$

   Acesta este un exemplu de ecuaţie liniară *autonomă* (în care variabila t nu apare explicit), şi poate cel mai important exemplu de ecuaţie diferenţială liniară pe care o vom întâlni. Pentru a evidenţia că necunoscuta este o funcţie care depinde de t, ecuaţia se mai scrie uneori şi sub forma

$$\frac{\mathsf{d}x}{\mathsf{d}t} = \lambda x$$

   Pentru $\lambda = 1$ avem ecuaţia diferenţială $x' = x$, pentru care o soluţie *celebră* este funcţia exponenţială $x : \mathbb{R} \to \mathbb{R}, x(t) = \mathsf{e}^t$. Pentru λ arbitrar, este uşor de verificat că funcţia $x : \mathbb{R} \to \mathbb{R}, x(t) = \mathsf{e}^{\lambda t}$ este soluţie a a ecuaţiei $x' = \lambda x$. Dar mai există şi alte soluţii?

   Pentru a putea răspunde, vom proceda astfel: înmulţim ambii membri ai ecuaţiei diferenţiale cu $\mathsf{e}^{-\lambda t}$, şi obţinem $x'\mathsf{e}^{-\lambda t} - \lambda x\mathsf{e}^{-\lambda t} = 0$. Dar membrul stâng este derivata funcţiei $e^{-\lambda t}x$, de unde deducem că **soluţia generală** a ecuaţiei diferenţiale

$$x' = \lambda x \tag{19}$$

este

$$\boxed{x(t) = C\mathrm{e}^{\lambda t}} \tag{20}$$

unde C este o constantă reală.

Mai departe, obţinem că soluţia problemei Cauchy

$$\begin{cases} x' &= \lambda x \\ x(t_0) &= x_0 \end{cases}$$

este

$$x(t) = x_0 \mathrm{e}^{\lambda(t-t_0)}$$

2) **Ecuaţia diferenţială** $x' = f(t)x$ (unde $f : I \subseteq \mathbb{R} \to \mathbb{R}$ este o funcţie continuă)

Fie F o primitivă a funcţiei f (care există datorită ipotezei de continuitate a lui f). Atunci multiplicând ecuaţia diferenţială $x' = f(t)x$ cu $\mathrm{e}^{-F(t)}$ şi procedând ca în cazul precedent, rezultă că **soluţia generală** este

$$x(t) = C\mathrm{e}^{F(t)}$$

unde $C \in \mathbb{R}$ este o constantă.

În cazul problemei Cauchy

$$\begin{cases} x' &= f(t)x \\ x(t_0) &= x_0 \end{cases}$$

putem obţine un rezultat mult mai precis (reamintim că primitiva unei funcţii continue este unică până la o constantă), alegând $F(t) = \int_{t_0}^{t} f(s)\mathrm{d}s$. Soluţia este atunci

$$x(t) = x_0 \mathrm{e}^{\int_{t_0}^{t} f(s)\mathrm{d}s}$$

Analizând soluţia generală a ecuaţiei diferenţiale $x' = f(t)x$, putem deduce următoarele:

Propoziţie 1.3. Fie $f : I \to \mathbb{R}$ o funcţie continuă, unde $I \subseteq \mathbb{R}$ este un interval. Atunci mulţimea soluţiilor ecuaţiei diferenţiale $x' = f(t)x$

$$\mathcal{S}_o = \{x : I \to \mathbb{R} \mid x \text{ funcţie derivabilă şi } x' = f(t)x\}$$

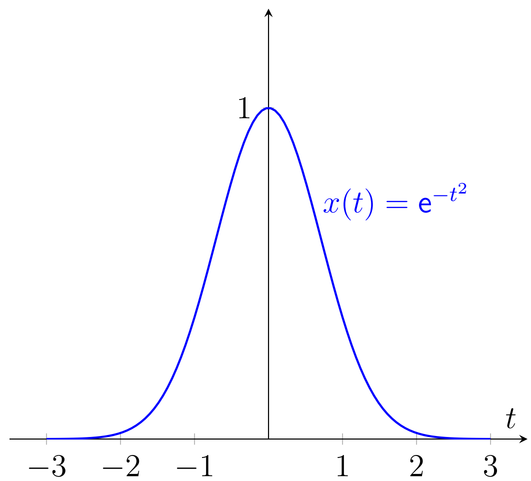
este un $\mathbb{R}$-spaţiu vectorial în raport cu operaţiile uzuale de adunare şi înmulţire cu scalari şi

$$\dim_{\mathbb{R}} \mathcal{S}_o = 1$$

O bază în acest spaţiu vectorial este $\{x(t) = \mathrm{e}^{F(t)}\}$, unde F este o primitivă a funcţiei f.

Indicele inferior din notaţia $\mathcal{S}_o$ utilizată în Propoziţia precedentă indică faptul că ne referim la **soluţiile** ecuaţiei diferenţiale **omogene** $x' = f(t)x$ (adică membrul drept este liniar în x şi nu conţine termen liber, spre deosebire de următorul caz – ecuaţia diferenţială liniară **neomogenă**).

Exemplu 1.4. Să determinăm soluţia problemei Cauchy $x' = -2tx, x(0) = 1$. Avem $f(t) = -2t$, cu primitiva $F(t) = -t^2$, deci soluţia generală a ecuaţiei diferenţiale $x' = -2tx$ este $x(t) = C\mathsf{e}^{F(t)} = C\mathsf{e}^{-t^2}$. Din condiţia iniţială $x(0) = 1$ obţinem $C = 1$, deci soluţia problemei Cauchy va fi $x(t) = \mathsf{e}^{-t^2}$.



Această funcţie, având graficul reprezentat mai sus şi cunoscut sub denumirea de *clopotul lui Gauss*, are multiple aplicaţii în statistică (repatiţia normală), în teoria semnalelor (filtre), în procesarea imaginilor şi nu în ultimul rând în matematică, de exemplu în rezolvarea ecuaţiei căldurii şi în procese de difuzie.

3) **Ecuaţia diferenţială** $x' = f(t)x + g(t)$ $\quad (f, g : I \subseteq \mathbb{R} \to \mathbb{R}$ funcţii continue)

Dacă notăm ca mai devreme cu F o primitivă a funcţiei f, şi multiplicăm ambii membri ai ecuaţiei $x' = f(t)x + g(t)$ cu $\mathsf{e}^{-F(t)}$, obţinem

$$x'\mathsf{e}^{-F(t)} - f(t)x\mathsf{e}^{-F(t)} = g(t)$$

sau, echivalent,

$$(x\mathsf{e}^{-F(t)})' = g(t)\mathsf{e}^{-F(t)}$$

Rezultă că soluţia generală a ecuaţiei diferenţiale $x' = f(t)x + g(t)$ este

$$x(t) = \mathsf{e}^{F(t)}\left(K + \int g(t)\mathsf{e}^{-F(t)}\mathsf{d}t\right) \tag{21}$$

cu $K \in \mathbb{R}$.

În practică, pentru obţinerea soluţiei de mai sus se aplică *metoda variaţiei constantelor (Lagrange)*:

Pasul I. Se determină soluţia generală a ecuaţiei omogene $x' = f(t)x$, şi anume

$$x_o(t) = C\mathsf{e}^{F(t)}$$

Pasul II. Se caută soluţia ecuaţiei diferenţiale neomogene $x' = f(t)x + g(t)$ de forma

$$x(t) = C(t)\mathsf{e}^{F(t)}$$

Punând condiţia ca aceasta să verifice ecuaţia diferenţială, se obţine că funcţia $C = C(t)$ satisface $C'(t) = g(t)\mathsf{e}^{-F(t)}$, de unde, prin integrare, se ajunge la formula (21).

Mai departe, soluţia problemei Cauchy

$$\begin{cases} x' & = & f(t)x + g(t) \\ x(t_0) & = & x_0 \end{cases}$$

este

$$x(t) = C(t)\mathsf{e}^{\int_{t_0}^{t} f(s)\mathsf{d}s}, \text{ unde } C(t) = x_0 + \int_{t_0}^{t} g(\tau)\mathsf{e}^{-\int_{t_0}^{\tau} f(s)\mathsf{d}s}\mathsf{d}\tau$$

Exemplu 1.5. Să rezolvăm problema Cauchy $x' = \frac{2t}{1+t^2}x + 1$, $x(0) = -2$. Aplicăm metoda variaţiei constantelor, descrisă mai sus. Rezolvăm mai întâi ecuaţia omogenă $x' = \frac{2t}{1+t^2}x$:

$$x_o(t) = C\mathsf{e}^{\int \frac{2t}{1+t^2}\mathsf{d}t} = C\mathsf{e}^{\mathsf{ln}|1+t^2|} = C(1+t^2)$$

Căutăm acum soluţia generală a ecuaţiei neomogene de forma $x(t) = C(t)(1+t^2)$. Derivând şi înlocuind în ecuaţia diferenţială, obţinem:

$$C'(t)(1+t^2) + 2tC(t) = \frac{2t}{1+t^2}C(t)(1+t^2) + 1$$

de unde $C'(t) = \dfrac{1}{1+t^2}$, $C(t) = \mathsf{arctg}t + K$, unde $K \in \mathbb{R}$. Deci

$$x(t) = (1+t^2)(\mathsf{arctg}t + K)$$

Punând condiţia $x(0) = -2$, rezultă $K = -2$, deci $x(t) = (1+t^2)(\mathsf{arctg}t - 2)$ este soluţia căutată. □

În paralel cu cazul omogen, putem enunţa acum următorul rezultat:

Propoziţie 1.6. Fie $f, g : I \to \mathbb{R}$ funcţii continue, unde $I \subseteq \mathbb{R}$ este un interval. Dacă notăm

$$\mathcal{S}_o = \{x : I \to \mathbb{R} \mid x \text{ funcţie derivabilă şi } x' = f(t)x\}$$

spaţiul vectorial al soluţiilor ecuaţiei diferenţiale liniare omogene $x' = f(t)x$, şi cu

$$\mathcal{S} = \{x : I \to \mathbb{R} \mid x \text{ funcţie derivabilă şi } x' = f(t)x + g(t)\}$$

mulţimea soluţiilor ecuaţiei diferenţiale neomogene $x' = f(t)x + g(t)$, atunci:

(a) Pentru orice soluţii $x_1, x_2 \in \mathcal{S}$, $x_1 - x_2 \in \mathcal{S}_o$.

(b) Dacă $x_p \in \mathcal{S}$ este o soluţie (particulară) a ecuaţiei diferenţiale liniare neomogene, atunci

$$\mathcal{S} = \{x_o + x_p \mid x_o \in \mathcal{S}_o\}$$

Demonstraţie. Prin verificare directă. □

În concluzie, pentru rezolvarea ecuaţiei diferenţiale neomogene $x' = f(t)x + g(t)$ este necesar şi suficient să cunoaştem soluţia generală x_o a ecuaţiei diferenţiale liniare omogene $x' = f(t)x$, şi o soluţie *particulară* x_p a ecuaţiei neomogene; atunci $x = x_o + x_p$ este soluţia generală a ecuaţiei neomogene. În multe situaţii, o soluţie particulară se poate determina prin observarea termenului liber $g(t)$.

Exemplu 1.7. Să considerăm ecuaţia $x' = 2x+t^2$. Soluţia ecuaţiei omogene asociate $x' = 2x$ este $x_o(t) = C\mathrm{e}^{2t}$. Căutăm o soluţie particulară a ecuaţiei neomogene $x' = 2x + t^2$ de aceeaşi formă cu termenul liber t^2, adică un polinom de gradul II $x_p(t) = at^2+bt+c$. Punând condiţia ca acesta să verifice ecuaţia, obţinem $a = b = -\frac{1}{2}, c = -\frac{1}{4}$, deci $x_p(t) = -\frac{1}{2}(t^2 + t + \frac{1}{2})$ şi soluţia generală a ecuaţiei date se obţine prin superpoziţia celor două soluţii:

$$x(t) = x_o(t) + x_p(t) = C\mathrm{e}^{2t} - \frac{1}{2}(t^2 + t + \frac{1}{2}),\ C \in \mathbb{R}$$

Exerciţiu 1.8. Aplicaţi metoda variaţiei constantelor pentru a determina soluţia ecuaţiei diferenţiale din exemplul precendent şi verificaţi că rezultatul obţinut este acelaşi.

1.3 Aplicaţii

Creşterea populaţiei. Rata de creştere a unei populaţii este diferenţa dintre rata natalităţii şi rata mortalităţii. Dacă ne referim la populaţia umană la nivel mondial, o estimare actuală este de 1.05% pe an, adică $\lambda = 0.0105$/an.[34] Ca model elementar al evoluţiei populaţiei, putem presupune că rata de creştere este constantă în timp. Dacă $x(t)$ reprezintă numărul de indivizi la momentul t, atunci evoluţia în timp (în acest model elementar) este descrisă de ecuaţia diferenţială $x' = \lambda x$.

După cum am văzut mai sus, soluţia este de forma $x(t) = C\mathrm{e}^{\lambda t}$.

În condiţiile dezvoltării economice actuale şi a creşterii cererii pentru resursele naturale, este normal să ne punem întrebarea *În ce interval de timp populaţia globului se va dubla?*. Cu alte cuvinte, $x(t) = 2x(t_0)$, unde t_0 este momentul faţă de care ne raportăm (anul 2020, de exemplu). Aceasta revine la $\mathrm{e}^{\lambda(t-t_0)} = 2$, de unde obţinem $t - t_0 = \frac{\ln 2}{\lambda} \approx 66$ ani.

Mai mult chiar, să observăm că după 1000 ani vom avea $x(t_0+1000) = \mathrm{e}^{10.5}x(t_0)$ indivizi, comparativ cu $x(t_0)$ indivizi la momentul raportării. Dar $\mathrm{e}^{10.5} \approx 36315$, ceea ce ar conduce la o creştere enormă a populaţiei mondiale.

În realitate, situaţia diferă faţă de modelul elementar prezentat, în sensul că ecuaţia diferenţială ce descrie evoluţia populaţiei este $x' = \lambda x - \mu x^2$, termenul neliniar $-\mu x^2$ având rolul de a atenua creşterea exponenţială observată anterior.

Ecuaţia diferenţială $x' = \lambda x - \mu x^2$ se numeşte **ecuaţia logistică**.

Exerciţiu 1.9. În acest exerciţiu vom vedea cum putem determina o soluţie a ecuaţiei logistice $x' = \lambda x - \mu x^2$ menţionate în exemplul precedent ($\lambda > 0, \mu \neq 0$).

1) Arătaţi că soluţiile numerice ale ecuaţiei $\lambda x - \mu x^2 = 0$ induc soluţiile constante ale ecuaţiei logistice $x_1(t) = 0$ şi $x_2(t) = \dfrac{\lambda}{\mu}$, numite *soluţii staţionare*.

[34] `https://www.worldometers.info/world-population/`.

2) Observaţi că prin schimbarea de funcţie $y(t) = \dfrac{1}{x(t)}$ ecuaţia logistică se transformă într-o *ecuaţie diferenţială liniară de ordinul I*

$$y' = -\lambda y + \mu$$

3) Rezolvând această ecuaţie diferenţială liniară, arătaţi că soluţia ecuaţiei logistice este

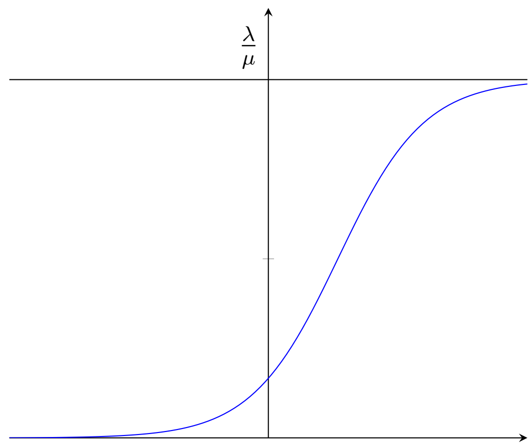
$$x(t) = \frac{\lambda}{\mu + K\mathsf{e}^{-\lambda t}} \tag{22}$$

unde $K \in \mathbb{R}$.

4) Observaţi următoarele:

(a) Limitele la $-\infty$, respectiv la ∞ ale soluţiei 22 corespund soluţiilor staţionare determinate anterior.

(b) $x(t)$ este o funcţie crescătoare, având graficul în forma literei S (reprezentat mai jos).



(c) Punctul de inflexiune corespunde momentului de timp în care creşterea este cea mai rapidă, după care rata de creştere scade spre 0. Determinaţi acest punct de inflexiune.

2 Sisteme de ecuaţii diferenţiale liniare

Forma generală a unui sistem de ecuaţii diferenţiale **liniare** cu coeficienţi constanţi este

$$\begin{cases} x_1' = a_{11}x_1 + \ldots + a_{1n}x_n + b_1(t) \\ \quad \vdots \\ x_n' = a_{n1}x_1 + \ldots + a_{nn}x_n + b_n(t) \end{cases} \tag{23}$$

unde:

- t este variabila independentă[35]

[35] Plecăm de la premiza că sistemul de ecuaţii diferenţiale de mai sus descrie evoluţia în **timp** a unor mărimi fizice măsurabile, de exemplu coordonatele unui punct material aflat în mişcare, sau parametrii presiune/temperatură/volum care descriu starea unui sistem termodinamic, etc.

- $x_1, \ldots, x_n$ sunt funcţiile necunoscute
- $a_{11}, \ldots, a_{nn} \in \mathbb{R}$ sunt constante reale
- $b_1, \ldots, b_n : I \to \mathbb{R}$ sunt funcţii continue pe un interval deschis $I \subseteq \mathbb{R}$.

Printr-o soluţie a acestui sistem vom înţelege n funcţii derivabile $x_1, \ldots, x_n : I \to \mathbb{R}$ care satisfac relaţiile $x_i'(t) = \sum_{j=1}^{n} a_{ij} x_j(t)$ pentru orice $t \in I$.

Pentru simplificare, vom utiliza notaţii matriceale: $A = (a_{ij})_{i,j=\overline{1,n}}$, $\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{b}(t) = \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix}$.

Atunci sistemul de ecuaţii diferenţiale (23) se poate scrie astfel:

$$\mathbf{X}' = A\mathbf{X} + \mathbf{b}(t)$$

Dacă $\mathbf{b}(t) = 0$ pentru orice $t \in I$, spunem că sistemul este **omogen**. În caz contrar, el se va numi **neomogen**.

2.1 Cazul omogen: $\mathbf{X}' = A\mathbf{X}$

Existenţa şi unicitatea soluţiei

Teoremă 2.1. Fie $A \in \mathcal{M}_n(\mathbb{R})$, $t_0 \in \mathbb{R}$ şi $\mathbf{x}_0 \in \mathbf{R}^n$. Atunci problema Cauchy

$$\begin{cases} \mathbf{X}' & = & A\mathbf{X} \\ \mathbf{X}(t_0) & = & \mathbf{x}_0 \end{cases}$$

are soluţie unică pe $\mathbb{R}$.

Demonstraţie. Rezultă din Teorema Cauchy-Lipschitz 1.2. □

Structura soluţiilor sistemului omogen de ecuaţii diferenţiale

Teoremă 2.2. Fie $\mathcal{S}_o = \{\mathbf{X} : \mathbb{R} \to \mathbb{R}^n \mid \mathbf{X}$ funcţie derivabilă şi $\mathbf{X}' = A\mathbf{X}\}$ mulţimea tuturor soluţiilor sistemului $\mathbf{X}' = A\mathbf{X}$. Atunci:

1) $\mathcal{S}_o$ este spaţiu vectorial real (**principiul superpoziţiei**).

2) Problema Cauchy $\begin{cases} \mathbf{X}' & = & A\mathbf{X} \\ \mathbf{X}(t_0) & = & \mathbf{0} \end{cases}$ are doar solutia banală $\mathbf{X}(t) = \mathbf{0}$.

3) Funcţiile $\mathbf{X}_1, \ldots, \mathbf{X}_p \in \mathcal{S}_o$ sunt liniar independente dacă şi numai dacă vectorii $\mathbf{X}_1(t_0)$, $\ldots$, $\mathbf{X}_p(t_0) \in \mathbb{R}^n$ sunt liniar independenţi.

4) $\dim_{\mathbb{R}} \mathcal{S}_o = n$.

Demonstraţie. 1) Prin verificare directă.

2) Rezultă din Teorema precedentă.

3) " $\Rightarrow$ " Fie $\mathbf{X}_1, \ldots, \mathbf{X}_p \in \mathcal{S}_o$ funcţii liniar independente. Fie $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ astfel încât $\alpha_1\mathbf{X}_1(t_0) + \ldots + \alpha_p\mathbf{X}_p(t_0) = \mathbf{0}$. Fie $\mathbf{Y} = \alpha_1\mathbf{X}_1 + \ldots + \alpha_p\mathbf{X}_p$. Atunci din punctul 1) rezultă $\mathbf{Y} \in \mathcal{S}_o$. Dar $\mathbf{Y}(t_0) = \mathbf{0}$. aplicând rezultatul de la punctul 2), obţinem $\mathbf{Y} = \mathbf{0}$. Deci $\alpha_1\mathbf{X}_1 + \ldots + \alpha_p\mathbf{X}_p = \mathbf{0}$, de unde, ţinând cont de liniar independenţa funcţiilor $\mathbf{X}_1, \ldots, \mathbf{X}_p$, rezultă $\alpha_1 = \ldots = \alpha_p = 0$.

   " $\Leftarrow$ " Fie acum $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ astfel încât $\alpha_1\mathbf{X}_1 + \ldots + \alpha_p\mathbf{X}_p = \mathbf{0}$. Aceasta înseamnă că pentru orice $t \in \mathbb{R}$, $\alpha_1\mathbf{X}_1(t) + \ldots + \alpha_p\mathbf{X}_p(t) = \mathbf{0}$. În particular, pentru $t = t_0$, rezultă $\alpha_1\mathbf{X}_1(t_0) + \ldots + \alpha_p\mathbf{X}_p(t_0) = \mathbf{0}$. Dar din ipoteză vectorii $\mathbf{X}_1(t_0)$, ..., $\mathbf{X}_p(t_0)$ sunt liniar independenţi, de unde obţinem $\alpha_1 = \ldots = \alpha_p = 0$.

4) Conform punctului 3), $\dim_{\mathbb{R}} \mathcal{S}_o \leq n$. Vom construi acum n soluţii liniar independente, de unde va rezulta că $\dim_{\mathbb{R}} \mathcal{S}_o = n$.

   Pentru fiecare $i \in \{1, \ldots, n\}$, să notăm cu $\mathbf{X}_i$ (unica) soluţie a următoarei probleme Cauchy:

$$\begin{cases} \mathbf{X}' &= A\mathbf{X} \\ \mathbf{X}(t_0) &= \mathbf{e}_i \end{cases}$$

   unde $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ sunt vectorii bazei canonice. Deoarece $\mathbf{X}_1(t_0), \ldots, \mathbf{X}_n(t_0)$ sunt vectori liniar independenţi, rezultă conform punctului anterior că şi funcţiile $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathcal{S}_o$ sunt liniar independente.

$\square$

Vom numi **sistem fundamental de soluţii** al sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$ o bază $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ a spaţiului soluţiilor $\mathcal{S}_o$. Atunci orice soluţie $\mathbf{X} \in \mathcal{S}_o$ a sistemului se scrie în mod unic sub forma

$$\mathbf{X}(t) = C_1\mathbf{X}_1(t) + \ldots C_n\mathbf{X}_n(t)\,,\ t \in \mathbb{R}$$

unde $C_1, \ldots C_n$ sunt constante reale.

Alternativ, putem scrie relaţia de mai sus matriceal

$$\mathbf{X}(t) = \big(\mathbf{X}_1(t) \mid \ \ldots \ \mid \mathbf{X}_n(t)\big) \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix} = \mathbf{M}(t)\mathbf{C}\,,\ \text{ unde } \mathbf{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}$$

iar matricea

$$\mathbf{M}(t) = \big(\mathbf{X}_1(t) \mid \ \ldots \ \mid \mathbf{X}_n(t)\big)$$

se numeşte **matrice fundamentală de soluţii** a sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$.

Observaţie 2.3. Ramâne totuşi o întrebare: cum determinăm un sistem fundamental de soluţii şi, implicit, soluţia generală a sistemului de ecuaţii diferenţiale?

Conform Teoremei 2.1, soluţia există şi este unică în prezenţa unei condiţii iniţiale, deci, independent de metoda aleasă, rezultatul va fi mereu acelaşi.

Putem proceda atunci astfel: întrucât sistemul de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$ generalizează cazul $n = 1$, şi anume ecuaţia diferenţială (19), este natural să căutăm soluţii ale sistemului de aceeaşi formă cu soluţiile ecuaţiei diferenţiale menţionate, şi anume $\mathbf{X}(t) = \mathsf{e}^{\lambda t}\mathbf{v}$, unde $\lambda \in \mathbb{R}$ şi $\mathbf{v} \in \mathbb{R}^n$.

Atunci $\mathbf{X}'(t) = \lambda \mathrm{e}^{\lambda t}\mathbf{v}$, şi punând condiţia ca $\mathbf{X}(t)$ să verifice sistemul $\mathbf{X}' = A\mathbf{X}$, rezultă $\lambda \mathrm{e}^{\lambda t}\mathbf{v} = A\mathrm{e}^{\lambda t}\mathbf{v}$, de unde obţinem $\lambda\mathbf{v} = A\mathbf{v}$. Deci λ este o valoare proprie a matricei A şi $\mathbf{v}$ un vector propriu asociat.

Rezultă că dacă putem determina n astfel de funcţii $\mathbf{X}_1(t) = \mathrm{e}^{\lambda_1 t}\mathbf{v}_1, \ldots, \mathbf{X}_n(t) = \mathrm{e}^{\lambda_n t}\mathbf{v}_n$ care să fie şi *liniar independente*, atunci am obţinut un sistem fundamental de soluţii. Dar conform Teoremei 2.2.3, este suficient să verificăm liniar independenţa vectorilor $\mathbf{X}_1(t_0), \ldots, \mathbf{X}_p(t_0)$. Luând $t_0 = 0$, obţinem vectorii (proprii) $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Dar matricea A admite n vectori proprii liniar independenţi dacă şi numai dacă este *diagonalizabilă*. Am obţinut astfel următorul rezultat:

Propoziţie 2.4. Fie $A \in \mathcal{M}_n(\mathbb{R})$ o matrice diagonalizabilă şi $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ vectori proprii liniar independenţi, corespunzători valorilor proprii $\lambda_1, \ldots, \lambda_n$ (nu neapărat distincte). Atunci funcţiile $\mathbf{X}_1(t) = \mathrm{e}^{\lambda_1 t}\mathbf{v}_1, \ldots, \mathbf{X}_n(t) = \mathrm{e}^{\lambda_n t}\mathbf{v}_n$ formează un sistem fundamental de soluţii ai sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$, iar soluţia generală este

$$\mathbf{X}(t) = C_1\mathrm{e}^{\lambda_1 t}\mathbf{v}_1 + \ldots + C_n\mathrm{e}^{\lambda_n t}\mathbf{v}_n$$

unde $C_1, \ldots, C_n \in \mathbb{R}$.

2.2 Cazul omogen: $\mathbf{X}' = A\mathbf{X}$ ($n = 2$)

Vom analiza în continuare cazul sistemelelor de ecuaţii diferenţiale planare (două ecuaţii şi două funcţii necunoscute, x şi y). Dacă $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$ şi $\mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}$, atunci sistemul $\mathbf{X}' = A\mathbf{X}$ devine

$$\begin{cases} x' &= ax + by \\ y' &= cx + dy \end{cases}$$

Cazul 0. Matricea sistemului este diagonală: $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. În acest caz, sistemul se scrie

$$\begin{cases} x' = \lambda_1 x \\ y' = \lambda_2 y \end{cases} \tag{24}$$

Se mai spune că ecuaţiile diferenţiale de mai sus sunt **decuplate**, deoarece putem rezolva separat fiecare dintre cele două ecuaţii diferenţiale, obţinând astfel soluţia generală

$$\begin{cases} x(t) = C_1\mathrm{e}^{\lambda_1 t} \\ y(t) = C_2\mathrm{e}^{\lambda_2 t} \end{cases}$$

unde $C_1, C_2 \in \mathbb{R}$ sunt constante.[36]

Matricial, putem scrie soluţia generală sub forma

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} C_1\mathrm{e}^{\lambda_1 t} \\ C_2\mathrm{e}^{\lambda_2 t} \end{pmatrix} = \underbrace{C_1\mathrm{e}^{\lambda_1 t}\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{\mathbf{X}_1(t)} + \underbrace{C_2\mathrm{e}^{\lambda_2 t}\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{\mathbf{X}_2(t)} = \underbrace{\begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix}}_{\mathbf{M}(t)}\underbrace{\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}}_{\mathbf{C}}$$

[36] Deşi ecuaţiile diferenţiale sunt decuplate, soluţiile acestora nu sunt, întrucât depind de aceeaşi variabilă t.

În particular, sistemul fundamental de soluţii este $\{\mathbf{X}_1(t) = \mathsf{e}^{\lambda_1 t}\begin{pmatrix}1\\0\end{pmatrix}, \mathbf{X}_2(t) = \mathsf{e}^{\lambda_2 t}\begin{pmatrix}0\\1\end{pmatrix}\}$.

Exemplu 2.5. Soluţia generală a sistemului de ecuaţii diferenţiale liniare

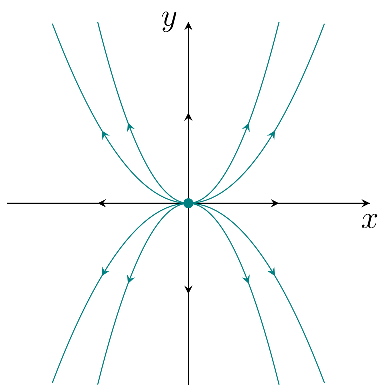
$$\begin{cases} x' = x \\ y' = 2y \end{cases}$$

conform observaţiilor de mai sus, este

$$\begin{cases} x(t) = C_1\mathsf{e}^{t} \\ y(t) = C_2\mathsf{e}^{2t} \end{cases}$$

unde $C_1, C_2 \in \mathbb{R}$.

Pentru a înţelege acest rezultat, să revenim la interpretarea soluţiilor ca evoluţii în timp (traiectorii) $(x(t), y(t))$ ale unui sistem dinamic. Pentru fiecare pereche de constante (C_1, C_2) se obţine o traiectorie. Cazul $C_1 = C_2 = 0$ produce soluţia staţionară $x(t) = y(t) = 0$, numită **soluţie de echilibru**. Apoi $C_1 = 0$ şi $C_2 \neq 0$ conduc la o traiectorie rectilinie verticală ($x(t) = 0, y(t) = C_2\mathsf{e}^{2t}$) care se îndepărtează de punctul de echilibru $(0,0)$, în sensul că $\lim_{t\to\infty} y(t) = \infty$ pentru $C_2 > 0$, respectiv $\lim_{t\to\infty} y(t) = -\infty$ pentru $C_2 < 0$. În sfârşit, pentru $C_1 \neq 0$ şi $C_2 \neq 0$ se obţine o familie de parabole $y = \frac{C_2}{C_1^2}x^2$, deplasarea pe fiecare dintre acestea efectuându-se prin îndepărtare de origine, conform graficului de mai jos. Această reprezentare a evoluţiei (în timp) a soluţiilor sistemului de ecuaţii diferenţiale se numeşte **portret de faze**:



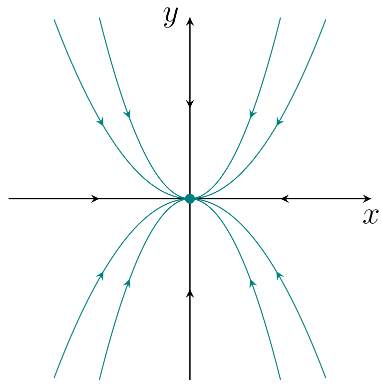
Coroborând cazul general al sistemului de ecuaţii diferenţiale decuplate (24) cu exemplul descris mai sus, observăm că putem întâlni următoarele situaţii:

1) $0 < \lambda_1 < \lambda_2$. Aceasta este situaţia descrisă în exemplul de mai sus, în care soluţia de echilibru $(0,0)$ este un **nod instabil**.[37]

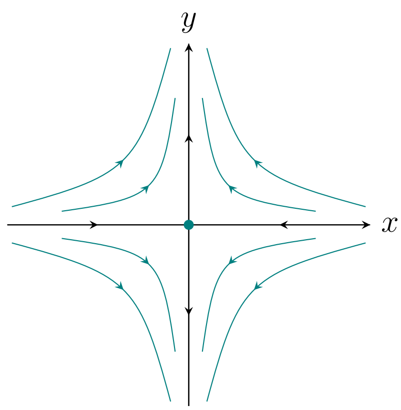
2) $\lambda_1 < \lambda_2 < 0$. Ceea ce se modifică faţă de cazul precedent este sensul de deplasare, de data aceasta către origine, pe fiecare dintre traiectorii. Soluţia de echilibru $(0,0)$ se numeşte de

[37] *Source*, în limba engleză.

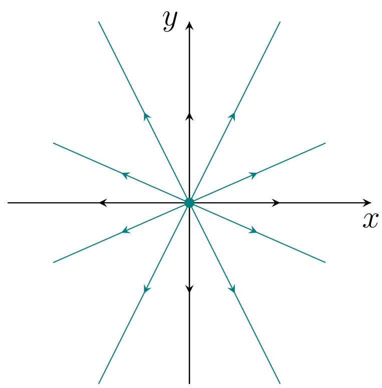
data aceasta **nod stabil atractor**.[38]



3) $\lambda_1 < 0 < \lambda_2$. Portretul de faze este reprezentat mai jos, iar originea (punctul de echilibru) este un **punct instabil de tip şa**. De exemplu, sistemul $x' = -x, y' = 2y$ conduce la traiectoriile curbilinii de tipul $x^2y =$ constant:



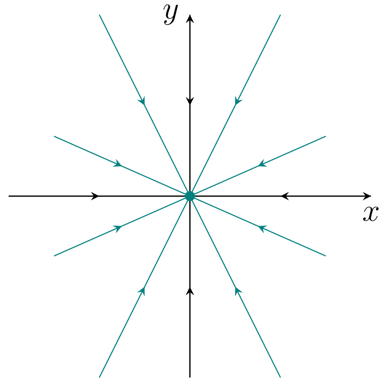
4) $\lambda_1 = \lambda_2 > 0$. În acest caz, soluţia generală este $x(t) = C_1\mathsf{e}^{\lambda t}, y(t) = C_2\mathsf{e}^{\lambda t}$, iar traiectoriile sunt rectilinii ($C_1y = C_2x$), originea este **nod instabil**, iar portretul de faze este următorul:



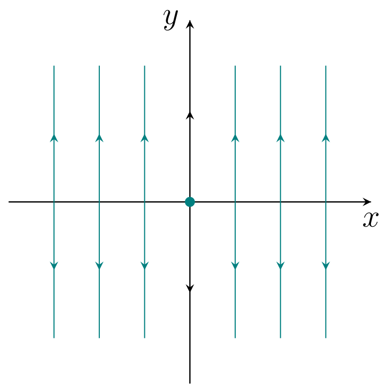
5) $\lambda_1 = \lambda_2 < 0$. Ceea ce se schimbă este direcţia de deplasare pe fiecare dintre traiectorii (de asemenea rectilinii), către punctul de echilibru (originea), care devine astfel un **nod stabil**

[38] *Sink*, în limba engleză.

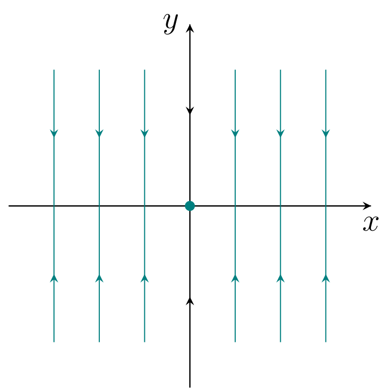
atractor:



6) $\lambda_1 = 0$, $\lambda_2 > 0$. Soluţiile sunt $x(t) = C_1, y(t) = C_2\mathrm{e}^{\lambda_2 t}$ şi determină o familie de drepte paralele cu axa Oy; fiecare punct de pe axa Ox este o soluţie de echilibru, de tip **nod instabil** (ne îndepărtăm în timp de acestea prin deplasare paralel cu Oy):



7) $\lambda_1 = 0$, $\lambda_2 < 0$: analog, fiecare punct de pe Ox este un **nod stabil atractor**, iar traiectoriile sunt rectilinii şi paralele cu Oy:



8) $\lambda_1 = \lambda_2 = 0$. Acest caz degenerat produce doar soluţii constante $x(t) = C_1, y(t) = C_2$, deci staţionare în timp (orice punct din plan este un punct de echilibru stabil).

Cazul I. Matricea A este diagonalizabilă $A = PDP^{-1}$, unde $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ este matricea diagonală a valorilor proprii şi $P = \big(\mathbf{v}_1 \mid \mathbf{v}_2\big)$ este matricea de trecere de la baza canonică la baza formată cu vectorii proprii. Putem obţine soluţia generală a sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$ utilizând Propoziţia 2.4. Pentru a putea interpreta însă soluţia obţinută şi a putea realiza portretul de faze, vom rezolva direct sistemul, utilizând diagonalizarea matricei A, astfel:

$$\mathbf{X}' = PDP^{-1}\mathbf{X}$$

devine

$$P^{-1}\mathbf{X}' = DP^{-1}\mathbf{X}$$

Dacă notăm $\begin{pmatrix} u \\ v \end{pmatrix} = P^{-1}\mathbf{X}$, se obţine sistemul cu ecuaţiile diferenţiale decuplate

$$\begin{cases} u' = \lambda_1 u \\ v' = \lambda_2 v \end{cases}$$

pentru care am văzut mai sus (**Cazul 0.**) că soluţia este

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} C_1 \mathrm{e}^{\lambda_1 t} \\ C_2 \mathrm{e}^{\lambda_2 t} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

Deci soluţia generală a sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$ va fi

$$\mathbf{X}(t) \quad = \quad P \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \quad = \quad C_1 \underbrace{\mathrm{e}^{\lambda_1 t}\mathbf{v}_1}_{\mathbf{X}_1(t)} + C_2 \underbrace{\mathrm{e}^{\lambda_2 t}\mathbf{v}_2}_{\mathbf{X}_2(t)}$$

şi $\{\mathbf{X}_1(t) = \mathrm{e}^{\lambda_1 t}\mathbf{v}_1, \mathbf{X}_2(t) = \mathrm{e}^{\lambda_2 t}\mathbf{v}_2\}$ este un sistem fundamental de soluţii, în concordanţă cu Propoziţia 2.4.

Trecerea de la x şi y la u, respectiv v, s-a realizat prin schimbarea bazei în $\mathbb{R}^2$: de la baza canonică la baza formată cu vectorii proprii ai matricei A. Vom ţine cont de această observaţie în realizarea portretelor de faze pentru exemplele din secţiunea 2.3.

Cazul II. Matricea A are o valoare proprie dublă şi nu se diagonalizează. În acest caz, Propoziţia 2.4 nu mai furnizează un sistem fundamental de soluţii (nu mai putem determina o bază formată cu vectori proprii). Cu toate acestea, reamintim din Secţiunea 6.3 că matricea se poate descompune sub forma

$$A = PJP^{-1}, \quad \text{unde } J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \text{ şi } P = \big(\mathbf{v}_1 \mid \mathbf{v}_2\big) \tag{25}$$

Ca în cazul precedent, sistemul $\mathbf{X}' = A\mathbf{X}$ devine

$$\mathbf{X}' = PDP^{-1}\mathbf{X} \iff P^{-1}\mathbf{X}' = DP^{-1}\mathbf{X}$$

Dacă notăm $\begin{pmatrix} u \\ v \end{pmatrix} = P^{-1}\mathbf{X}$, se obţine

$$\begin{cases} u' = \lambda u + v \\ v' = \qquad \lambda v \end{cases}$$

a cărui soluţie se determină în două etape:

1) Mai întâi, rezolvăm ecuaţia independentă $v' = \lambda v$, cu soluţia generală $v(t) = C_2 \mathrm{e}^{\lambda t}$.

2) Apoi, să observăm că prima ecuaţie a sistemului devine $u' = \lambda u + C_2 \mathrm{e}^{\lambda t}$, o ecuaţie diferenţială liniară, cu soluţia $u(t) = (C_1 + C_2 t)\mathrm{e}^{\lambda t}$.

În concluzie, soluţia noului sistem este

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} (C_1 + C_2 t)\mathrm{e}^{\lambda t} \\ C_2 \mathrm{e}^{\lambda t} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\lambda t} & t\mathrm{e}^{\lambda t} \\ 0 & \mathrm{e}^{\lambda t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

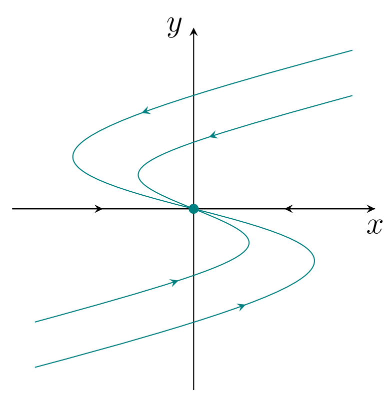
iar soluţia generală a sistemului $\mathbf{X}' = A\mathbf{X}$ va fi

$$\mathbf{X}(t) = P \begin{pmatrix} \mathrm{e}^{\lambda t} & t\mathrm{e}^{\lambda t} \\ 0 & \mathrm{e}^{\lambda t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = C_1 \underbrace{\mathrm{e}^{\lambda t}\mathbf{v}_1}_{\mathbf{X}_1(t)} + C_2 \underbrace{\mathrm{e}^{\lambda t}(t\mathbf{v}_1 + \mathbf{v}_2)}_{\mathbf{X}_2(t)}$$

Exemplu 2.6. 1) Să rezolvăm problema Cauchy

$$\begin{cases} x' = -x + y & x(0) = 1 \\ y' = \quad\;\; - y & y(0) = 2 \end{cases}$$

În acest exemplu, matricea sistemului este deja de forma $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, deci putem rezolva mai întâi a doua ecuaţie independentă, urmând apoi să utilizăm soluţia obţinută pentru a determina şi cealaltă funcţie necunoscută. Aşadar, ecuaţia diferenţială liniară $y' = -y$ conduce la soluţia generală $y(t) = C_2 \mathrm{e}^{-t}$, de unde obţinem $x' = -x + C_2 \mathrm{e}^{-t}$, cu soluţia $x(t) = \mathrm{e}^{-t}(C_1 + C_2 t)$. Înainte de a utiliza condiţiile iniţiale, să analizăm şi comportamentul soluţiilor. Avem $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} y(t) = 0$, deci soluţiile tind asimptotic către soluţia de echilibru $(0, 0)$. Reprezentarea grafică a fiecărei soluţii (fiecărei traiectorii), mai puţin intuitivă decât în cazurile precedente, are forma literei "S":



Revenind la problema Cauchy, condiţiile iniţiale $x(0) = 1, y(0) = 2$ conduc la $C_1 = 1, C_2 = 2$, deci soluţia va fi $x(t) = (1 + 2t)\mathrm{e}^{-t}, y(t) = 2\mathrm{e}^{-t}$.

2) Să rezolvăm acum problema Cauchy

$$\begin{cases} x' = -1.4x + 0.8y & x(0) = 0 \\ y' = -0.2x - 0.6y & y(0) = -1 \end{cases}$$

Matricea sistemului $A = \begin{pmatrix} -1.4 & 0.8 \\ -0.2 & -0.6 \end{pmatrix}$ admite valoarea proprie dublă $\lambda = -1$, cu $V_\lambda = \mathrm{sp}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$, deci nu este diagonalizabilă. Vom proceda atunci ca în Secţiunea 6.3 pentru a obţine o descompunere a matricei de forma (25). Fie $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.[39] Atunci $\mathbf{v}_1 = (A - \lambda I_2)\mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in V_\lambda$ şi $\mathbf{v}_1, \mathbf{v}_2$ sunt vectorii căutaţi. Putem scrie $A = PJP^{-1}$, unde

$$P = (\mathbf{v}_1 \mid \mathbf{v}_2) = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \quad \text{şi} \quad J = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

Efectuăm schimbarea $\begin{pmatrix} u \\ v \end{pmatrix} = P^{-1} \begin{pmatrix} x \\ y \end{pmatrix}$. Atunci u şi v verifică sistemul

$$\begin{cases} u' = -u + v \\ v' = - v \end{cases}$$

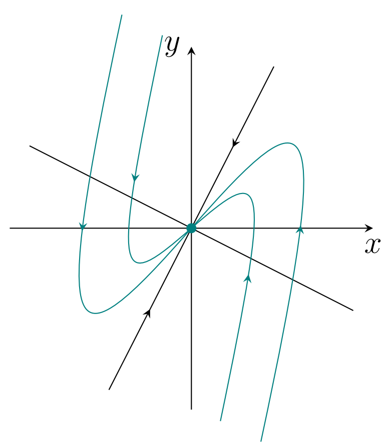
Am determinat soluţia acestui sistem în exemplul precedent, şi anume

$$\begin{cases} u(t) = (C_1 + C_2 t)\mathrm{e}^{-t} \\ v(t) = C_2 \mathrm{e}^{-t} \end{cases}$$

Rezultă

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = P \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} (C_1 + C_2 t)\mathrm{e}^{-t} \\ C_2 \mathrm{e}^{-t} \end{pmatrix} = \mathrm{e}^{-t} \begin{pmatrix} (2C_1 - C_2) + C_2 t \\ (C_1 + 2C_2) + C_2 t \end{pmatrix}$$

cu portretul de faze[40]



Cazul III. Matricea A are valori proprii complexe conjugate $\lambda, \overline{\lambda} \in \mathbb{C} \setminus \mathbb{R}$. Fie $\mathbf{v} \in \mathbb{C}^2$ un vector propriu corespunzător valorii proprii λ. Atunci $\overline{\mathbf{v}}$ este vector propriu corespunzător valorii proprii $\overline{\lambda}$, matricea A este diagonalizabilă (peste corpul numerelor complexe), şi putem aplica Propoziţia 2.4: funcţiile (complexe) vectoriale $\mathbf{Z}(t) = \mathrm{e}^{\lambda t}\mathbf{v}$ şi $\overline{\mathbf{Z}(t)} = \mathrm{e}^{\overline{\lambda} t}\overline{\mathbf{v}}$ formează un sistem fundamental de soluţii.

[39] Conform algoritmului descris în Secţiunea 6.3, putem alege *orice* vector $\mathbf{v}_2 \in \mathbb{R}^2$ cu proprietatea că $\mathbf{v}_2 \notin V_\lambda$. Vectorul $\mathbf{v}_2$ ales mai sus are proprietatea adiţională de a fi ortogonal pe V_λ, deci vom transforma baza canonică nu într-o bază arbitrară $\{\mathbf{v}_1, \mathbf{v}_2\}$, ci într-o bază ortogonală (axele rămân perpendiculare).

[40] Obţinut prin scalare cu $\sqrt{5}$ şi apoi rotaţie cu unghiul $\mathrm{arctg} 2$ în sens trigonometric în jurul originii din portretul anterior de faze. Aceste transformări geometrice corespund trecerii de la baza canonică la baza ortogonală $\{\mathbf{v}_1, \mathbf{v}_2\}$ – observaţi că ambii vectori au norma $\sqrt{5}$, de unde factorul de scalare menţionat anterior. Puteţi preciza cum s-a obţinut unghiul de rotaţie?

Dar întrucât suntem interesaţi să determinăm soluţiile **reale** ale sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$, să observăm că

$$\mathbf{X}_1(t) = \mathsf{Re}(\mathsf{e}^{\lambda t}\mathbf{v}) \qquad \mathbf{X}_2(t) = \mathsf{Im}(\mathsf{e}^{\lambda t}\mathbf{v})$$

sunt soluţii reale ale sistemului, deoarece A este o matrice reală, $A \in \mathcal{M}_2(\mathbb{R})$.

În plus, vectorii $\mathbf{X}_1(0) = \mathsf{Re}(\mathbf{v}), \mathbf{X}_2(0) = \mathsf{Im}(\mathbf{v})$ sunt liniar independenţi (Exerciţiul 6.23). Conform Teoremei 2.2.3, rezultă că funcţiile $\mathbf{X}_1(t), \mathbf{X}_2(t)$ sunt liniar independente, deci formează un sistem fundamental de soluţii reale ale sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$, iar soluţia generală a sistemului va fi

$$\mathbf{X}(t) = C_1\mathbf{X}_1(t) + C_2\mathbf{X}_2(t)$$

unde $C_1, C_2 \in \mathbb{R}$.

Dacă notăm $\lambda = \alpha + \mathsf{i}\beta$ şi $\mathbf{v} = \mathbf{v}_1 + \mathsf{i}\mathbf{v}_2$, atunci din formula lui Euler rezultă $\mathbf{Z}(t) = \mathsf{e}^{\alpha t}(\cos\beta t + \mathsf{i}\sin\beta t)(\mathbf{v}_1 + \mathsf{i}\mathbf{v}_2)$, deci $\mathbf{X}_1(t) = \mathsf{Re}(\mathbf{Z}(t)) = \mathsf{e}^{\alpha t}(\cos\beta t\,\mathbf{v}_1 - \sin\beta t\,\mathbf{v}_2)$ şi $\mathbf{X}_2(t) = \mathsf{Im}(\mathbf{Z}(t)) = \mathsf{e}^{\alpha t}(\sin\beta t\,\mathbf{v}_1 + \cos\beta t\,\mathbf{v}_2)$. Putem scrie atunci soluţia generală a sistemului de ecuaţii diferenţiale sub o formă matriceală, mai uşor de reţinut, astfel:

$$\begin{aligned} \mathbf{X}(t) &= C_1\mathbf{X}_1(t) + C_2\mathbf{X}_2(t) \\ &= C_1\mathsf{e}^{\alpha t}(\cos\beta t\,\mathbf{v}_1 - \sin\beta t\,\mathbf{v}_2) + C_2\mathsf{e}^{\alpha t}(\sin\beta t\,\mathbf{v}_1 + \cos\beta t\,\mathbf{v}_2) \\ &= \mathsf{e}^{\alpha t}\left(\mathbf{v}_1 \mid \mathbf{v}_2\right)\begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \end{aligned} \tag{26}$$

Cum arată portretul de faze în această situaţie? Să considerăm mai întâi cazul $(\mathbf{v}_1 \mid \mathbf{v}_2) = I_2$[41] Deci

$$\mathbf{X}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \mathsf{e}^{\alpha t}\begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

În loc să scriem explicit cele două componente $x(t)$ şi $y(t)$, să observăm că

$$\|\mathbf{X}(t)\| = \mathsf{e}^{\alpha t}\left\|\begin{pmatrix} C_1 \\ C_2 \end{pmatrix}\right\|$$

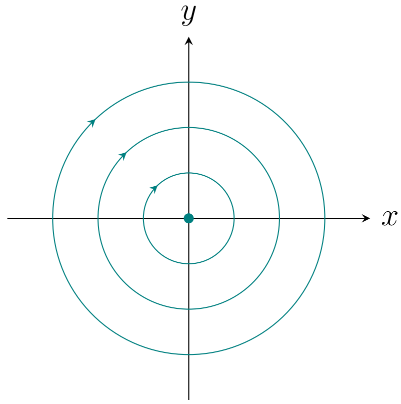
deoarece matricea $\begin{pmatrix} \cos\beta t & \sin\beta t \\ -\sin\beta t & \cos\beta t \end{pmatrix}$ este ortogonală. În particular, aceasta revine la

$$x^2(t) + y^2(t) = \mathsf{e}^{2\alpha t}(C_1^2 + C_2^2) \quad , \quad \forall\, t$$

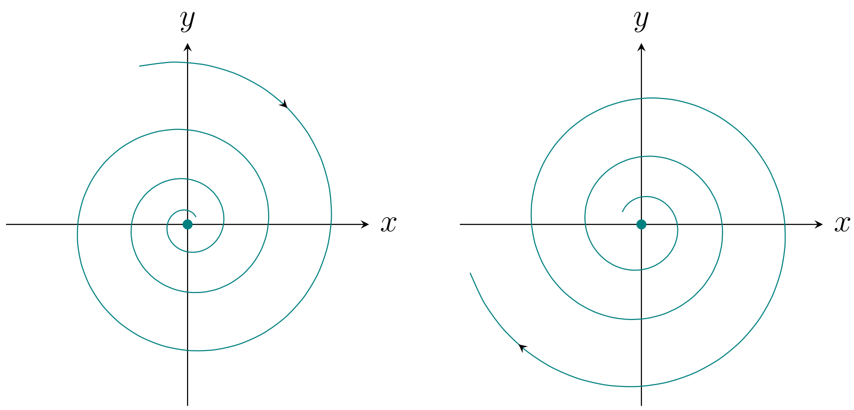
Distingem două situaţii:

[41] Atunci matricea sistemului este $A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, conform Secţiunii 6.3.

1) $\alpha = 0$. Atunci traiectoriile sunt circulare: $x^2(t) + y^2(t) = C_1^2 + C_2^2$ (periodice în timp):



2) $\alpha \neq 0$. Atunci $\lim_{t\to\infty}(x^2(t) + y^2(t)) = \begin{cases} 0 & \alpha < 0 \\ \infty & \alpha > 0 \end{cases}$. Se obţin ca traiectorii *spirale logaritmice* pentru care originea $(0,0)$ este un **focar stabil** ($\alpha < 0$ şi traiectoriile se apropie de origine), respectiv **focar instabil** (pentru $\alpha > 0$):



Revenim acum la cazul general, în care $(\mathbf{v}_1 \mid \mathbf{v}_2)$ nu este neapărat matricea identică. Traiectoriile vor fi atunci cercuri/spirale logaritmice deformate, corespunzător trecerii de la baza canonică la baza $\{\mathbf{v}_1, \mathbf{v}_2\}$.

2.3 Aplicaţii şi exemple ale sistemelor de ecuaţii diferenţiale liniare

Exemplu 2.7. Acesta este un model inspirat din biologie: un ecosistem în care trăiesc două specii aflate în competiţie pentru hrană. Fie $x(t)$ şi $y(t)$ numărul de exemplare din cele două specii la momentul de timp t, exprimat în ani. În absenţa celei de-a doua specii, rata de creştere a primei specii este $x' = 1.6x$, dar dacă ambele specii sunt prezente şi în competiţie pentru resurse, rata de creştere se modifică astfel: $x' = 1.6x - 0.3y$. Analog se întâmplă şi cu a doua specie, pentru care rata de creştere este $y' = -0.8x + 1.4y$. Dacă iniţial sunt câte 10 exemplare din fiecare specie, să se determine evoluţia în timp a celor două specii.

Avem deci de rezolvat problema Cauchy

$$\begin{cases} x' = 1.6x - 0.3y & x(0) = 10 \\ y' = -0.8x + 1.4y & y(0) = 10 \end{cases}$$

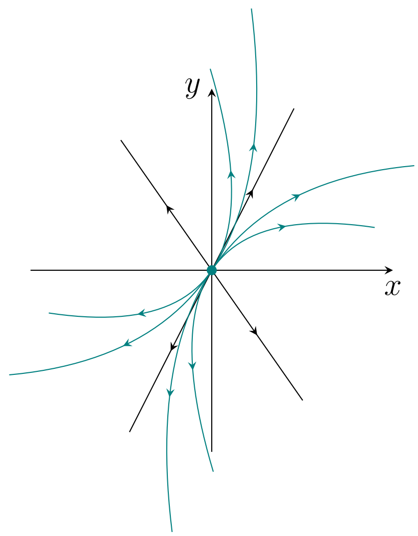
Matricea sistemului $A = \begin{pmatrix} 1.6 & -0.3 \\ -0.8 & 1.4 \end{pmatrix}$ admite valorile proprii simple $\lambda_1 = 1$, $\lambda_2 = 2$, deci este diagonalizabilă, iar $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -3 \\ 4 \end{pmatrix}$ sunt vectori proprii corespunzători celor două valori proprii. În aceste condiţii, soluţia generală a sistemului va fi

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = C_1 \mathsf{e}^{\lambda_1 t}\mathbf{v}_1 + C_2 \mathsf{e}^{\lambda_2 t}\mathbf{v}_2 = \begin{pmatrix} C_1\mathsf{e}^t - 3C_2\mathsf{e}^{2t} \\ 2C_1\mathsf{e}^t + 4C_2\mathsf{e}^{2t} \end{pmatrix}$$

Cum putem realiza portretul de faze? Scriind soluţia generală sub formă matriceală

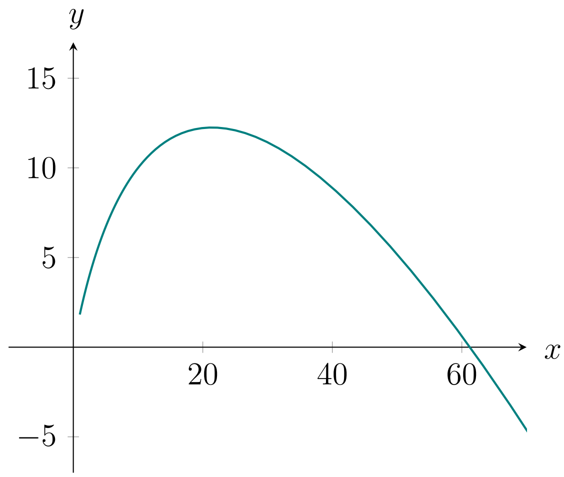
$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} \mathsf{e}^t & 0 \\ 0 & \mathsf{e}^{2t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

să observăm că $\begin{pmatrix} \mathsf{e}^t & 0 \\ 0 & \mathsf{e}^{2t} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$ este soluţia generală a sistemului de ecuaţii diferenţiale decuplate din Exemplul 2.5, iar matricea $P = \begin{pmatrix} 1 & -3 \\ 2 & 4 \end{pmatrix}$ este matricea de trecere de la baza canonică la baza formată cu cei doi vectori proprii ai matricei A. Portretul de faze, reprezentat mai jos, se obţine deci prin "deformarea" portretului de faze ilustrat în Exemplul 2.5



unde "deformarea" corespunde schimbării bazei $\left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \longrightarrow \left\{ \mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -3 \\ 4 \end{pmatrix} \right\}$ (pentru o mai bună intuiţie geometrică, a se vedea şi exerciţiul următor).

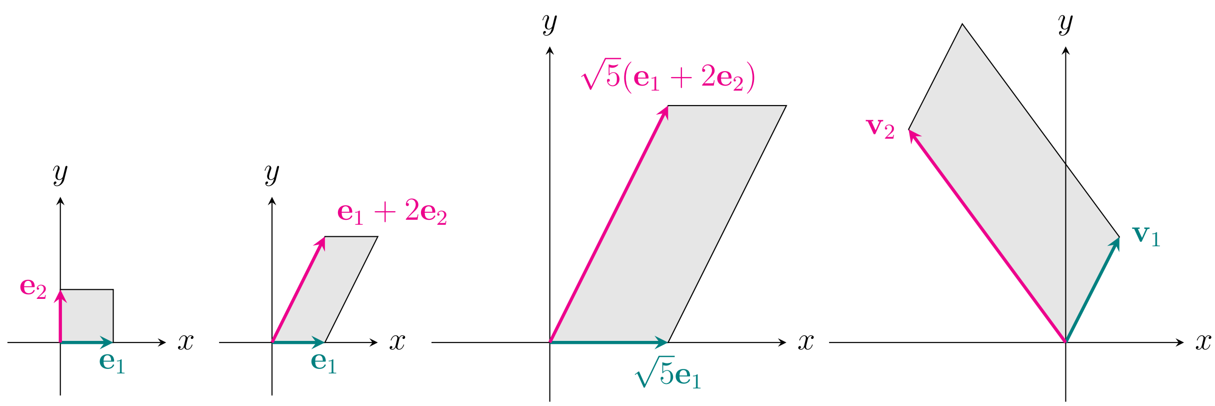
Dar să revenim la modelul biologic al problemei: condiţiile iniţiale $x(0) = y(0) = 10$ conduc la $C_1 = 7, C_2 = -1$, deci la soluţia $x(t) = 7\mathsf{e}^t + 3\mathsf{e}^{2t}$, $y(t) = 14\mathsf{e}^t - 4\mathsf{e}^{2t}$ reprezentată în graficul de mai jos. În particular, să observăm că a doua specie se va stinge după mai puţin de 2 ani (puteţi preciza momentul exact?).



Exerciţiu 2.8. Arătaţi cu ajutorul descompunerii QR că matricea de trecere $P = \begin{pmatrix} 1 & -3 \\ 2 & 4 \end{pmatrix}$ de la baza canonică la baza formată cu vectorii proprii din Exemplul precedent, se poate scrie astfel:

$$P = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \sqrt{5} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

Cu alte cuvinte, efectul schimbării bazei (coordonatelor) corespunde compunerii următoarelor transformări geometrice: o forfecare, urmată de o scalare (cu factorul $\sqrt{5}$) şi de o rotaţie în sens trigonometric cu un unghi $\mathsf{arctg}2 \approx 63°$. Reprezentarea grafică de mai jos ilustrează pas cu pas aceste transformări geometrice aplicate vectorilor din baza canonică (şi pătratului determinat de aceştia):



Exemplu 2.9 (Modelul pradă-prădător (Lotka-Volterra))**.** Vom prezenta un alt model originar din biologie, şi anume interacţiunea dintre două specii ce convieţuiesc în acelaşi mediu, de tipul **pradă-prădător**. Una dintre specii (**predator**) se hrăneşte cu indivizi din a doua specie (**prada**), iar aceştia din urmă îşi procură hrana din mediul înconjurător.

Modelul matematic a fost elaborat în anii '20 ai secolului precedent de matematicianul italian Vito Volterra, într-o analiză a populaţiei de rechini şi a celorlalte specii piscicole cu care aceştia se hrăneau din Marea Adriatică, plecând de la un sistem de ecuaţii diferenţiale obţinut anterior de Alfred Lotka în teoria reacţiilor chimice.

Vom prezenta o variantă simplificată: pentru început, să notăm cu $x(t)$ numărul de exemplare din specia pradă la momentul de timp t, şi cu $y(t)$ numărul prădătorilor la acelaşi moment de timp. Vom face următoarele presupuneri:

1) În absenţa prădătorilor, specia pradă va creşte în mod natural, conform ecuaţiei diferenţiale $x' = ax$, unde $a > 0$.

2) În absenţa prăzii, numărul de exemplare din specia prădător va scădea: $y' = -by$, unde $b > 0$.

3) Când ambele specii sunt prezente, este natural să ne aşteptăm la un declin al numărului de exemplare din specia pradă, respectiv la o creştere a numărului de exemplare din specia prădător, proporţional cu numărul întâlnirilor dintre indivizii celor două specii. Vom presupune că frecvenţa acestor întâlniri este proporţională cu produsul $x(t)y(t)$. În consecinţă, consumarea prăzii de către predatori va conduce la o rată de declin $-pxy$ în populaţia pradă, respectiv la o rată de creştere qxy în populaţia de tip prădător.

Rezultă următorul sistem de ecuaţii diferenţiale care modelează un ecosistem de tip pradă-prădător:

$$\begin{cases} x' &= ax - pxy \\ y' &= -by + qxy \end{cases} \tag{27}$$

unde $a, b, p, q > 0$.

Ca o primă observaţie, acesta nu este un sistem **liniar** de ecuaţii diferenţiale. Pentru a putea aplica tehnicile dezvoltate în acest capitol dedicat ecuaţiilor diferenţiale, vom proceda astfel:

1) Determinăm punctele de echilibru ale sistemului. Acestea sunt funcţiile staţionare (constante în timp) care verifică $ax - pxy = 0, -by + qxy = 0$. Obţinem soluţiile $(0, 0)$[42] şi $(\frac{b}{q}, \frac{a}{p})$.[43]

2) Liniarizăm sistemul în jurul punctelor de echilibru. Vom considera doar punctul $(\frac{b}{q}, \frac{a}{p})$, celălalt punct de echilibru rămânând ca exerciţiu.

   Prin liniarizare înţelegem aproximare cu polinomul Taylor de ordinul I. Pentru a fixa notaţiile, fie $F(x, y) = ax - pxy$ şi $G(x, y) = -by + qxy$. Atunci polinoamele Taylor de ordinul I (aproximările liniare) în jurul punctului $(x_0, y_0) = (\frac{b}{q}, \frac{a}{p})$ asociate funcţiilor F, respectiv G, vor fi

$$\begin{aligned} F(x,y) \approx&\ F(x_0,y_0) + \frac{1}{1!}\left(\frac{\partial F}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial F}{\partial y}(x_0,y_0)(y-y_0)\right) \\ =&\ (a - py_0)(x - x_0) + (-px_0)(y - y_0) \\ =&\ -\frac{bp}{q}(y - y_0) \\ G(x,y) \approx&\ G(x_0,y_0) + \frac{1}{1!}\left(\frac{\partial G}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial G}{\partial y}(x_0,y_0)(y-y_0)\right) \\ =&\ \frac{aq}{p}(x - x_0) \end{aligned}$$

3) Aproximăm sistemul neliniar iniţial (27) cu sistemul obţinut prin liniarizare

$$\begin{cases} x' &= -\frac{bp}{q}(y - y_0) \\ y' &= \frac{aq}{p}(x - x_0) \end{cases} \tag{28}$$

[42] Extincţia celor două populaţii.

[43] Aceasta din urmă reprezintă singura variantă în care numărul de indivizi din specia pradă şi respectiv prădător care pot *coexista* permanent, sunt constante în timp.

4) Deşi am obţinut sistemul liniar (28), acesta nu este omogen. Îl putem transforma într-un sistem omogen printr-o translaţie: $u = x - x_0, v = y - y_0$. Atunci $u' = x', v' = y'$ şi (28) devine

$$\begin{cases} u' &= -\frac{bp}{q}v \\ v' &= \frac{aq}{p}x \end{cases} \tag{29}$$

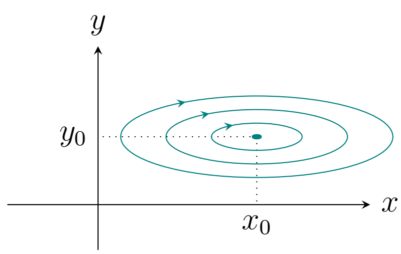
Am obţinut în final un sistem omogen de ecuaţii liniare diferenţiale, cu matricea sistemului $A = \begin{pmatrix} 0 & -\frac{bp}{q} \\ \frac{aq}{p} & 0 \end{pmatrix}$. Valorile proprii sunt pur imaginare $\pm\mathrm{i}\sqrt{ab}$. Pentru $\lambda = \mathrm{i}\sqrt{ab}$ obţinem vectorul propriu $\mathbf{v} = \begin{pmatrix} p\sqrt{b} \\ -\mathrm{i}q\sqrt{a} \end{pmatrix}$, deci cu notaţiile din (26) avem $\mathbf{v}_1 = \begin{pmatrix} p\sqrt{b} \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ -q\sqrt{a} \end{pmatrix}$ şi soluţia sistemului de ecuaţii diferenţiale (29) va fi

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} p\sqrt{b} & 0 \\ 0 & -q\sqrt{a} \end{pmatrix} \begin{pmatrix} \cos\sqrt{ab}t & \sin\sqrt{ab}t \\ -\sin\sqrt{ab}t & \cos\sqrt{ab}t \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

de unde

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \begin{pmatrix} p\sqrt{b} & 0 \\ 0 & -q\sqrt{a} \end{pmatrix} \begin{pmatrix} \cos\sqrt{ab}t & \sin\sqrt{ab}t \\ -\sin\sqrt{ab}t & \cos\sqrt{ab}t \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

Reprezentarea grafică a traiectoriilor (portretul de faze) este ilustrată mai jos:



Se observă că cele două specii *coexistă*; nici una dintre specii nu creşte indefinit şi există o perioadă $T = \frac{2\pi}{\sqrt{ab}}$ astfel încât $x(t+T) = x(t)$, $y(t+T) = y(t)$, deci numărul de indivizi din fiecare specie variază ciclic în timp.

Exerciţiu 2.10. Continuaţi analiza modelului pradă prădător cu liniarizarea în jurul punctului de echilibru $(0,0)$. Ce observaţi? Care este natura punctului de echilibru $(0,0)$?

Exemplu 2.11 (Modelul SIR)**.** SIR este un model matematic (Suspect-Infectat-Recuperat) al dinamicii unei epidemii în rândul unei populatii [3]. Modelul ilustrat mai jos este o varianta simplă deterministă (în semestrul II veţi întâlni şi cazul probabilistic), cu următoarele ipoteze:

- Există trei categorii de indivizi (trei stări ale sistemului dinamic): s (suspect/posibil sănătos pana la proba contrarie), i (infectat) şi r (recuperat/a fost infectat, s-a vindecat şi a dobândit imunitate/a decedat). În particular, un individ, odată infectat şi recuperat, nu se poate reinfecta.[44]

- Dinamica are loc în timp continuu, deci pentru fiecare $t \in \mathbb{R}$ vom nota cu $s(t)$ numărul indivizilor sănătoşi la momentul t, şi analog pentru $i(t)$ şi $r(t)$. Ne interesează, bineînţeles, doar soluţiile pozitive pentru $s(t), i(t)$ şi $r(t)$. Acesta este un model simplificat, dar cu toate acestea, putem obţine concluzii relevante asupra evoluţiei epidemice.

[44] Acesta este un model simplificat, dar aplicabil multor maladii infecţioase, cum ar fi rujeola, oreionul, varicela, etc. Presupunerea că indivizii recuperaţi se pot reinfecta creşte complexitatea sistemului de ecuaţii diferenţiale.

- Populaţia este omogenă, de volum constant N (fără naşteri/emigranţi); are deci loc relaţia $s(t) + i(t) + r(t) = N$, $\forall t \in \mathbb{R}$.
- Rata de infecţie $b > 0$ şi rata de recuperare $a > 0$ sunt constante.

În studiul oricărei maladii infecţioase, este esenţial studiul evoluţiei sale în timp: dacă epidemia se va împrăştia sau nu, şi, poate cel mai important, când va începe regresul. Vom urmări să determinăm aceste informaţii cunoscând rata de infecţie b şi cea de recuperare a, numărul iniţial al indivizilor sănătoşi s_0 şi numărul iniţial al celor infectaţi i_0.

$$\boxed{s} \xrightarrow{b} \boxed{i} \xrightarrow{a} \boxed{r}$$

Dinamica populaţiei este reprezentată prin următorul sistem de ecuaţii diferenţiale:

$$\begin{cases} s' &= -bsi \\ i' &= bsi - ai \\ r' &= \phantom{-bsi - {}} ai \end{cases} \tag{30}$$

Prima ecuaţie a sistemului diferenţial (30) reprezintă rata de variaţie a numărului de persoane sănătoase ($s(t)$ scade). Aceasta este proporţională cu numărul de indivizi sănătoşi (deci susceptibili a fi infectaţi) şi cu numărul indivizilor deja infectaţi (care la rândul lor, pot infecta alte persoane sănătoase).

A doua ecuatie indică variaţia numărului persoanelor infectate: o creştere în funcţie de indivizii nou infectaţi, şi o scădere corespunzând celor recuperati (vindecaţi şi imunizaţi/decedaţi).

În sfârşit, ultima relaţie este acum clară: variaţia în rândul persoanelor recuperate se datorează complet celor infectaţi care s-au vindecat.

Observaţi că însumând cele trei ecuaţii obţinem că volumul total al populaţiei nu variază în timp $s' + i' + r' = 0$, deci sistemul de mai sus este consistent cu presupunerea ca volumul total al populaţiei este constant.

Însă (30) nu este un sistem de ecuaţii diferenţiale *liniare*. Vom proceda atunci astfel: pentru început, să observăm că ne putem dispensa de ultima ecuaţie, soluţia acesteia fiind $r(t) = N - s(t) - i(t)$.

Pasul I: determinăm punctele de echilibru ale sistemului

$$\begin{cases} s' &= -\, bsi \\ i' &= bsi - ai \end{cases}$$

Găsim $i = 0$ şi s arbitrar (în intervalul $[0, N]$). Deci punctele de echilibru se găsesc pe un segment de pe axa orizontală (într-un sistem de coordonate (s, i)). Cazul extrem $s = 0, i = 0$ corespunde unei situaţii post-epidemice (când $r = N$), iar situaţia $s = N, i = 0, r = 0$ indică echilibrul în absenţa infecţiei.

Pasul II: liniarizăm sistemul în jurul unui punct de forma $(s_0, 0)$, cu $s_0 \in [0, N]$. Pentru aceasta, vom nota $F(s, i) = -bsi$ şi $G(s, i) = bsi - ai$.

Procedând ca în Exemplul 2.9 (modelul pradă-prădător), obţinem aproximarea liniara:

$$\begin{cases} F(s,i) \approx F(s_0,0) + \frac{\partial F}{\partial s}(s_0,0)(s-s_0) + \frac{\partial F}{\partial i}(s_0,0)(i-0) = -\, bs_0 i \\ G(s,i) \approx G(s_0,0) + \frac{\partial G}{\partial s}(s_0,0)(s-s_0) + \frac{\partial G}{\partial i}(s_0,0)(i-0) = (bs_0 - a)i \end{cases}$$

şi sistemul de ecuaţii diferenţiale liniare

$$\begin{cases} s' &= \quad\quad - bs_0 i \\ i' &= (bs_0 - a) i \end{cases}$$

Matricea sistemului este $A = \begin{pmatrix} 0 & -bs_0 \\ 0 & bs_0 - a \end{pmatrix}$ şi admite valorile proprii 0 şi $bs_0 - a$, deci este diagonalizabilă pentru $bs_0 \neq a$.[45] Este instructiv însă să observăm că a doua ecuaţie este independentă de prima şi o putem rezolva direct; obţinem că numărul indivizilor infectaţi la momentul t este

$$i(t) = C\mathrm{e}^{(bs_0 - a)t}$$

unde C este o constantă reală (pozitivă).

Pentru $bs_0 > a$, $\lim_{t\to\infty} i(t) = \infty$, deci infecţia se răspândeşte; cantitatea $R_0 = \frac{bs_0}{a}$ se numeşte **rata de reproducere de bază a infecţiei**) şi ilustrează raportul dintre infecţiile secundare (cele produse de prima infecţie) versus cazurile de vindecare/imunizare.

Întorcându-ne la sistem, obţinem că

$$s' = -bs_0 C\mathrm{e}^{(bs_0 - a)t}$$

de unde rezultă că numărul persoanelor sănătoase la momentul t este

$$s(t) = -\frac{bs_0}{bs_0 - a} C\mathrm{e}^{(bs_0 - a)t} + D$$

cu C şi D constante reale.

Dacă la sistemul de ecuaţii diferenţiale 30 adăgăm condiţiile iniţiale $s(0) = s_0, i(0) = i_0$, obţinem

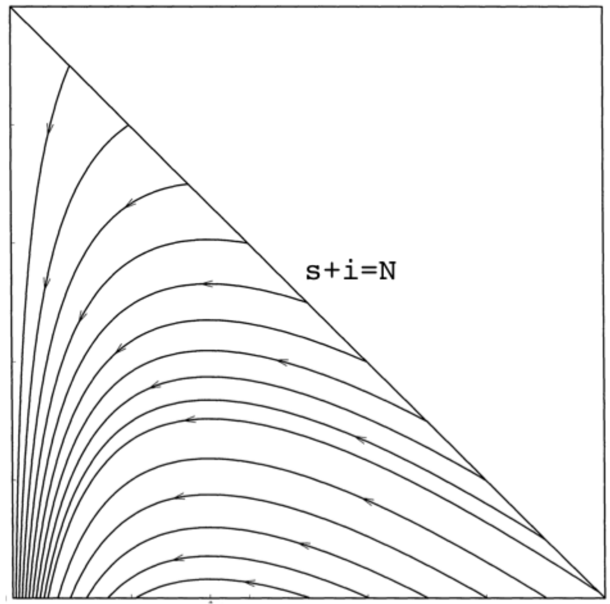
$$\begin{cases} s(t) &= s_0 - \frac{bs_0}{bs_0 - a} i_0 \left(\mathrm{e}^{(bs_0 - a)t} - 1\right) \\ i(t) &= i_0 \mathrm{e}^{(bs_0 - a)t} \end{cases}$$

Exerciţiu 2.12. Interpretaţi rezultatul obţinut pentru $s(t)$.

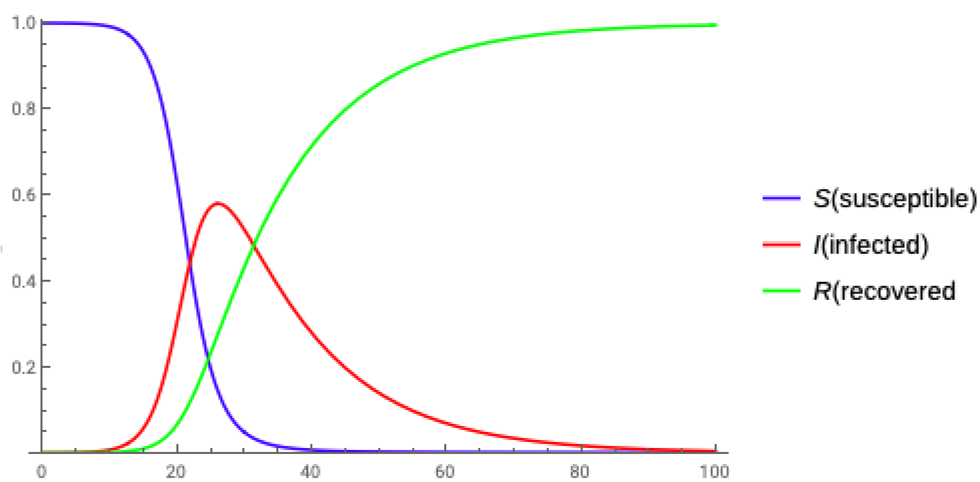
În general, pentru sistemul de ecuaţii diferenţiale 30 vom căuta soluţie în domeniul $\{(s, i) \mid s \geq 0, i \geq 0, s + i \leq N\}$. Conform Teoremei Cauchy-Lipschitz, în prezenţa condiţiilor iniţiale soluţia există şi este unică; în graficul de mai jos am reprezentat un model al traiectoriilor $(s(t), i(t))$. Se poate arăta că pentru $R_0 \leq 1$, $\lim_{t\to\infty} i(t) = 0$ (deci acelaşi comportament ca pentru sistemul liniarizat), iar pentru $R_0 > 1$, numărul indivizilor infectaţi $i(t)$ creşte până la o valoare maximă

[45] Cum interpretaţi condiţia $bs_0 = a$?

i_{max}, apoi tinde către 0 pentru $t \to \infty$.



Exerciţiu 2.13. Să presupunem că în medie, fiecare persoană infectată distribuie boala mai departe unei alte persoane o dată la două zile ($b = \frac{1}{2}$) şi că timpul mediu de recuperare este 14 zile (de unde $a = \frac{1}{14}$). Dacă la momentul iniţial $t = 0$, se cunoaşte că 0.01% din volumul total al populaţiei este infectată (pentru simplitate, vom lucra cu procente, $s(0) = 0.9999$ şi $r(0) = 0$), cum va evolua epidemia peste $t = 100$ zile?



În reprezentarea grafică de mai sus am indicat evoluţia celor trei componente ale populaţiei în $t = 100$ zile. Observaţi că numărul celor infectaţi atinge un punct de maxim după aproximativ 4 săptămâni, apoi scade către 0.

2.4 Sisteme de ecuaţii diferenţiale liniare cu coeficienţi constanţi – cazul neomogen

Reamintim că un sistem **neomogen** de ecuaţii diferenţiale liniare de ordin I cu coeficienţi constanţi are forma

$$\begin{cases} x_1' = a_{11}x_1 + \ldots + a_{1n}x_n + b_1(t) \\ \quad\vdots \\ x_n' = a_{n1}x_1 + \ldots + a_{nn}x_n + b_n(t) \end{cases}$$

unde

- t este variabila independentă;
- $x_1 = x_1(t), \ldots, x_n = x_n(t)$ sunt funcţiile necunoscute;
- $a_{ij} \in \mathbb{R}$, $i, j = \overline{1, n}$
- $b_1, \ldots, b_n : I \subseteq \mathbb{R} \to \mathbb{R}$ sunt funcţii continue

Dacă notăm $A = (a_{ij})_{i,j=1,n} \in \mathcal{M}_n(\mathbb{R})$, $\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{b}(t) = \begin{pmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{pmatrix}$, putem scrie sistemul sub formă matricială astfel:

$$\mathbf{X}' = A\mathbf{X} + \mathbf{b}(t)$$

Existenţa şi unicitatea soluţiei

Teoremă 2.14. Fie $A \in \mathcal{M}_n(\mathbb{R})$, $b_1, \ldots, b_n : I \subseteq \mathbb{R} \to \mathbb{R}$ funcţii continue, $t_0 \in I$ şi $\mathbf{x}_0 \in \mathbb{R}^n$. Atunci problema Cauchy

$$\begin{cases} \mathbf{X}' = A\mathbf{X} + \mathbf{b}(t) \\ \mathbf{X}(t_0) = \mathbf{x}_0 \end{cases}$$

admite soluţie pe tot intervalul I şi soluţia este unică.

Demonstraţie. Ca şi în cazul omogen, rezultă din Teorema Cauchy-Lipschitz 1.2. □

Rezolvarea sistemelor liniare neomogenen de ecuaţii diferenţiale

Propoziţie 2.15. Fie $A \in \mathcal{M}_n(\mathbb{R})$, $b_1, \ldots, b_n : I \subseteq \mathbb{R} \to \mathbb{R}$ funcţii continue.
Considerăm următoarele sisteme de ecuaţii diferenţiale:

$$\mathbf{X}' = A\mathbf{X} \tag{31}$$

$$\mathbf{X}' = A\mathbf{X} + \mathbf{b}(t) \tag{32}$$

Fie $\mathcal{S}_o$ spaţiul vectorial de dimensiune n al al soluţiilor sistemului omogen (31), şi $\mathcal{S}_{no}$ mulţimea soluţiilor sistemului neomogen (32).
Fie $\mathbf{X}_p \in \mathcal{S}_{no}$ o soluţie (particulară) a sistemului neomogen.

Atunci pentru orice funcţie derivabilă $\mathbf{X} : I \to \mathbb{R}^n$ are loc echivalenţa

$$\mathbf{X} \in \mathcal{S}_{no} \ \Leftrightarrow \ \mathbf{X} - \mathbf{X}_p \in \mathcal{S}_o$$

Demonstraţie. Exerciţiu. $\square$

Rezultatul de mai sus ne spune că este suficient a determina o *singură* soluţie pentru sistemul neomogen de ecuaţii diferenţiale, pentru a le cunoaşte pe toate (prin suprapoziţie), dacă ştim soluţia sistemului omogen asociat. Prezentăm mai jos o metodă pentru aceasta:

Metoda variaţiei constantelor

Propoziţie 2.16. Fie $\mathbf{X}_1, \ldots, \mathbf{X}_n : \mathbb{R} \to \mathbb{R}^n$ un sistem fundamental de soluţii pentru sistemul de ecuaţii diferenţiale liniare şi omogene $\mathbf{X}' = A\mathbf{X}$.
Atunci soluţia generală a sistemului neomogen $\mathbf{X}' = A\mathbf{X} + \mathbf{b}(t)$ este

$$\mathbf{X}(t) = C_1(t)\mathbf{X}_1(t) + \ldots + C_n(t)\mathbf{X}_n(t) = \mathbf{M}(t)\begin{pmatrix} C_1(t) \\ \vdots \\ C_n(t) \end{pmatrix}$$

unde $\mathbf{M}(t) = (\mathbf{X}_1(t) \mid \ldots \mid \mathbf{X}_n(t))$ este matricea fundamentală de soluţii, iar funcţiile $C_1, \ldots, C_n : I \to \mathbb{R}^n$ verifică relaţia

$$C_1'(t)\mathbf{X}_1(t) + \ldots + C_n'(t)\mathbf{X}_n(t) = \mathbf{M}(t)\begin{pmatrix} C_1'(t) \\ \vdots \\ C_n'(t) \end{pmatrix} = \mathbf{b}(t)$$

Deci

$$\mathbf{X}(t) = \mathbf{M}(t)\left(\mathbf{K} + \int \mathbf{M}^{-1}(t)\mathbf{b}(t)\mathsf{d}t\right)$$

unde $\mathbf{K} \in \mathbb{R}^n$ este un vector constant.
În particular, soluţia problemei Cauchy

$$\begin{cases} \mathbf{X}' &= A\mathbf{X} + \mathbf{b}(t) \\ \mathbf{X}(t_0) &= \mathbf{X}_0 \end{cases}$$

este

$$\mathbf{X}(t) = \mathbf{M}(t)\left(M^{-1}(t_0)\mathbf{X}_0 + \int_{t_0}^{t} \mathbf{M}^{-1}(s)\mathbf{b}(s)\mathsf{d}s\right)$$

Demonstraţie. Conform propoziţiei precedente, este suficient să determinăm doar o soluţie particulară pentru sistemul neomogen de ecuaţii diferenţiale (32). Căutăm o soluţie de forma $\mathbf{X}_p(t) = C_1(t)\mathbf{X}_1(t) + \ldots + C_n(t)\mathbf{X}_n(t)$. Punând condiţia ca aceasta să verifice sistemul neomogen (32), obţinem

$$C_1'(t)\mathbf{X}_1(t) + C_1(t)(\mathbf{X}_1)'(t) + \ldots + C_n'(t)\mathbf{X}_n(t) + C_n(t)(\mathbf{X}_n)'(t) = C_1(t)A\mathbf{X}_1(t) + \ldots + C_n(t)A\mathbf{X}_n(t) + \mathbf{b}(t)$$

Dar fiecare dintre funcţiile $\mathbf{X}_1(t), \ldots, \mathbf{X}_n(t)$ satisface sistemul de ecuaţii diferenţiale (31), deci putem simplifica şi obţinem

$$C_1'(t)\mathbf{X}_1(t) + \ldots + C_n'(t)\mathbf{X}_n(t) = \mathbf{b}(t)$$

Deoarece funcţiile vectoriale $\mathbf{X}_1(t), \ldots, \mathbf{X}_n(t)$ sunt liniar independente, matricea

$$\mathbf{M}(t) = (\mathbf{X}_1(t) \mid \ldots \mid \mathbf{X}_n(t))$$

este inversabilă pentru orice t, deci

$$\begin{pmatrix} C_1(t) \\ \vdots \\ C_n(t) \end{pmatrix} = \int \mathbf{M}^{-1}(t)\mathbf{b}(t)\mathsf{d}t$$

şi obţinem astfel formula dorită

$$\mathbf{X}(t) = \mathbf{X}_o(t) + \mathbf{X}_p(t) = \mathbf{M}(t)\mathbf{K} + \mathbf{M}(t)\int \mathbf{M}^{-1}(t)\mathbf{b}(t)\mathsf{d}t = \mathbf{M}(t)\left(\mathbf{K} + \int \mathbf{M}^{-1}(t)\mathbf{b}(t)\mathsf{d}t\right)$$

□

Exemplu 2.17. Să determinăm soluţia problemei Cauchy

$$\begin{cases} x' &= \quad 2x + 3y + 2 \quad , \quad x(0) = -2 \\ y' &= -x - 2y + 2t \quad , \quad y(0) = 0 \end{cases}$$

Vom rezolva mai întâi sistemul omogen de ecuaţii diferenţiale

$$\begin{cases} x' &= \quad 2x + 3y \\ y' &= -x - 2y \end{cases}$$

Matricea sistemului $A = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}$ admite valorile proprii $\lambda_1 = -1$, $\lambda_2 = 1$ (deci este diagonalizabilă), cu vectorii proprii corespunzători $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$. Conform Propoziţiei 2.4, soluţia sistemului omogen de mai sus este

$$\begin{pmatrix} x_o(t) \\ y_o(t) \end{pmatrix} = K_1\underbrace{\mathsf{e}^{\lambda_1 t}\mathbf{v}_1}_{\mathbf{X}_1(t} + K_2\underbrace{\mathsf{e}^{\lambda_2 t}\mathbf{v}_2}_{\mathbf{X}_2(t)} = K_1\mathsf{e}^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + K_2\mathsf{e}^{t}\begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

unde $K_1, K_2 \in \mathbb{R}$.

Trecem mai departe: determinăm o soluţie particulară a sistemului neomogen

$$\begin{pmatrix} x_p(t) \\ y_p(t) \end{pmatrix} = C_1(t)\mathbf{X}_1(t) + C_2(t)\mathbf{X}_2(t) = C_1(t)\mathsf{e}^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + C_2(t)\mathsf{e}^{t}\begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

cu metoda variaţiei constantelor. Deci funcţiile $C_1(t)$ şi $C_2(t)$ sunt soluţii ale sistemului

$$C_1'(t)\mathsf{e}^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + C_2'(t)\mathsf{e}^{t}\begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2t \end{pmatrix}$$

Obţinem $C_1'(t) = -\mathsf{e}^t(3t+1)$ şi $C_2'(t) = -\mathsf{e}^{-t}(t+1)$, care prin integrare conduc la $C_1(t) = \mathsf{e}^t(2-3t)$ şi $C_2(t) = \mathsf{e}^{-t}(t+2)$. Deci

$$\begin{pmatrix} x_p(t) \\ y_p(t) \end{pmatrix} = (2-3t)\begin{pmatrix} 1 \\ -1 \end{pmatrix} + (t+2)\begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} -6t-4 \\ 4t \end{pmatrix}$$

Atunci soluţia generală este

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_o(t) \\ y_o(t) \end{pmatrix} + \begin{pmatrix} x_p(t) \\ y_p(t) \end{pmatrix} = K_1\mathsf{e}^{-t}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + K_2\mathsf{e}^{t}\begin{pmatrix} -3 \\ 1 \end{pmatrix} + \begin{pmatrix} -6t-4 \\ 4t \end{pmatrix} = \begin{pmatrix} K_1\mathsf{e}^{-t} - 3K_2\mathsf{e}^t - 6t - 4 \\ -K_1\mathsf{e}^{-t} + K_2\mathsf{e}^t + 4t \end{pmatrix}$$

unde $K_1, K_2 \in \mathbb{R}$. Din condiţiile iniţiale $x(0) = -2, y(0) = 0$ obţinem $K_1 = K_2 = -1$, deci soluţia problemei Cauchy este

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} -\mathsf{e}^{-t} + 3\mathsf{e}^t - 6t - 4 \\ \mathsf{e}^{-t} - \mathsf{e}^t + 4t \end{pmatrix}$$

3 Ecuaţii diferenţiale liniare de ordin superior

3.1 Noţiuni generale

Definiţie 3.1. O *ecuaţie diferenţială liniară de ordin n cu coeficienţi constanţi* are forma

$$x^{(n)} + a_1x^{(n-1)} + \ldots + a_{n-1}x' + a_nx = b(t)$$

unde $a_1, \ldots a_n \in \mathbb{R}$ şi $b : I \subseteq \mathbb{R} \to \mathbb{R}$ este o funcţie continuă depinzând de variabila t, iar $x, x', \ldots x^{(n)}$ sunt funcţia necunoscută, respectiv derivatele sale până la ordinul n. O **soluţie** este o funcţie $x = x(t) : I \subseteq \mathbb{R} \to R$ derivabilă de n ori, astfel încât

$$x^{(n)}(t) + a_1x^{(n-1)}(t) + \ldots + a_{n-1}x'(t) + a_nx(t) = b(t)\,,\ \forall\, t \in I$$

Dacă $b(t) = 0, \forall\ t \in I$, ecuaţia se numeşte *omogenă*; în caz contrar spunem că este *neomogenă*.

3.2 Ecuatii diferenţiale liniare de ordin superior – cazul omogen

Existenţa/unicitatea/structura soluţiei

Teoremă 3.2. 1) Fie $a_1, \ldots a_n \in \mathbb{R}$. Atunci pentru orice $t_0 \in I$ şi α_0, $\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{R}$, problema Cauchy

$$\begin{cases} x^{(n)} + a_1x^{(n-1)} + \ldots + a_{n-1}x' + a_nx = 0 \\ x(t_0) = \alpha_0 \\ x'(t_0) = \alpha_1 \\ \vdots \\ x^{(n-1)}(t_0) = \alpha_{n-1} \end{cases}$$

admite soluţie pe $\mathbb{R}$ şi soluţia este unică.

2) Fie $\mathcal{E}_o$ mulţimea soluţiilor ecuaţiei diferenţiale omogene

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = 0 \tag{33}$$

Atunci $\mathcal{E}_o$ este un spaţiu vectorial real de dimensiune n.

Demonstraţie. 1) Rezultă din demonstraţia punctului 2) şi din Teorema 2.1.

2) Faptul că $\mathcal{E}_o$ este un spaţiu vectorial rezultă prin verificare directă. Pentru determinarea dimensiunii, vom proceda astfel: asociem ecuaţiei diferenţiale (33) următorul sistem de ecuaţii diferenţiale liniare de ordinul I:

$$\begin{cases} x_1' &= x_2 \\ x_2' &= x_3 \\ \vdots & \vdots \\ x_n' &= -a_n x_1 - a_{n-1}x_2 - \ldots - a_1 x_n \end{cases} \tag{34}$$

Atunci $x(t)$ este o soluţie a ecuaţiei diferenţiale (33) dacă şi numai dacă funcţiile $x_1(t) = x(t), x_2(t) = x'(t), \ldots, x_n(t) = x^{(n-1)}(t)$ sunt soluţii ale sistemului de ecuaţii diferenţiale liniare (34). Cu alte cuvinte, funcţia $f : \mathcal{E}_o \to \mathcal{S}_o$, $f(x(t)) = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix}$ este un izomorfism de spaţii vectoriale (cu inversa $f^{-1}\begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = x_1(t)$).

□

Ca în cazul sistemelor de ecuaţii diferenţiale, vom numi **sistem fundamental de soluţii** al ecuaţiei

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = 0$$

o bază $\{x_1(t), \ldots, x_n(t)\}$ în spaţiul vectorial al soluţiilor $\mathcal{E}_o$.

Cunoscând un sistem fundamental de soluţii, soluţia generală a ecuaţiei va fi

$$x(t) = C_1 x_1(t) + \ldots + C_n x_n(t)$$

unde $C_1, \ldots C_n \in \mathbb{R}$ sunt constante reale.

Observaţie 3.3. În teorema precedentă am văzut cum putem reduce rezolvarea oricărei ecuaţii diferenţiale liniare de ordin superior la un sistem de ecuaţii diferenţiale liniare de ordinul I. De fapt, procesul este reversibil: **orice** sistem de ecuaţii diferenţiale liniare de ordinul I se poate reduce la o ecuaţie diferenţială liniară de ordin superior. Pentru a arăta aceasta, fie $\mathbf{X}' = A\mathbf{X}$ un asemenea sistem, şi

$$P_A(X) = (-1)^n (X^n + a_1 X^{n-1} + \ldots + a_{n-1}X + a_n)$$

polinomul caracteristic al matricei A. Conform Teoremei Hamilton-Cayley,

$$A^n + a_1 A^{n-1} + \ldots + a_{n-1} A + a_n I_n = 0_n$$

Fie acum $\mathbf{X}(t)$ o soluţie a sistemului de ecuaţii diferenţiale $\mathbf{X}' = A\mathbf{X}$. Derivând succesiv, obţinem $\mathbf{X}'' = A\mathbf{X}' = A^2\mathbf{X}$, $\mathbf{X}''' = A^3\mathbf{X}$, $\ldots$, $\mathbf{X}^{(n)} = A^n\mathbf{X}$. Atunci

$$\mathbf{X}^{(n)} + a_1\mathbf{X}^{(n-1)} + \ldots + a_{n-1}\mathbf{X}' + a_n\mathbf{X} = (A^n + a_1 A^{n-1} + \ldots + a_{n-1}A + a_n I_n)\mathbf{X} = \mathbf{0}$$

În particular, fiecare componentă a funcţiei vectoriale $\mathbf{X}(t)$ verifică ecuaţia (scalară) diferenţială de mai sus, ai cărei coeficienţi provin din polinomul caracteristic al matricei A. Această procedură de transformare a unui sistem de ecuaţii diferenţiale liniare într-o singură ecuaţie diferenţială liniară, dar de ordin superior, se numeşte **metoda derivării şi eliminării.** Vom ilustra această metodă în Secţiunea 3.6, după ce vom studia complet rezolvarea ecuaţiilor diferenţiale liniare de ordin superior.

3.3 Ecuaţii diferenţiale liniare de ordin $n = 2$ omogene

În acest caz, ecuaţia diferenţială este $x'' + ax' + bx = 0$, unde $a, b \in \mathbb{R}$. Vom căuta soluţii de forma $x(t) = \mathsf{e}^{\lambda t}$. Înlocuind în ecuaţia diferenţială, obţinem $\mathsf{e}^{\lambda t}(\lambda^2 + a\lambda + b) = 0$. Vom numi ecuaţia

$$\lambda^2 + a\lambda + b = 0$$

ecuaţia caracteristică asociată ecuaţiei diferenţiale.

Fie λ_1, λ_2 rădăcinile ecuaţiei caracteristice. Distingem trei situaţii:

Cazul I. $\lambda_1 \neq \lambda_2 \in \mathbb{R}$ sunt rădăcini distincte reale ale ecuaţiei caracteristice.

Propoziţie 3.4. 1) Funcţiile

$$x_1(t) = \mathsf{e}^{\lambda_1 t}\ ,\ x_2(t) = \mathsf{e}^{\lambda_2 t}$$

formează un sistem fundamental de soluţii pentru ecuaţia diferenţială $x'' + ax' + bx = 0$.

2) Soluţia generală a ecuaţiei diferenţiale $x'' + ax' + bx = 0$ este

$$x(t) = C_1 x_1(t) + C_2 x_2(t) = C_1\mathsf{e}^{\lambda_1 t} + C_2\mathsf{e}^{\lambda_2 t}, \quad \text{unde} \quad C_1, C_2 \in \mathbb{R}$$

Demonstraţie. Din modul în care s-au obţinut λ_1, λ_2 rezultă că funcţiile $x_1(t)$ şi $x_2(t)$ sunt soluţii ale ecuaţiei diferenţiale $x'' + ax' + bx = 0$. Conform Teoremei 3.2, este suficient deci să stabilim liniar independenţa celor două funcţii. Dacă $\alpha_1, \alpha_2 \in \mathbb{R}$ satisfac $\alpha_1\mathsf{e}^{\lambda_1 t} + \alpha_2\mathsf{e}^{\lambda_2 t} = 0$, $\forall\, t \in \mathbb{R}$, atunci prin derivare obţinem $\alpha_1\lambda_1\mathsf{e}^{\lambda_1 t} + \alpha_2\lambda_2\mathsf{e}^{\lambda_2 t} = 0$, $\forall\, t \in \mathbb{R}$. Rezultă sistemul liniar omogen

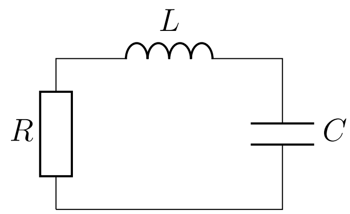
$$\begin{cases} \alpha_1\mathsf{e}^{\lambda_1 t} + \alpha_2\mathsf{e}^{\lambda_2 t} = 0 \\ \alpha_1\lambda_1\mathsf{e}^{\lambda_1 t} + \alpha_2\lambda_2\mathsf{e}^{\lambda_2 t} = 0 \end{cases}$$

cu necunoscutele α_1, α_2. Deoarece $\begin{vmatrix} \mathsf{e}^{\lambda_1 t} & \mathsf{e}^{\lambda_2 t} \\ \lambda_1\mathsf{e}^{\lambda_1 t} & \lambda_2\mathsf{e}^{\lambda_2 t} \end{vmatrix} = \mathsf{e}^{(\lambda_1+\lambda_2)t}(\lambda_2 - \lambda_1) \neq 0$, $\alpha_1 = \alpha_2 = 0$. $\square$

Exemplu 3.5. Într-un circuit RLC sunt cuplate în serie următoarele componente: o rezistenţă de 3 Ω, o bobină cu inductanţa de 1 H şi un condensator de 0.5 F. La momentul $t = 0$ condensatorul este încărcat cu sarcina de 2 C, iar curentul în circuit este $I(0) = 4$ A.

1) Să se determine evoluţia sarcinii electrice $Q(t)$ în acest circuit.

2) Să se compare rezultatul obţinut cu cel corespunzător cazurilor $I_1(0) = 0$ A şi $I_2(0) = 8$ A. Ce observaţi?



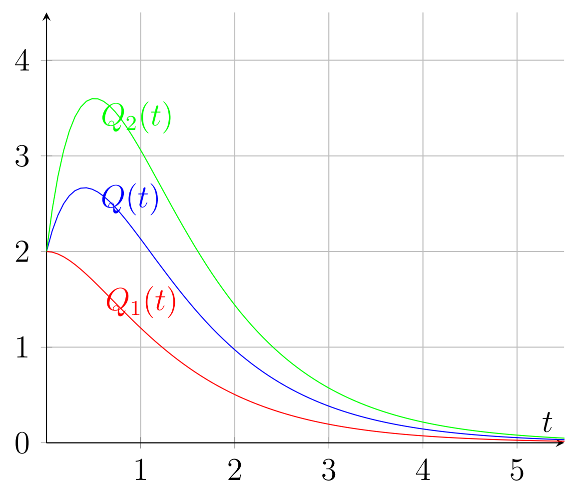
1) Legea lui Kirchhoff a tensiunilor pentru acest circuit se exprimă, în cazul de faţă, sub forma

$$LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = 0$$

Introducând datele problemei obţinem ecuaţia diferenţială liniară $Q''(t) + 3Q'(t) + 2Q(t) = 0$, cu condiţiile iniţiale $Q(0) = 2$ C, $Q'(0) = I(0) = 4$ A.

Ecuaţia caracteristică asociată este $\lambda^2 + 3\lambda + 2 = 0$ cu rădacinile -1 şi -2. Deci soluţia generală va fi $Q(t) = C_1\mathsf{e}^{-t} + C_2\mathsf{e}^{-2t}$. Din condiţiile iniţiale rezultă $C_1 = 8$, $C_2 = -6$, deci soluţia problemei Cauchy este $Q(t) = 8\mathsf{e}^{-t} - 6\mathsf{e}^{-2t}$. Se observă că sarcina electrică creşte până la momentul $t_0 = \ln\frac{3}{2} \approx 0.4$ secunde când atinge valoarea maximă $\frac{8}{3}$ C, apoi scade $\lim_{t\to\infty} Q(t) = 0$, deci are loc o descărcare a sarcinii electrice în circuitul dat.

2) Condiţia iniţială $I_1(0) = 0$ A produce soluţia $Q_1(t) = 4\mathsf{e}^{-t} - 2\mathsf{e}^{-2t}$. Pentru $I_2(0) = 8$ A obţinem $Q_2(t) = 12\mathsf{e}^{-t} - 10\mathsf{e}^{-2t}$. Graficele celor trei funcţii sunt reprezentate mai jos.



Exerciţiu 3.6. Interpretaţi rezultatele obţinute.

Cazul II. $\lambda \in \mathbb{R}$ este rădăcină dublă a ecuaţiei caracteristice.

Propoziţie 3.7. 1) Funcţiile

$$x_1(t) = \mathsf{e}^{\lambda t}\ ,\ x_2(t) = t\mathsf{e}^{\lambda t}$$

formează un sistem fundamental de soluţii pentru ecuaţia diferenţială $x'' + ax' + bx = 0$.

2) Soluţia generală a ecuaţiei diferenţiale $x'' + ax' + bx = 0$ este

$$x(t) = C_1 x_1(t) + C_2 x_2(t) = C_1 \mathsf{e}^{\lambda t} + C_2 t\mathsf{e}^{\lambda t}\ ,\quad \text{unde}\ \ C_1, C_2 \in \mathbb{R}$$

Demonstraţie. Ştim că funcţia $x_1(t) = \mathsf{e}^{\lambda t}$ este soluţie a ecuaţiei diferenţiale $x'' + ax' + bx = 0$. Verificăm acum că şi $x_2(t) = t\mathsf{e}^{\lambda t}$ este soluţie a aceleiaşi ecuaţii. Avem

$$x'(t) = (\lambda t + 1)\mathsf{e}^{\lambda t}\ ,\ x''(t) = (\lambda^2 t + 2\lambda)\mathsf{e}^{\lambda t}$$

de unde rezultă

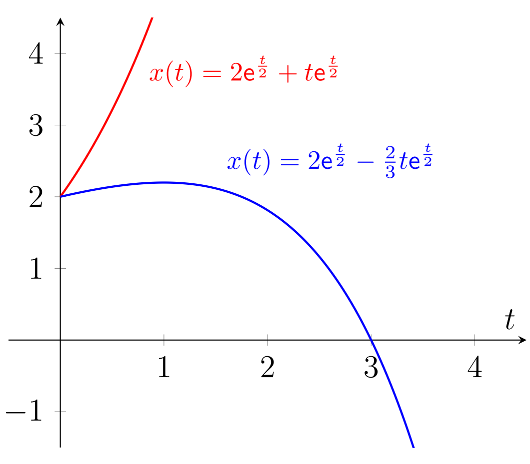
$$(\lambda^2 t + 2\lambda)\mathsf{e}^{\lambda t} + a(\lambda t + 1)\mathsf{e}^{\lambda t} + bt\mathsf{e}^{\lambda t} = ((\lambda^2 + a\lambda + b)t + (2\lambda + a))\mathsf{e}^{\lambda t} = 0$$

Mai sus am utilizat că λ este rădăcină *dublă* a ecuaţiei caracteristice. Lăsăm ca exerciţiu verificarea liniar independenţei celor două funcţii. $\square$

Exemplu 3.8. Să considerăm ecuaţia diferenţială $x'' - x' + \frac{1}{4}x = 0$, cu condiţiile iniţiale $x(0) = 2, x'(0) = \frac{1}{3}$.

Ecuaţia caracteristică asociată este $\lambda^2 - \lambda + \frac{1}{4} = 0$, cu rădăcina dublă $\lambda = \frac{1}{2}$. Atunci $x(t) = C_1\mathsf{e}^{\frac{t}{2}} + C_2 t\mathsf{e}^{\frac{t}{2}}$, cu $C_1, C_2 \in \mathbb{R}$, este soluţia generală a ecuaţiei diferenţiale. Condiţiile iniţiale $x(0) = 2, x'(0) = \frac{1}{3}$ conduc la $x(t) = 2\mathsf{e}^{\frac{t}{2}} - \frac{2}{3}t\mathsf{e}^{\frac{t}{2}}$. Soluţia obţinută se anulează pentru $t = 3$, apoi scade spre $-\infty$.

Dacă modificăm condiţiile iniţiale ale problemei Cauchy, mai precis dacă luăm $x'(0) = 2$, atunci obţinem soluţia $x(t) = 2\mathsf{e}^{\frac{t}{2}} + t\mathsf{e}^{\frac{t}{2}}$. Spre deosebire de soluţia precedentă, aceasta este o funţie crescătoare şi tinde spre ∞. Reprezentând grafic cele două soluţii, obţinute prin variaţia valorii $x'(0)$ (panta tangentei la graficul soluţiei în $t = 0$), observăm că există o pantă iniţială $x'(0)$ între $\frac{1}{3}$ şi 2, care separă soluţiile nemărginite crescătoare de cele descrescătoare, care vor tinde către $-\infty$ pentru $t \to \infty$.



Exerciţiu 3.9. 1) Arătaţi că soluţia ecuaţiei diferenţiale precedente depinde continuu de condiţiile iniţiale.

2) Deduceţi că pentru $x'(0) > \frac{x(0)}{2}$, soluţia problemei Cauchy verifică $\lim_{t\to\infty} x(t) = \infty$, iar pentru $x'(0) < \frac{x(0)}{2}$ are loc relaţia $\lim_{t\to\infty} x(t) = -\infty$. Ce se întâmplă pentru $x'(0) = \frac{x(0)}{2}$, în particular pentru $x'(0) = 1$?

Cazul III. $\lambda_1 = \alpha + \mathrm{i}\beta$, $\lambda_2 = \alpha - \mathrm{i}\beta \in \mathbb{C} \setminus \mathbb{R}$ sunt rădăcini complexe conjugate ale ecuaţiei caracteristice.

Propoziţie 3.10. 1) Funcţiile

$$x_1(t) = \mathrm{e}^{\alpha t}\cos\beta t\ ,\ \ x_2(t) = \mathrm{e}^{\alpha t}\sin\beta t$$

formează un sistem fundamental de soluţii pentru ecuaţia diferenţială $x'' + ax' + bx = 0$.

2) Soluţia generală a ecuaţiei diferenţiale $x'' + ax' + bx = 0$ este

$$x(t) = C_1x_1(t) + C_2x_2(t) = C_1\mathrm{e}^{\alpha t}\cos\beta t + C_2\mathrm{e}^{\alpha t}\sin\beta t\ , \quad \text{unde} \ \ C_1, C_2 \in \mathbb{R}$$

Demonstraţie. Deoarece $\lambda = \alpha + \mathrm{i}\beta$ este rădăcină a ecuaţiei caracteristice, funcţia (complexă) $x(t) = \mathrm{e}^{\lambda t} = \mathrm{e}^{\alpha t}(\cos\beta t + \mathrm{i}\sin\beta t) = x_1(t) + \mathrm{i}x_2(t)$ este soluţie a ecuaţiei diferenţiale $x'' + ax' + b = 0$. Cum această ecuaţie diferenţială are coeficienţi reali, rezultă că şi $x_1(t) = \mathrm{Re}\, x(t)$ şi $x_2(t) = \mathrm{Im}\, x(t)$ sunt soluţii (reale, de data aceasta). Lăsăm ca exerciţiu verificarea liniar independenţei. □

Exemplu 3.11. Să considerăm problema Cauchy $16x'' - 8x' + 145x = 0\,,\ x(0) = -2, x'(0) = 1$.

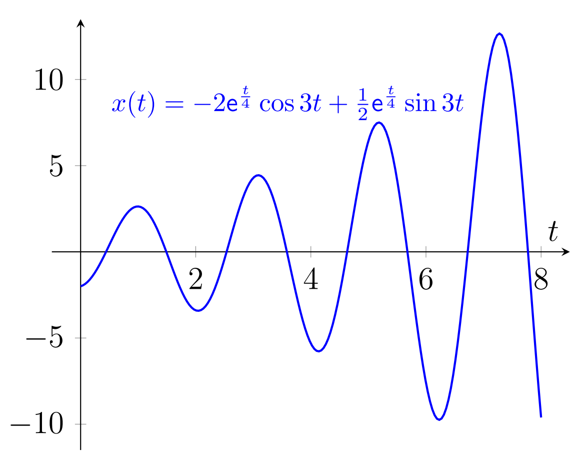
Ecuaţia caracteristică este $\lambda^2 - \frac{1}{2} + \frac{145}{16} = 0$, cu rădăcinile $\lambda_{1,2} = \frac{1}{4} \pm 3\,\mathrm{i}$. Rezultă că soluţia generală a ecuaţiei diferenţiale este

$$x(t) = C_1\mathrm{e}^{\frac{t}{4}}\cos 3t + C_2\mathrm{e}^{\frac{t}{4}}\sin 3t$$

unde C_1, C_2 sunt constante reale. Din condiţiile iniţiale obţinem $C_1 = -2, C_2 = \frac{1}{2}$. Deci soluţia problemei Cauchy este

$$x(t) = -2\mathrm{e}^{\frac{t}{4}}\cos 3t + \frac{1}{2}\mathrm{e}^{\frac{t}{4}}\sin 3t$$

Se observă că soluţia obţinută este oscilatorie, cu perioada $\frac{2\pi}{3}$ şi amplitudinea crescătoare exponenţial:



3.4 Ecuaţii diferenţiale liniare de ordin superior omogene (cazul general)

Fie ecuaţia diferenţială liniară de ordin n cu coeficienţi constanţi:

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = 0$$

În vederea determinării soluţiei generale, vom proceda astfel:

Pasul I. Scriem ecuaţia caracteristică asociată:

$$\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0$$

şi aflăm rădăcinile acesteia: $\lambda_1, \ldots, \lambda_p$ cu ordinele de multiplicitate $m_1, \ldots, m_p$ (unde $m_1+\ldots+m_p = n$).

Pasul II. Determinăm un sistem fundamental de soluţii $x_1(t), \ldots, x_n(t)$:

- Dacă $\lambda \in \mathbb{R}$ este rădăcină reală de multiplicitate m, atunci acesteia îi corespund m soluţii liniar independente
$$\mathsf{e}^{\lambda t}, t\mathsf{e}^{\lambda t}, \ldots, t^{m-1}\mathsf{e}^{\lambda t}$$
- Dacă $\lambda = \alpha + \mathsf{i}\beta, \overline{\lambda} = \alpha - \mathsf{i}\beta \in \mathbb{C}$ sunt rădăcini complexe conjugate de multiplicitate m, atunci acestora le vor corespunde $2m$ soluţii liniar independente
$$\begin{gathered}\mathsf{e}^{\alpha t}\cos\beta t, t\mathsf{e}^{\alpha t}\cos\beta t, \ldots, t^{m-1}\mathsf{e}^{\alpha t}\cos\beta t\\ \mathsf{e}^{\alpha t}\sin\beta t,\ t\mathsf{e}^{\alpha t}\sin\beta t,\ \ldots,\ t^{m-1}\mathsf{e}^{\alpha t}\sin\beta t\end{gathered}$$
- Rezultă astfel n funcţii (întrucât $m_1 + \ldots m_p = n$) liniar independente, soluţii ale ecuaţiei diferenţiale.

Pasul III. Soluţia generală va fi atunci

$$x(t) = C_1x_1(t) + \ldots + C_nx_n(t)\,,\ C_1, \ldots C_n \in \mathbb{R}$$

Exemplu 3.12. Să rezolvăm următoarele ecuaţii diferenţiale liniare de ordin superior:

1) $x^{(\mathsf{iv})} - 5x''' + 6x'' + 4x' - 8x = 0$. Ecuaţia caracteristică asociată este $\lambda^4 - 5\lambda^3 + 6\lambda^2 + 4\lambda - 8 = 0$, cu rădăcinile $\lambda_1 = -1$, cu multiplicitate 1, şi $\lambda_2 = 2$, rădăcină triplă. Acestora le vor corespunde următoarele funcţii liniar independente, soluţii ale ecuaţiei diferenţiale date:
$$\lambda_1 = -1 \Rightarrow x_1(t) = \mathsf{e}^{-t} \qquad \lambda_2 = 2 \Rightarrow \begin{cases} x_2(t) = \mathsf{e}^{2t}\\ x_3(t) = t\mathsf{e}^{2t}\\ x_4(t) = t^2\mathsf{e}^{2t}\end{cases}$$
Rezultă că soluţia generală a ecuaţiei diferenţiale este
$$x(t) = \sum_{i=1}^{4} C_i x_i(t) = C_1\mathsf{e}^{-t} + C_2\mathsf{e}^{2t} + C_3t\mathsf{e}^{2t} + C_4t^2\mathsf{e}^{2t}$$

2) $x^{(\mathsf{iv})} + 4x''' + 8x'' + 8x' + 4x = 0$, $x(0) = 1$, $x'(0) = 0$, $x''(0) = 2$, $x'''(0) = 0$. De data aceasta avem de rezolvat o problemă Cauchy. Ecuaţia caracteristică asociată ecuaţiei diferenţiale este $\lambda^4+4\lambda^3+8\lambda^2+8\lambda+4 = 0$, cu rădăcinile duble $-1\pm\mathsf{i}$. Acestora le vor corespunde următoarele soluţii fundamentale:

$$x_1(t) = \mathsf{e}^{-t}\cos t, x_2(t) = \mathsf{e}^{-t}\sin t, x_3(t) = t\mathsf{e}^{-t}\cos t, x_4(t) = t\mathsf{e}^{-t}\sin t$$

Rezultă că soluţia generală a ecuaţiei diferenţiale este

$$x(t) = \sum_{i=1}^{4} C_i x_i(t) = \mathsf{e}^{-t}\left((C_1 + C_3 t)\cos t + (C_2 + C_4 t)\sin t\right)$$

Din condiţiile iniţiale rezultă $C_1 = 1$, $C_2 = 5$, $C_3 = -4$, $C_4 = 2$, deci soluţia problemei Cauchy este

$$x(t) = \mathsf{e}^{-t}\left((-4t + 1)\cos t + (2t + 5)\sin t\right)$$

3) $x''' + 3x'' + 4x' + 12x = 0$, $x(0) = 0$, $x'(0) = 0$, $x''(0) = 13$. Ecuaţia caracteristică este $\lambda^3 + 3\lambda^2 + 4\lambda + 12 = 0$, cu rădăcinile simple $\lambda_1 = -3$, $\lambda_{2,3} = \pm 2i$. Atunci

$$\lambda_1 = -3 \Rightarrow x_1(t) = \mathsf{e}^{-3t} \qquad \lambda_{2,3} = \pm 2i \Rightarrow \begin{cases} x_2(t) = \cos 2t \\ x_3(t) = \sin 2t \end{cases}$$

şi deci soluţia generală va fi

$$x(t) = C_1 x_1(t) + C_2 x_2(t) + C_3 x_3(t) = C_1 \mathsf{e}^{-3t} + C_2 \cos 2t + C_3 \sin 2t$$

Dar $x(0) = 0$, $x'(0) = 0$, $x''(0) = 13$, de unde rezultă $C_1 = 1$, $C_2 = -1$, $C_3 = \frac{3}{2}$ şi

$$x(t) = \mathsf{e}^{-3t} - \cos 2t + \frac{3}{2}\sin 2t$$

3.5 Ecuaţii diferenţiale liniare neomogene de ordin superior cu coeficienţi constanţi

Existenţa şi unicitatea soluţiei

Teoremă 3.13. Fie $a_1, \ldots, a_n \in R$ şi $b : I \subseteq \mathbb{R} \to \mathbb{R}$ o funcţie continuă. Atunci pentru orice $t_0 \in I$ şi $\alpha_0, \alpha_1, \ldots, \alpha_{n-1} \in \mathbb{R}$, problema Cauchy

$$\begin{cases} x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = b(t) \\ x(t_0) = \alpha_0 \\ x'(t_0) = \alpha_1 \\ \vdots \\ x^{(n-1)}(t_0) = \alpha_{n-1} \end{cases}$$

admite soluţie pe întreg intervalul I şi soluţia este unică.

Demonstraţie. Rezultă din Teoremele 1.2 şi 2.14, coroborate cu demonstraţia Teoremei 3.2. □

Structura soluţiilor

Propoziţie 3.14. Fie $a_1, \dots, a_n \in \mathbb{R}$ şi $b : I \subseteq \mathbb{R} \to \mathbb{R}$ o funcţie continuă.
Considerăm următoarele ecuaţii diferenţiale:

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = 0 \qquad (35)$$

$$x^{(n)} + a_1 x^{(n-1)} + \ldots + a_{n-1}x' + a_n x = b(t) \qquad (36)$$

Fie $\mathcal{E}_o$ spaţiul vectorial de dimensiune n al al soluţiilor ecuaţiei (35) şi $\mathcal{E}_{no}$ mulţimea soluţiilor ecuaţiei neomogene (36).
Fie $x_p \in \mathcal{E}_{no}$ o soluţie (particulară) a ecuaţiei neomogene.
Atunci pentru orice funcţie n-derivabilă $x : I \to \mathbb{R}^n$ are loc echivalenţa

$$x \in \mathcal{E}_{no} \;\Leftrightarrow\; x - x_p \in \mathcal{E}_o$$

Demonstraţie. Exerciţiu. □

Informal, rezultatul de mai sus se mai poate scrie

$$\mathcal{E}_{no} = \mathcal{E}_o + \{x_p\}$$

Pentru determinarea unei soluţii particulare şi implicit a soluţiei generale a ecuaţiei diferenţiale 36, vom proceda ca la sistemele de ecuaţii diferenţiale:

Metoda variaţiei constantelor

Propoziţie 3.15. Fie $\{x_1(t), \ldots x_n(t)\}$ un sistem fundamental de soluţii pentru ecuaţia diferenţială omogenă (35) (o bază în spaţiul vectorial al soluţiilor $\mathcal{E}_o$).
Atunci soluţia generală a ecuaţiei neomogene (36) este

$$x(t) = C_1(t)x_1(t) + \ldots + C_n(t)x_n(t)$$

unde funcţiile $C_1, \ldots C_n : I \to \mathbb{R}$ verifică relaţiile

$$\begin{cases} C_1'(t)x_1(t) & + & \ldots & + & C_n'(t)x_n(t) & = & 0 \\ C_1'(t)x_1'(t) & + & \ldots & + & C_n'(t)x_n'(t) & = & 0 \\ \vdots & & & & & & \vdots \\ C_1'(t)x_1^{(n-1)}(t) & + & \ldots & + & C_n'(t)x_n^{(n-1)}(t) & = & b(t) \end{cases}$$

Demonstraţie. Rezultă din Propoziţia 2.16 şi din demonstraţia Teoremei 3.2. □

Exemplu 3.16. $x'' + x = \mathsf{tg}t$, $t \in \left(0, \frac{\pi}{2}\right)$. Ecuaţia omogenă asociată $x'' + x = 0$ admite ecuaţia caracteristică $\lambda^2 + 1 = 0$, cu rădăcinile simple $\pm\mathrm{i}$, deci soluţia generală a ecuaţiei diferenţiale omogene este $x_o(t) = C_1 \cos t + C_2 \sin t$. Aplicăm acum metoda variaţiei constantelor pentru a determina soluţia ecuaţiei diferenţiale neomogene de forma $x(t) = C_1(t)\cos t + C_2(t)\sin t$, unde funcţiile $C_1(t)$ şi $C_2(t)$ satisfac sistemul

$$\begin{cases} C_1'(t)\cos t + C_2'(t)\sin t = 0 \\ -C_1'(t)\sin t + C_2'(t)\cos t = \mathsf{tg}t \end{cases}$$

Obţinem soluţia $C_1'(t) = -\frac{\sin^2 t}{\cos t}$, $C_2'(t) = \sin t$ care prin integrare conduce la $C_1(t) = \sin t + \frac{1}{2}\ln\frac{1-\sin t}{1+\sin t} + D_1$, $C_2(t) = -\cos t + D_2$. Deci soluţia generală a ecuaţiei diferenţiale este $x(t) = D_1 \cos t + D_2 \sin t + \frac{1}{2}\cos t \ln\frac{1-\sin t}{1+\sin t}$.

Exerciţiu 3.17. Arătaţi că soluţia problemei Cauchy

$$x'' + a_1 x' + a_2 x = b(t)\,,\; x(0) = x_1, x'(0) = x_1$$

se poate scrie sub forma

$$x(t) = y(t) + z(t)$$

unde $y(t)$ şi $z(t)$ sunt soluţiile problemelor Cauchy

$$y'' + a_1 y' + a_2 y = 0\,,\; y(0) = x_1, y'(0) = x_1$$

respectiv

$$z'' + a_1 z' + a_2 z = b(t)\,,\; z(0) = 0, z'(0) = 0$$

Cu alte cuvinte, non-omogenitatea din ecuaţia diferenţială şi din condiţiile iniţiale pot fi considerate separat.

Observaţie 3.18. Uneori, pentru determinarea unei soluţii particulare $x_p(t)$ (şi implicit a soluţiei generale $x(t)$) a ecuaţiei diferenţiale 36, putem aplica şi o altă metodă: **metoda coeficienţilor nedeterminaţi**. Aceasta este valabilă doar când termenul liber $b(t)$ are anumite forme:

1) $b(t) = P(t)$ polinom.
   - Dacă $\lambda = 0$ nu este rădăcină a a ecuaţiei caracteristice, se caută $x_p(t)$ de forma $Q(t)$ polinom, cu $\mathsf{grad}Q = \mathsf{grad}P$.
   - Dacă $\lambda = 0$ este rădăcină a ecuaţiei caracteristice cu multiplicitatea s, se caută $x_p(t)$ de forma $t^s Q(t)$, cu $Q(t)$ polinom, $\mathsf{grad}Q = \mathsf{grad}P$.

2) $b(t) = \mathsf{e}^{\lambda t} P(t)$, cu $P(t)$ polinom.
   - Dacă λ nu este rădăcină a ecuaţiei caracteristice, se caută $x_p(t)$ de forma $\mathsf{e}^{\lambda t} Q(t)$ polinom, cu $\mathsf{grad}Q = \mathsf{grad}P$.
   - Dacă λ este rădăcină a ecuaţiei caracteristice cu multiplicitatea s, se caută $x_p(t)$ de forma $t^s \mathsf{e}^{\lambda t} Q(t)$, cu $Q(t)$ polinom, $\mathsf{grad}Q = \mathsf{grad}P$.

3) $b(t) = \mathsf{e}^{\alpha t}(C_1 \cos(\beta t) + C_2 \sin(\beta t))$, unde $C_1, C_2 \in \mathbb{R}$.
   - Dacă $\lambda = \alpha + \mathsf{i}\beta$ nu este rădăcină a ecuaţiei caracteristice, se caută $x_p(t)$ de forma $\mathsf{e}^{\alpha t}(D_1 \cos(\beta t) + D_2 \sin(\beta t))$, cu $D_1, D_2 \in \mathbb{R}$.
   - Dacă $\lambda = \alpha + \mathsf{i}\beta$ este rădăcină a ecuaţiei caracteristice cu multiplicitatea s, se caută $x_p(t)$ de forma $t^s \mathsf{e}^{\alpha t}(D_1 \cos(\beta t) + D_2 \sin(\beta t))$, cu $D_1, D_2 \in \mathbb{R}$.

4) $b(t) = \mathsf{e}^{\alpha t}(P_1(t) \cos(\beta t) + P_2(t) \sin(\beta t))$, $P_1(t), P_2(t)$ polinoame.
   - Dacă $\lambda = \alpha + \mathsf{i}\beta$ nu este rădăcină a ecuaţiei caracteristice, se caută $x_p(t)$ de forma $\mathsf{e}^{\alpha t}(Q_1(t) \cos(\beta t) + Q_2(t) \sin(\beta t))$, cu $Q_1(t), Q_2(t)$ polinoame de grad $\mathsf{max}(\mathsf{grad}P_1, \mathsf{grad}P_2)$.

- Dacă $\lambda = \alpha + \mathrm{i}\beta$ este rădăcină a ecuaţiei caracteristice cu multiplicitatea s, se caută $x_p(t)$ de forma $t^s \mathrm{e}^{\alpha t}(Q_1(t)\cos(\beta t) + Q_2(t)\sin(\beta t))$, cu $Q_1(t), Q_2(t)$ polinoame de grad $\mathsf{max}(\mathsf{grad} P_1, \mathsf{grad} P_2)$.

Exemplu 3.19. 1) Să determinăm soluţia generală a ecuaţiei diferenţiale $x'' - 3x' - 4x = 3\mathrm{e}^{2t}$. Ecuaţia diferenţială omogenă asociată admite soluţia generală $x_o(t) = C_1\mathrm{e}^{4t} + C_2\mathrm{e}^{-t}$ (rădăcinile ecuaţiei caracteristice asociate fiind $\lambda_1 = 4$ şi $\lambda_2 = -1$). Termenul liber al ecuaţiei diferenţiale neomogene este $b(t) = 3\mathrm{e}^{2t}$. Deoarece $\lambda = 2$ nu este rădăcină a ecuaţiei caracteristice, conform Observaţiei 3.18.2, căutăm o soluţie particulară de forma $x_p(t) = A\mathrm{e}^{2t}$, unde $A \in \mathbb{R}$ este o constantă ce trebuie determinată. Punând condiţia ca $x_p(t)$ să verifice ecuaţia diferenţială, obţinem

$$4A\mathrm{e}^{2t} - 6A\mathrm{e}^{2t} - 4A\mathrm{e}^{2t} = 3\mathrm{e}^{2t}$$

de unde $A = -\frac{1}{2}$ şi $x_p(t) = -\frac{1}{2}\mathrm{e}^{2t}$. Deci soluţia generală a ecuaţiei diferenţiale date este $x(t) = x_o(t) + x_p(t) = C_1\mathrm{e}^{4t} + C_2\mathrm{e}^{-t} - \frac{1}{2}\mathrm{e}^{2t}$.

2) Să găsim o soluţie particulară a ecuaţiei diferenţiale $x'' - 3x' - 4x = 13\mathrm{e}^t \cos 2t$. Dacă urmărim doar forma termenului liber $b(t) = 13\mathrm{e}^t \cos 2t$, am fi tentaţi să căutăm o soluţie particulară de forma $x_p(t) = A\mathrm{e}^t \cos 2t$. Derivând şi înlocuind în ecuaţia diferenţială, obţinem:

$$A(-4\mathrm{e}^t \sin 2t - 3\mathrm{e}^t \cos 2t) - 3A(\mathrm{e}^t \cos 2t - 2\mathrm{e}^t \sin 2t) - 4A\mathrm{e}^t \cos 2t = 13\mathrm{e}^t \sin 2t$$

Simplificând şi grupând, rezultă

$$(-10A)\cos 2t + (2A - 13)\sin(2t) = 0$$

Dar funcţiile $\cos 2t, \sin 2t$ sunt liniar independente, deci $-10A = 0, 2A - 13 = 0$, sistem incompatibil. Deci nu are sens să cătăm o soluţie x_p care să depindă doar de $\cos 2t$. Conform Observaţiei 3.18.3, vom căuta soluţia de forma $x_p(t) = A\mathrm{e}^t \cos 2t + B\mathrm{e}^t \sin 2t$. Avem

$$\begin{aligned} x_p'(t) &= (A + 2B)\mathrm{e}^t \cos 2t + (-2A + B)\mathrm{e}^t \sin 2t \\ x_p''(t) &= (-3A + 4B)\mathrm{e}^t \cos 2t + (-4A - 3B)\mathrm{e}^t \sin 2t \end{aligned}$$

deci

$$\big((-3A + 4B)\mathrm{e}^t \cos 2t + (-4A - 3B)\mathrm{e}^t \sin 2t\big) - 3\big((A + 2B)\mathrm{e}^t \cos 2t + (-2A + B)\mathrm{e}^t \sin 2t\big) - 4(A\mathrm{e}^t \cos 2t + B\mathrm{e}^t \sin 2t) = 13\mathrm{e}^t \cos 2t$$

Obţinem

$$10A + 2B = 13\,,\ 2A - 10B = 0$$

cu soluţia $A = \frac{5}{4}, B = \frac{1}{4}$. Deci o soluţie particulară a ecuaţiei diferenţiale date este $x_p(t) = \frac{5}{4}\mathrm{e}^t \cos 2t + \frac{1}{4}\mathrm{e}^t \sin 2t$.

3) **Principiul superpoziţiei soluţiilor**. Să analizăm acum ecuaţia diferenţială

$$x'' - 3x' - 4x = 3\mathrm{e}^{2t} + 13\mathrm{e}^t \cos 2t + 5\mathrm{e}^{-t}$$

Termenul liber $b(t) = 3\mathrm{e}^{2t} + 13\mathrm{e}^t \cos 2t + 5\mathrm{e}^{-t}$ nu corespunde nici unuia dintre cazurile menţionate în Observaţia 3.18; mai degrabă, este o *combinaţie liniară* a acestor cazuri, $b(t) = b_1(t) + b_2(t) + b_3(t)$, unde $b_1(t) = 3\mathrm{e}^{2t}$, $b_2(t) = 13\mathrm{e}^t \cos 2t$, $b_3(t) = 5\mathrm{e}^{-t}$. Să observăm atunci următoarele:

Dacă x_i este o soluţie (particulară) a ecuaţiei diferenţiale

$$x'' - 3x' - 4x = b_i(t)$$

pentru $i \in \{1,2,3\}$, atunci $x(t) = x_1(t) + x_2(t) + x_3(t)$ este soluţie a ecuaţiei diferenţiale

$$x'' - 3x' - 4x = b_1(t) + b_2(t) + b_3(t)$$

În consecinţă, vom considera separat cel trei ecuaţii diferenţiale neomogene.

Pentru $x'' - 3x' - 4x = b_1(t) = 3\mathsf{e}^{2t}$, am determinat o soluţie la punctul 1), şi anume $x_1(t) = -\frac{1}{2}\mathsf{e}^{2t}$. De asemenea, la punctul 2) am găsit că o soluţie a a ecuaţiei diferenţiale $x''-3x'-4x = b_2(t) = 13\mathsf{e}^t\cos 2t$ este $x_2(t) = \frac{5}{4}\mathsf{e}^t\cos 2t + \frac{1}{4}\mathsf{e}^t\sin 2t$.

Rămâne să determinăm o soluţie pentru a treia ecuaţie $x'' - 3x' - 4x = b_3(t) = 5\mathsf{e}^{-t}$. Dacă alegem $x_3(t) = A\mathsf{e}^{-t}$ (aceeaşi formă ca şi termenul liber), derivând şi înlocuind în ecuaţia diferenţială obţinem

$$A\mathsf{e}^{-t} + 3A\mathsf{e}^{-t} - 4A\mathsf{e}^{-t} = 5\mathsf{e}^{-t}$$

relaţie imposibilă. Aceasta pentru că $\lambda = -1$ (coeficientul lui t de la exponent din $b_3(t) = 5\mathsf{e}^{-t}$) este rădăcină cu multiplicitatea $s = 1$ a ecuaţiei caracteristice $\lambda^2-3\lambda-4 = 0$ asociată ecuaţiei diferenţiale. Conform Observaţiei 3.18.2, vom căuta atunci o soluţie de forma $t^sA\mathsf{e}^{-t} = At\mathsf{e}^{-t}$. Derivăm şi înlocuim în ecuaţia diferenţială:

$$(-2A + At)\mathsf{e}^{-t} - 3(A - At)\mathsf{e}^{-t} - 4At\mathsf{e}^{-t} = 5\mathsf{e}^{-t}$$

de unde obţinem $A = -1$. Deci o soluţie a ecuaţiei diferenţiale $x'' - 3x' - 4x = 5\mathsf{e}^{-t}$ este $x_3(t) = -t\mathsf{e}^{-t}$.

Coroborând Propoziţia 3.14 cu observaţia de mai sus, rezultă că $x_1(t)+x_2(t)+x_3(t) = -\frac{1}{2}\mathsf{e}^{2t} + \frac{5}{4}\mathsf{e}^t\cos 2t + \frac{1}{4}\mathsf{e}^t\sin 2t - t\mathsf{e}^{-t}$ este o soluţie (particulară) a ecuaţiei diferenţiale $x'' - 3x' - 4x = 3\mathsf{e}^{2t} + 13\mathsf{e}^t\cos 2t + 5\mathsf{e}^{-t}$, iar soluţia generală va fi

$$\underbrace{C_1\mathsf{e}^{4t} + C_2\mathsf{e}^{-t}}_{x_o(t)} + \underbrace{\left(-\frac{1}{2}\mathsf{e}^{2t} + \frac{5}{4}\mathsf{e}^t\cos 2t + \frac{1}{4}\mathsf{e}^t\sin 2t - t\mathsf{e}^{-t}\right)}_{x_1(t)+x_2(t)+x_3(t)}$$

Exerciţiu 3.20. Rezolvaţi exerciţiile de mai sus şi cu metoda variaţiei constantelor şi comparaţi cele două metode.

3.6 Rezolvarea sistemelor de ecuaţii diferenţiale cu coeficienţi constanţi, prin metoda derivării şi eliminării

Prin această metodă un sistem de n ecuatii diferentiale va fi redus la o singura ecuaţie diferenţială de ordinul n, a cărei ecuaţie caracteristică coincide cu ecuaţia asociată polinomului caracteristic al matricei sistemului. Este utilă mai ales când matricea sistemului nu se diagonalizează. Vom ilustra acesta metodă prin exemple:

Exemplu 3.21. 1) Să determinăm soluţia următoarei probleme Cauchy cu metoda derivării şi eliminării:

$$\begin{cases} x' = 2x + y - 2\mathsf{e}^t & x(0) = 0 \\ y' = -x + 4y - 2\mathsf{e}^t & y(0) = 1 \end{cases}$$

Alegem una dintre ecuaţiile diferenţale ale sistemului, de exemplu prima, şi o derivăm:

$$x'' = 2x' + y' - 2\mathrm{e}^t$$

Înlocuim y' din a doua ecuaţie a sistemului:

$$x'' = 2x' + (-x + 4y - 2\mathrm{e}^t) - 2\mathrm{e}^t = 2x' - x + 4y - 4\mathrm{e}^t$$

În sfârşit, din prima ecuaţie a sistemului obţinem $y = x' - 2x + 2\mathrm{e}^t$, de unde

$$x'' = 2x' - x + 4x' - 8x + 4\mathrm{e}^t$$

Deci funcţia necunoscută $x(t)$ satisface ecuaţia diferenţială

$$x'' - 6x' + 9x = 4\mathrm{e}^t$$

Să remarcăm că polinomul caracteristic al matricei sistemului $A = \begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix}$ este $P_A(X) = X^2 - 6X + 9$, în concordanţă cu Observaţia 3.3.

Scriem acum ecuaţia caracteristică

$$\lambda^2 - 6\lambda + 9 = 0$$

cu rădăcina dublă $\lambda_{1,2} = 3$. Rezultă că soluţia generală a ecuaţiei diferenţiale omogene $x'' - 6x' + 9x = 0$ este $x_o(t) = C_1\mathrm{e}^{3t} + C_2\mathrm{e}^{3t}$. Pentru a rezolva ecuaţia neomogenă, aplicăm metoda coeficienţilor nedeterminaţi. Cum termenul liber este $b(t) = 4\mathrm{e}^t$ şi $\lambda = 1$ nu este rădăcină a ecuaţiei caracteristice, căutăm o soluţie particulară de forma $x_p(t) = a\mathrm{e}^t$, cu $a \in \mathbb{R}$. Derivând şi înlocuind în ecuaţie, obţinem $a = 1$, deci $x_p(t) = \mathrm{e}^t$ şi soluţia generală a ecuaţiei diferenţiale va fi $x(t) = x_o(t) + x_p(t) = C_1\mathrm{e}^{3t} + C_2 t\mathrm{e}^{3t} + \mathrm{e}^t$. Revenind la sistemul iniţial, să observăm că putem determina a doua funcţie necunoscută direct din prima ecuaţie a sistemului, şi anume $y(t) = x'(t) - 2x(t) + 2\mathrm{e}^t = (C_1 + C_2)\mathrm{e}^{3t} + C_2 t\mathrm{e}^{3t} + \mathrm{e}^t$.

Din condiţiile iniţiale $x(0) = 0, y(0) = 1$ rezultă $C_1 = -1, C_2 = 1$, deci soluţia problemei Cauchy este

$$\begin{cases} x(t) = -\mathrm{e}^{3t} + t\mathrm{e}^{3t} + \mathrm{e}^t \\ y(t) = t\mathrm{e}^{3t} + \mathrm{e}^t \end{cases}$$

2) Vom proceda analog şi pentru următoarea problemă Cauchy:

$$\begin{cases} x' = x + 2y + 2t & x(0) = 1 \\ y' = -x + 3y + 3t - 1 & y(0) = 0 \end{cases}$$

Alegem una dintre ecuaţiile diferenţale ale sistemului, de exemplu prima, şi o derivăm:

$$x'' = x' + 2y' + 2$$

Înlocuind y' din a doua ecuaţie a sistemului, rezultă $x'' = x' + 2(-x + 3y + 3t - 1) + 2 = x' - 2x + 6y + 6t$, şi cum $y = \frac{1}{2}(x' - x - 2t)$, obţinem $x'' = x' - 2x + 3(x' - x - 2t) + 6t = 4x' - 5x$. Deci funcţia necunoscută $x(t)$ satisface ecuaţia diferenţială

$$x'' - 4x' + 5x = 0$$

cu ecuaţia caracteristică asociată $\lambda^2 - 4\lambda + 5 = 0$. Rădăcinile sunt $\lambda_{1,2} = 2 \pm \mathsf{i}$. Rezultă $x(t) = C_1 \mathsf{e}^{2t} \cos t + C_2 \mathsf{e}^{2t} \sin t$. Revenind la sistemul de ecuaţii diferenţiale, obţinem şi $y(t)$ din prima ecuaţie:

$$\begin{aligned} y(t) &= \frac{1}{2}(x' - x - 2t) \\ &= \frac{1}{2}(2C_1 \mathsf{e}^{2t} \cos t - C_1 \mathsf{e}^{2t} \sin t + 2C_2 \mathsf{e}^{2t} \sin t + C_2 \mathsf{e}^{2t} \cos t - C_1 \mathsf{e}^{2t} \cos t - C_2 \mathsf{e}^{2t} \sin t) - t \\ &= \frac{C_1 + C_2}{2} \mathsf{e}^{2t} \cos t + \frac{-C_1 + C_2}{2} \mathsf{e}^{2t} \sin t - t \end{aligned}$$

Din condiţiile iniţiale $x(0) = 0, y(0) = 1$ rezultă $C_1 = 1, C_2 = -1$, deci soluţia problemei Cauchy este

$$\begin{cases} x(t) = \mathsf{e}^{2t}(\cos t - \sin t) \\ y(t) = -\mathsf{e}^{2t} \sin t - t \end{cases}$$

Exerciţiu 3.22. Rezolvaţi sistemele precedente de ecuaţii diferenţiale cu ajutorul valorilor proprii şi al vectorilor proprii asociaţi matricelor sistemelor şi verificaţi soluţiile determinate mai sus.

Bibliografie

[1] St. Boyd, L. Vandenberghe. *Introduction to applied linear algebra*. Cambridge University Press (2018)

[2] P. R. Halmos. *Finite-dimensional vector spaces*. 2nd ed., Dover Publications (2017)

[3] W.O. Kermack, A.G. McKendrick. *Contributions to the mathematical theory of epidemics*. Proc. R. Soc. Lond. A, 115:700–721, 1927. `https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1927.0118`

[4] S. Lang. *Linear algebra*. 3rd ed., Springer-Verlag (1987)

[5] K. Rosen. *Discrete mathematics and its applications*. 8th ed., McGraw Hill (2019)

[6] O. Stănăşilă. *Analiza matematică a semnalelor şi undinelor*. Matrix Rom (1997)

[7] O. Stănăşilă. *Matematici speciale, ecuaţii diferenţiale şi analiză complexă*, vol II. ALL Educational (2001)

[8] G. Strang, *Linear algebra and its applications*. 4th ed., Brooks/Cole/Cengage (2006)