

Cursul 3

Spații vectoriale

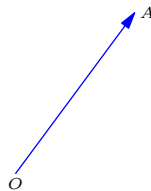
3.1 Motivația introducerii spațiului vectorial

Array-urile (tablourile) constituie structura de date de bază în *Computer Science*. Din punct de vedere matematic un tablou este o matrice de diverse dimensiuni. Tablourile cu o singură linie se mai numesc tablouri 1D sau vectori. Un vector $v = [x_0 \ x_1 \ \dots \ x_{n-1}]$, este deci un set ordonat de n numere de același tip (întreg, real).

În C/C++ un vector, respectiv o matrice se declară astfel:

```
double v[100];  
double A[20][35];
```

Această definiție a vectorului pare să nu aibă legătură cu vectorii din fizică definiți ”ca o mărime caracterizată prin direcție, sens și modul” și vizualizată printr-un segment orientat (adică segment ce unește două puncte O, A , și o săgeată ce indică orientarea). Un astfel de vector este notat \overrightarrow{OA} .



În cele ce urmează vom evidenția în ce fel operațiile și proprietățile vectorilor folosiți în fizică pentru a reprezenta forța, viteza, accelerația, au influențat definirea vectorilor abstracți, definiți ca seturi de n numere reale.

Vectorii priviți ca tablouri 1D au numeroase aplicații în *Computer Graphics*, *Gaming*, *Robotică*. Ei se folosesc pentru modelarea obiectelor 2D și 3D, vizualizarea și manipularea acestora, în generarea mișcării roboților, a creaturilor din jocuri. Imaginile digitale, care fizic sunt matrici de pixeli, se stochează și se transmit pe canale de comunicație ca vectori (seturi de numere reale ce codifică culoarea pixelilor). Matricile interpretate ca vectori intervin în compresia datelor, a imaginilor, a informației audio și video.

Volumul imens de date ce se generează pe WEB sau în diverse domenii de cercetare (fizica particulelor, fuziune nucleară, meteorologie), în domeniul financiar (bănci), în biologie/medicină, se înregistrează ca seturi de numere reale. Vectorii generați în acest fel sunt apoi analizați prin

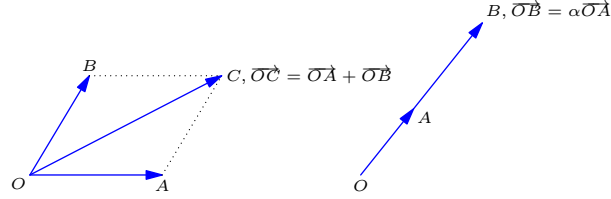


Fig.3.1: Regula paralelogramului de adunare a vectorilor cu același punct de aplicație (stânga) și produsul unui vector cu un scalar (dreapta)

diverse proceduri, tehnici și algoritmi din inteligența artificială, care extrag informație din datele pe care aceștia le conțin.

Operațiile elementare ce se definesc între acești vectori abstracți sunt definite în așa fel încât să aibă aceleași proprietăți ca și vectorii din fizică ce au același punct de aplicație.

Reamintim că pe mulțimea vectorilor cu același punct de aplicație, O , se definește operația de adunare a doi vectori, prin regula paralelogramului și respectiv produsul cu un scalar (adică număr real) al unui astfel de vector (Fig.3.1). Exploataând proprietățile celor două operații cu vectori având același punct de aplicație, s-a introdus noțiunea de spațiu vectorial peste un corp, care va permite ulterior ca vectorii reprezentați prin segmente orientate, să poată fi reprezentați și ei prin coloane numerice, respectiv printr-o pereche de numere, dacă sunt vectori într-un plan și respectiv de un triplet de numere, dacă sunt vectori în spațiul fizic tri-dimensional.

3.2 Definiția spațiului vectorial. Exemple

Considerăm un corp comutativ \mathbb{K} în care notăm cu $+$ operația de adunare, cu \cdot operația multiplicativă și cu 0 , respectiv 1 , elementul neutru al grupului $(\mathbb{K}, +)$, respectiv elementul unitate al grupului $(\mathbb{K} \setminus \{0\}, \cdot)$. Preponderent corpul \mathbb{K} va fi corpul numerelor reale, \mathbb{R} , corpul \mathbb{Z}_2 al claselor de resturi modulo 2 și eventual, corpul numerelor complexe, \mathbb{C} .

Precizăm că adunarea modulo 2 din \mathbb{Z}_2 se numește în computer science `exclusive or` și se notează XOR (derivat din `eXclusive OR`). În C și C++ operația pe biți XOR se notează \wedge , iar în criptografie \oplus .

Definiția 3.2.1 Fie V o mulțime nevidă pe care definim o operație de adunare, care asociază fiecărei perechi $(v, w) \in V \times V$ un element din V , notat $v + w$:

$$(v, w) \mapsto v + w$$

și o operație de înmulțire cu elemente din corpul \mathbb{K} , care asociază unui număr $\alpha \in \mathbb{K}$ și unui element $v \in V$ elementul notat $\alpha v \in V$:

$$(\alpha, v) \mapsto \alpha v.$$

V are structură de spațiu vectorial peste corpul \mathbb{K} dacă cele două operații verifică condițiile următoare:

SV1. $(v + w) + u = v + (w + u)$, $\forall v, w, u \in V$ (asociativitatea adunării);

SV2. există un element în V , notat θ , cu proprietatea că $v + \theta = \theta + v = v, \forall v \in V$;

SV3. pentru orice $v \in V$ există un element $v' \in V$ astfel încât $v + v' = v' + v = \theta$;

SV4. $v + w = w + v, \forall v, w \in V$;

SV5. $\alpha(v + w) = \alpha v + \alpha w, \forall \alpha \in \mathbb{K}, \text{ și } \forall v, w \in V$;

SV6. $(\alpha + \beta)v = \alpha v + \beta v, \forall \alpha, \beta \in \mathbb{K}, \forall v \in V$;

SV7. $\alpha(\beta v) = (\alpha\beta)v, \forall \alpha, \beta \in \mathbb{K}, \forall v \in V$;

SV8. $1v = v, \forall v \in V$.

Observații.

- 1. Dacă V are structură de spațiu vectorial peste corpul \mathbb{K} , atunci elementele lui V se numesc *vectori*, iar cele din corpul \mathbb{K} , *scalari*;
- 2. Operația de adunare a vectorilor este o operație internă, deoarece asociază la orice doi vectori, tot un vector;
- 3. Operația a doua, αv , se numește operație de înmulțire a vectorilor cu scalari. Ea este o operație externă, în sensul că unui scalar α (deci un element din "exteriorul" mulțimii de vectori) și unui vector v i se asociază vectorul αv ;
- 4. Condițiile **SV1–SV4** asigură că mulțimea V are o structură de grup față de adunare. Elementul neutru, θ , față de adunarea vectorilor se numește *vectorul nul* al spațiului;
- 5. Simetricul v' al vectorului v , față de adunare, se va nota $-v$ și îl numim *opusul vectorului* v ;
- 6. Condițiile **SV5–SV8** din definiția spațiului vectorial stabilesc legături dintre operațiile pe corpul \mathbb{K} și cele din spațiul V .

În continuare vom nota (indica) prin V/\mathbb{K} , spațiul vectorial peste corpul \mathbb{K} . Dacă $\mathbb{K} = \mathbb{R}$, atunci V/\mathbb{R} se numește *spațiu vectorial real*, iar dacă $\mathbb{K} = \mathbb{C}$, atunci V/\mathbb{C} se numește *spațiu vectorial complex*. Cu unele excepții, noi vom studia spații vectoriale reale. Spațiul vectorial real, fundamental în algebra liniară este spațiul $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$.

Exemplul 1. Spațiul vectorial pe care îl studiem cel mai mult în acest curs este spațiul vectorial \mathbb{R}^n/\mathbb{R} este definit ca mulțimea n -uplurilor de numere reale (adică a seturilor de n numere reale).

$$\mathbb{R}^n = \{v = (x_1, x_2, \dots, x_n)^T \mid x_i \in \mathbb{R}, i = \overline{1, n}\}$$

Simbolul de transpunere din notația $v = (x_1, x_2, \dots, x_n)^T$ ilustrează că cele n elemente se înregistrează într-o matrice linie transpusă, adică într-o matrice coloană:

$$v = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

În limbajele de programare C/C++, Java, Python, un vector v din \mathbb{R}^n sau tablou uni-dimensional de n elemente reale, are elementele indexate astfel $v[0], v[1], \dots, v[n-1]$ și memoria alocată vectorului este adresată pe linie. Interpretarea vectorilor ca matrici coloană, în algebra liniară se datorează însă simplității operațiilor și a demonstrației diverselor relații în care intervin, în descrierea unor algoritmi.

Dacă $n = 2$, \mathbb{R}^2 este mulțimea cuplurilor (perechilor), $(x_1, x_2)^T$, de numere reale, iar pentru $n = 3$, \mathbb{R}^3 este mulțimea tripletelor de numere reale, $(x_1, x_2, x_3)^T$. Deci pentru n arbitrar, folosim termenul de n -uplu.

Operațiile ce înzestrează pe \mathbb{R}^n cu o structură de spațiu vectorial real sunt definite astfel:

- adunarea: pentru orice $x = (x_1, x_2, \dots, x_n)^T, y = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$,

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)^T \in \mathbb{R}^n$$

- înmulțirea cu scalari: pentru orice $\alpha \in \mathbb{R}$ și $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$,

$$\alpha x = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)^T \in \mathbb{R}^n$$

Vectorul nul al spațiului vectorial \mathbb{R}^n/\mathbb{R} este $\theta_n = \underbrace{(0, 0, \dots, 0)^T}_n$. În particular vectorul nul din \mathbb{R}^2/\mathbb{R} este $\theta_2 = (0, 0)^T$, iar din \mathbb{R}^3 , $\theta_3 = (0, 0, 0)^T$.

Exemplul 2. Spațiul vectorial \mathbb{C}^n peste corpul numerelor complexe se definește analog cu spațiul \mathbb{R}^n/\mathbb{R} . Și anume, \mathbb{C}^n este mulțimea n -uplurilor de numere complexe, $z = (z_1, z_2, \dots, z_n)^T$, $z_i \in \mathbb{C}$, $i = \overline{1, n}$. De exemplu, cvadruplul $z = (-1 + 2i, 3, -5i, 1/2 + i\sqrt{3}/2)^T$ este un vector din \mathbb{C}^4 . Deși pare mai abstract la prima vedere, spațiul vectorial \mathbb{C}^n/\mathbb{C} are aplicații în domenii practice ca de exemplu, studiul semnalelor și imaginilor analogice și digitale.

Exemplul 3. Spațiul vectorial \mathbb{Z}_2^n peste corpul \mathbb{Z}_2 . \mathbb{Z}_2^n este mulțimea n -uplurilor de elemente din $\mathbb{Z}_2 = \{0, 1\}$, numite adesea stringuri de n biți sau vectori binari:

$$\mathbb{Z}_2^n = \{b = (b_1, b_2, \dots, b_n)^T \mid b_i \in \mathbb{Z}_2, i = \overline{0, n}\}$$

De exemplu $b = (0, 1, 1, 0, 1, 0, 0)^T$ este un string de biți din \mathbb{Z}_2^7 .

Precizăm că spre deosebire de \mathbb{R}^n , \mathbb{C}^n , care conțin o infinitate de vectori, \mathbb{Z}_2^n conține doar 2^n vectori, deci este un spațiu vectorial cu un număr finit de vectori.

Suma a două stringuri de biți, $b = (b_1, b_2, \dots, b_n)^T, c = (c_1, c_2, \dots, c_n)^T$ este stringul

$$d = b + c = (b_1 \oplus c_1, b_2 \oplus c_2, \dots, b_n \oplus c_n)^T$$

unde \oplus este adunarea modulo 2.

Produsul dintre un "scalar" $\alpha \in \mathbb{Z}_2$ și un string de biți $b = (b_1, b_2, \dots, b_n)^T$ este definit astfel:

$$\alpha b = (\alpha \cdot b_1, \alpha \cdot b_2, \dots, \alpha \cdot b_n)^T$$

unde înmulțirea $\alpha \cdot b_i$ este înmulțirea modulo 2 dată în tabelul:

·	0	1
0	0	0
1	0	1

Exemplul 4. Mulțimea matricilor de tip $m \times n$ cu elemente din corpul $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{Z}_2$,

$$\mathbb{K}^{m \times n} = \{A = (a_{ij}) | a_{ij} \in \mathbb{K}, i = \overline{1, m}, j = \overline{1, n}\}$$

are structură de spațiu vectorial peste corpul \mathbb{K} , relativ la operația uzuală de adunare a matricilor, respectiv înmulțirea cu un scalar (atenție la matricile binare, cu elemente din \mathbb{Z}_2 , unde adunarea și înmulțirea este modulo 2!). Vectorul nul într-un spațiu vectorial de matrici este matricea nulă, având toate elementele egale cu 0.

Există limbaje de programare numite *array programming languages*, care implementează direct operațiile cu vectori și matrici.

Limbajul C/C++ este un limbaj de programare scalar pentru că adunarea, de exemplu, a doi vectori din \mathbb{R}^n se efectuează adunând iterativ elementele din aceeași poziție a vectorilor.

Python este limbaj vectorizat, adică operațiile cu *array*-uri nu se efectuează element cu element, ci global. Mai precis dacă v și w au fost declarați ca *array*-uri și inițializați, atunci pentru a efectua suma lor, u , se scrie simplu:

$u = v + w$

În ultimii ani au crescut în popularitate și arhitecturile vectoriale pt hardware. Chiar și laptop-urile conțin unități vectoriale (reduse), destinate aplicațiilor multimedia.

În calculul paralel (http://en.wikipedia.org/wiki/Parallel_computing) se realizează de asemenea vectorizarea, adică se convertește implementarea scalară în una vectorială.

Exemplul 5. Spațiul vectorial real al funcțiilor definite pe o mulțime D , cu valori reale. În acest caz V este definit astfel:

$$V = \mathcal{F}(D) = \{f : D \rightarrow \mathbb{R}\}$$

- suma a două funcții din $\mathcal{F}(D)$ se definește punctual, și anume dacă $f, g \in \mathcal{F}(D)$, atunci $f + g$ este funcția definită pe D prin $(f + g)(x) = f(x) + g(x), \forall x \in D$;

- produsul unei funcții $f \in \mathcal{F}(D)$ cu un scalar $\alpha \in \mathbb{R}$, este funcția definită pe D prin $(\alpha f)(x) = \alpha f(x), \forall x \in D$.

În Fig. 3.2 ilustrăm aceste operații în cazul în care D este un interval $[a, b] \subset \mathbb{R}$.

Vectorul nul în spațiul funcțiilor $\mathcal{F}(D)$ este funcția $\theta : D \rightarrow \mathbb{R}$, constantă și egală cu zero: $\theta(x) = 0, \forall x \in D$.

Exemplul 6. În electronică se lucrează preponderent cu spațiile vectoriale

$$\mathcal{F}(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R}\}, \quad \mathcal{S} = \{f : \mathbb{R} \rightarrow \mathbb{C}\}$$

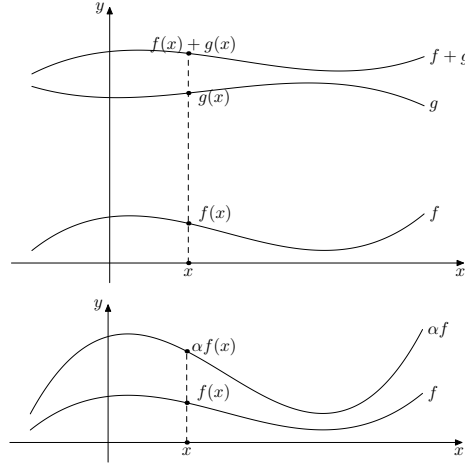


Fig.3.2: Ilustrarea grafică a sumei a două funcții, respectiv a produsului unei funcții cu un scalar.

Primul este spațiul vectorial real al semnalelor în timp continuu, cu valori reale, iar al doilea este spațiul vectorial complex al semnalelor în timp continuu, cu valori complexe.

Exemplul 7. Mulțimea vectorilor cu același punct de aplicație O din spațiul fizic S :

$$V = \{\overrightarrow{OA}, A \in S\}$$

este spațiu vectorial real, relativ la operațiile definite la începutul cursului.

Operații particulare într-un spațiu vectorial. Într-un spațiu vectorial V/\mathbb{K} în care vectorul nul este θ , iar 0 este scalarul zero și 1 unitatea din corpul \mathbb{K} , avem următoarele proprietăți:

- $0v = \theta, \forall v \in V$ și $\alpha\theta = \theta, \forall \alpha \in \mathbb{K}$;
- $(-1)v = -v, \forall v \in V$ (produsul scalarului -1 cu vectorul v este opusul lui v);
- dacă $\alpha v = \theta$, atunci $\alpha = 0$ sau $v = \theta$.

Aceste proprietăți rezultă combinând convenabil condiții din definiția spațiului vectorial.

3.3 Dependență și independență liniară

Fie V/\mathbb{K} un spațiu vectorial peste corpul \mathbb{K} , v_1, v_2, \dots, v_m un număr finit de vectori din V și $\alpha_1, \alpha_2, \dots, \alpha_m, m$ scalari din \mathbb{K} . Cum produsul unui scalar cu un vector este vector și suma unor vectori este vector, rezultă că $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m$ este un vector u . Spunem că vectorul

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m, \quad (3.1)$$

este o **combinație liniară** a vectorilor v_1, v_2, \dots, v_m . Cu alte cuvinte, vectorul u depinde liniar de vectorii v_1, v_2, \dots, v_m .

Având o familie finită de vectori $\{v_1, v_2, \dots, v_m\}$ ne întrebăm în ce condiții un vector v_i din familie depinde liniar de ceilalți, adică se poate exprima ca o combinație liniară a celorlalți vectori, și în ce condiții v_i este independent de ceilalți vectori.

Definiția 3.3.1 Vectorii v_1, v_2, \dots, v_m din spațiul vectorial V/\mathbb{K} sunt vectori liniar dependenți dacă există n scalari nu toți nuli, $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{K}$, astfel încât:

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = \theta \quad (3.2)$$

Această definiție pare a nu avea legătură cu dependența ilustrată mai sus. Analizând însă condițiile din Definiția 3.3.1, rezultă că dacă scalarii $\alpha_1, \alpha_2, \dots, \alpha_m$ nu sunt toți nuli, atunci cel puțin unul din ei este nenul. Dacă, de exemplu, α_1 este nenul, atunci există $\alpha_1^{-1} := 1/\alpha_1$. Înmulțind relația (3.2) cu $1/\alpha_1$ obținem:

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \frac{\alpha_3}{\alpha_1} v_3 - \dots - \frac{\alpha_m}{\alpha_1} v_m, \quad (3.3)$$

adică v_1 se exprimă ca o combinație liniară a vectorilor v_2, v_3, \dots, v_m .

Remarcăm că oricare ar fi vectorii v_1, v_2, \dots, v_m , combinația lor liniară cu coeficienții 0 este vectorul nul:

$$0v_1 + 0v_2 + \dots + 0v_m = \theta$$

Atunci când se studiază dependența unui sistem de vectori v_1, v_2, \dots, v_m , practic se determină dacă o combinație liniară a lor dă vectorul nul doar pentru toți scalarii zero sau și pentru scalari nu toți nuli.

Definiția 3.3.2 Vectorii v_1, v_2, \dots, v_m din spațiul vectorial V/\mathbb{K} sunt vectori liniar independenți dacă o combinație liniară a lor poate fi egală cu vectorul nul:

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = \theta, \quad (3.4)$$

doar dacă toți scalarii $\alpha_1, \alpha_2, \dots, \alpha_m$ sunt zero.

Exemplul 8. În spațiul vectorial \mathbb{R}^3/\mathbb{R} se dau vectorii $v_1 = (-1, 2, 3)^T$, $v_2 = (-9, 16, 7)^T$, $v_3 = (3, -5, 1)^T$. Să se verifice dacă acești vectori sunt liniar dependenți sau independenți și în caz de dependență să se determine relația dintre ei.

Rezolvare: Presupunem că există trei scalari $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ astfel încât:

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = \theta \quad (3.5)$$

Evident că relația (3.5) are loc dacă $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Pentru a stabili dacă vectorii dați sunt liniar independenți sau dependenți trebuie să investigăm dacă relația (3.5) poate avea loc doar pentru cei trei scalari, simultan zero sau și pentru un set de scalari $\alpha_1, \alpha_2, \alpha_3$, nu toți nuli.

În acest scop înlocuim în (3.5) fiecare vector prin tripletul reprezentativ și efectuăm operațiile corespunzătoare din \mathbb{R}^3 :

$$\alpha_1 \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} -9 \\ 16 \\ 7 \end{bmatrix} + \alpha_3 \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow \quad (3.6)$$

$$\begin{bmatrix} -1 & -9 & 3 \\ 2 & 16 & -5 \\ 3 & 7 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (3.7)$$

Observăm că problema independenței sau dependenței vectorilor v_1, v_2, v_3 s-a transformat în problema care cere să stabilim dacă sistemul liniar și omogen (3.7) admite numai soluția banală $\alpha_1 = \alpha_2 = \alpha_3 = 0$ sau și soluții nebanale. Dacă sistemul admite doar soluția banală, atunci vectorii sunt liniar independenți, iar dacă admite și soluții nebanale, atunci vectorii sunt liniar dependenți.

Pentru a decide natura soluțiilor calculăm rangul matricii sistemului:

$$A = \begin{bmatrix} -1 & -9 & 3 \\ 2 & 16 & -5 \\ 3 & 7 & 1 \end{bmatrix}$$

Cum determinantul matricii A este $\det(A) = 0$, rezultă că sistemul admite și soluții nebanale, cu alte cuvinte există trei scalari nu toți nuli, $(\alpha_1, \alpha_2, \alpha_3)$, soluție a sistemului omogen, astfel încât pe baza șirului de echivalențe (3.6) avem relația (3.5), adică vectorii sunt liniar dependenți. Pentru a determina relația dintre ei, rezolvăm efectiv sistemul omogen pentru a găsi o soluție nebanală. Rangul matricii sistemului este 2 și un determinant principal este:

$$\Delta = \begin{vmatrix} -1 & -9 \\ 2 & 16 \end{vmatrix} \neq 0,$$

constituit din coeficienții necunoscutelor α_1, α_2 . Prin urmare rezolvăm doar sistemul format din primele două ecuații în raport cu necunoscutele principale α_1, α_2 . Necunoscuta secundară α_3 o notăm cu β :

$$\begin{aligned} -\alpha_1 - 9\alpha_2 &= -3\beta \\ 2\alpha_1 + 16\alpha_2 &= 5\beta \end{aligned}$$

Rezolvând obținem $\alpha_1 = -3\beta/2$, $\alpha_2 = \beta/2$, $\alpha_3 = \beta$, $\beta \in \mathbb{R}$. Deci familia soluțiilor este $(-3\beta/2, \beta/2, \beta)$, $\beta \in \mathbb{R}$. O soluție nebanală obținem pentru $\beta \neq 0$, fixat. Luând $\beta = 2$ avem $\alpha_1 = -3$, $\alpha_2 = 1$, $\alpha_3 = 2$ și deci relația de dependență a celor trei vectori devine:

$$-3v_1 + v_2 + 2v_3 = \theta_3 \Leftrightarrow v_2 = 3v_1 - 2v_3 \Leftrightarrow v_1 = \frac{1}{3}v_2 + \frac{2}{3}v_3$$

Observația 3.3.1 Orice sistem (mulțime) de vectori ce conține vectorul nul este sistem de vectori liniar dependenți.

Într-adevăr, fie v_1, v_2, \dots, v_m un sistem de vectori din spațiul vectorial V/\mathbb{K} . Presupunem că vectorul v_i este vectorul nul θ . Evident că sistemul de n scalari $\alpha_1 = 0, \dots, \alpha_{i-1} = 0, \alpha_i, \alpha_{i+1} = 0, \dots, \alpha_m = 0$ cu $\alpha_i \neq 0$, conduce la:

$$0v_1 + \dots + 0v_{i-1} + \alpha_i\theta + 0v_{i+1} + \dots + 0v_m = \theta,$$

adică vectorii $v_1, v_2, \dots, v_{i-1}, \theta, v_{i+1}, \dots, v_m$ sunt liniar dependenți.

Criteriul practic de determinare a independenței sau dependenței liniare a unui sistem de vectori din \mathbb{K}^n/\mathbb{K} .

Exemplele precedente au ilustrat că problema independenței sau dependenței liniare a unui sistem de vectori din \mathbb{R}^n/\mathbb{R} ($n=3,4$) se reduce la a deduce dacă un sistem de ecuații liniare și omogene cu coeficienți din \mathbb{R} admite doar soluția banală sau și soluții nebanale. În continuare vom arăta că natura soluțiilor unui astfel de sistem de ecuații cu coeficienți în corpul \mathbb{K} ($= \mathbb{R}, \mathbb{C}, \mathbb{Z}_2$) asociat unei mulțimi finite de vectori rezultă din relația dintre rangul matricii sistemului și numărul de vectori.

Fie sistemul de k vectori din \mathbb{K}^n/\mathbb{K} :

$$v_1 = (a_{11}, a_{12}, \dots, a_{1n})^T, v_2 = (a_{21}, a_{22}, \dots, a_{2n})^T, \dots, v_k = (a_{k1}, a_{k2}, \dots, a_{kn})^T$$

Să deducem condițiile în care acești vectori sunt liniar independenți, respectiv liniar dependenți.

Presupunem că:

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \theta_n$$

Detaliat această egalitate se scrie astfel:

$$\alpha_1 \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1n} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2n} \end{bmatrix} + \dots + \alpha_k \begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{kn} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Membrul stâng al relației înșă reprezintă produsul dintre matricea $A = [v_1 | v_2 | \dots | v_k]$ și matricea coloană de elemente $\alpha_1, \alpha_2, \dots, \alpha_k$ (vezi Cursul 1.). Astfel combinația liniară a vectorilor, egală cu vectorul nul, este echivalentă cu :

$$\begin{bmatrix} a_{11} & a_{21} & \dots & a_{k1} \\ a_{12} & a_{22} & \dots & a_{k2} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{kn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

adică cu un sistem liniar și omogen de n ecuații cu k necunoscute.

Dacă acest sistem admite doar soluția banală atunci vectorii v_1, v_2, \dots, v_k sunt liniar independenți, iar dacă admite și soluții nebanale, atunci vectorii sunt liniar dependenți. Tipul

soluției sistemului liniar și omogen depinde de rangul matricii A a sistemului care are drept coloane n -uplurile ce definesc cei k vectori, $A = [v_1 | v_2 | \dots | v_k]$

Având sistemul liniar și omogen $A\alpha = 0$, unde $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$, matricea prelungită este:

$$\overline{A} = \left[\begin{array}{c|c} A & \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \right]$$

Prin transformări elementare pe linie, coloana ultimă de zerouri rămâne tot coloană de zerouri astfel că forma scară a matricii prelungite este:

$$S_{\overline{A}} = \left[\begin{array}{c|c} S_A & \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \end{array} \right]$$

Prin urmare sistemul liniar și omogen, $A\alpha = 0$, asociat setului de vectori v_1, v_2, \dots, v_k din \mathbb{R}^n este echivalent cu sistemul liniar și omogen $S_A\alpha = 0$ sau $S_A^0\alpha = 0$.

Astfel deducem dacă vectorii sunt liniar independenți sau dependenți analizând natura soluțiilor sistemului $S_A\alpha = 0$, adică dacă sistemul admite doar soluția banală sau și soluții nebanale.

Pentru aceasta reamintim de la liceu că după ce se stabilește rangul r al matricii unui sistem omogen $A\alpha = 0$, atunci se găsește un determinant principal Δ , de ordinul rangului. Dacă în determinantul Δ intră elemente din coloanele j_1, j_2, \dots, j_r ale matricii A , atunci necunoscutele $\alpha_{j_1}, \alpha_{j_2}, \dots, \alpha_{j_r}$ sunt necunoscute principale, iar restul până la k sunt necunoscute secundare.

În cazul sistemului echivalent $S_A\alpha = 0$, rangul este dat de numărul de pivoți, iar determinantul principal, Δ , conține elementele de intersecție ale primelor r linii, cu coloanele j_1, j_2, \dots, j_r , în care se găsesc pivoții.

Considerăm două cazuri:

1. Dacă $k > n$, în matricea A numărul de coloane este mai mare decât numărul de linii. și astfel A poate avea rangul cel mult n , adică $r \leq n < k$.

În acest caz forma scară (redușă) a lui A conține cel mult n coloane cu pivoți ($r \leq n < k$). Deci în rezolvarea sistemului omogen $S_A\alpha = 0$, vom avea r necunoscute principale și $k - r$ necunoscute secundare. Dar existența necunoscutelor secundare implică faptul că sistemul admite și soluții nebanale. Rezultă astfel că în acest caz, vectorii v_1, v_2, \dots, v_k sunt liniar dependenți.

De exemplu, 5 vectori v_1, v_2, \dots, v_5 din \mathbb{R}^3 sunt sigur liniar dependenți, deoarece $k = 5 > 3 = n$.

2. Dacă $k \leq n$ atunci rangul matricii $A = [v_1 | v_2 | \dots | v_k]$ poate fi cel mult k .

- când $\text{rang}(A) = k$ forma scară redusă S_A^0 conține pivoți pe toate cele k coloane și deci

$$S_A^0 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

În acest caz sistemul echivalent $S_A^0 \alpha = 0$ admite doar soluția banală $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ și deci cei k vectori sunt liniar independenți.

- Dacă rangul matricii este mai mic decât k , atunci matricea sistemului $S_A^0 \alpha = 0$ are $r < k$ pivoți, deci pentru rezolvarea lui se aleg r necunoscute principale și $k - r$ necunoscute secundare. Prin urmare sistemul are și soluții nebanale și în concluzie vectorii sunt liniar dependenți.

Sintetizând, rezultă că:

Dacă rangul matricii sistemului omogen $A\alpha = 0$ este egal cu numărul de coloane ale matricii A , atunci sistemul admite doar soluția banală, iar dacă rangul este strict mai mic decât numărul de coloane ale lui A , atunci sistemul admite și soluții nebanale.

Avem astfel următoarea regulă practică de determinare a dependenței sau independenței liniare a k vectori din \mathbb{R}^n , numită în continuare **Criteriul practic de determinare a dependenței sau independenței unui sistem de vectori**:

Vectorilor $v_1, v_2, \dots, v_k \in \mathbb{K}^n$ li se asociază matricea $A = [v_1 | v_2 | \dots | v_k]$, ce are drept coloane, n -uplurile ce definesc vectorii:

$$A = \begin{array}{cccc} v_1 & v_2 & & v_k \\ \downarrow & \downarrow & & \downarrow \\ \begin{bmatrix} a_{11} & a_{21} & \dots & a_{k1} \\ a_{12} & a_{22} & \dots & a_{k1} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{kn} \end{bmatrix} \end{array}$$

- Dacă rangul matricii A este egal cu numărul de vectori (adică cu numărul de coloane din A), atunci vectorii sunt liniar independenți.
- Dacă rangul matricii este diferit de numărul de vectori, atunci aceștia sunt liniar dependenți.

O problemă de bază în *machine learning* este aceea de a identifica dintr-o matrice de date, $A = [v_1|v_2|\dots|v_k]$, care coloane conțin informația relevantă și care conțin informație redundantă (în plus).

O modalitate de a găsi răspunsul la această întrebare este să identificăm care subset din cele k coloane este format din coloane independente.

Pentru aceasta precizăm că se poate demonstra că:

Prin transformări elementare pe linie aplicate unei matrici $A = [v_1|v_2|\dots|v_k]$, eventualele relații liniare între coloane matricii nu se schimbă. Adică dacă anumite coloane din A formează un sistem independent de vectori, exact aceleași coloane din forma scară redusă sunt independente și reciproc. Dacă unele coloane din A se exprimă ca o combinație liniară a altor coloane, atunci exact aceleași particularitate o au și coloanele corespunzătoare din S_A^0 și reciproc.

Astfel în loc să analizăm rangul matricii inițiale, A , asociate unui sistem de vectori $v_1, v_2, \dots, v_k \in \mathbb{K}^n$, deducem rangul r din forma scară redusă S_A^0 .

Coloanele ce conțin pivoți din forma scară redusă sunt coloanele $j_1 < j_2 < \dots < j_r$. Vectorii din aceste coloane sunt:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_r = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{linia } r$$

Acești vectori sunt linear independenți pentru că matricea asociată, $M = [e_1|e_2|\dots|e_r]$ are rangul r egal cu numărul de vectori (vezi criteriul practic):

$$M = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Restul coloanelor din forma scară redusă sunt combinații liniare ale acestor coloane. Să ilustrăm mai întâi printr-un exemplu și apoi dăm cazul general. Fie $v_1, v_2, v_3, v_4, v_5 \in \mathbb{R}^4$, cinci vectori din \mathbb{R}^4 , a căror matrice asociată, $A = [v_1|v_2|v_3|v_4|v_5]$, este:

$$A = \begin{bmatrix} 2 & 1 & -2 & 5 & -15 \\ -3 & -2 & 5 & 1 & -4 \\ 1 & 0 & 1 & -4 & 11 \\ 5 & 4 & -11 & 0 & 3 \end{bmatrix}$$

Forma scară redusă a matricii A este:

$$S_A^0 = \begin{bmatrix} \mathbf{1} & 0 & 1 & 0 & -1 \\ 0 & \mathbf{1} & -4 & 0 & 2 \\ 0 & 0 & 0 & \mathbf{1} & -3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Matricea S_A^0 conține 3 pivoți. Deci rangul lui S_A^0 (deci și al matricii A) este $r = 3$, mai mic decât numărul de vectori $k = 5$. Prin urmare vectorii v_1, v_2, \dots, v_5 sunt liniar dependenți. Există oare un subsistem de vectori printre cei 5 care să fie liniar independenți?

Răspuns: Coloane pivoților sunt coloanele 1, 2, 4. Vectorii din aceste coloane sunt:

$$C_1 = e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad C_2 = e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad C_4 = e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Deci coloanele C_1, C_2, C_4 din S_A^0 sunt coloane independente.

Observăm că orice coloană diferită de coloanele 1, 2, 4, se poate exprima ca o combinație liniară a coloanelor 1, 2, 4.

Coloana 3:

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

Coloana 5:

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

adică vectorul din coloana 3 se exprimă ca o combinație liniară a vectorilor din coloanele 1, 2, 4 și coeficienții acestei combinații liniare sunt primele 3 elemente din coloana 3: $-1, -4, 0$. Analog coeficienții din exprimarea coloanei 5 în funcție de coloanele 1, 2, 4 sunt primele 3 elemente din coloana 5: $-1 : 2, -3$.

Analizând acum coloanele 1, 2, 4 din matricea inițială, $A = [v_1|v_2|v_3|v_4|v_5]$, remarcăm că ele sunt liniar independente (deoarece $\text{rang}([v_1|v_2|v_4])=3$) iar coloanele 3 și 5 ale matricii inițiale, adică vectorii v_3, v_5 , se exprimă ca și combinații liniare cu aceleași coeficienți (ca și în cazul matricii S_A^0) ale coloanelor 1, 2, 4, adică ale vectorilor v_1, v_2, v_4 :

$$v_3 = \begin{bmatrix} -2 \\ 5 \\ 1 \\ -11 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ -3 \\ 1 \\ 5 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} = 1v_1 - 4v_2$$

$$v_5 = \begin{bmatrix} -15 \\ -4 \\ 11 \\ 3 \end{bmatrix} = -1 \begin{bmatrix} 2 \\ -3 \\ 1 \\ 5 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 4 \end{bmatrix} - 3 \begin{bmatrix} 5 \\ 1 \\ -4 \\ 0 \end{bmatrix} = -1v_1 + 2v_2 - 3v_3$$

În concluzie, între coloanele matricii A există aceleași relații ca și între coloanele matricii în forma scară redusă, S_A^0 .

Dacă matricea A ar fi o matrice de date analizată în *machine learning*, atunci ar rezulta că informația relevantă este cea din coloanele 1,2,4. Restul coloanelor sunt doar combinații liniare ale acestora.

În concluzie forma scară redusă a unei matrici înglobează o informație în plus față de ce am dedus în cursurile precedente.

1. Rangul r al matricii inițiale A este egal cu numărul pivoților, adică numărul liniilor nenule din forma scară redusă.
2. Dacă pivoții se găsesc în pozițiile $(1, j_1), (2, j_2), \dots, (r, j_r)$, atunci coloanele j_1, j_2, \dots, j_r sunt linear independente (conțin informația de bază) și orice coloană din matricea A diferită de coloanele j_1, j_2, \dots, j_r se exprimă ca o combinație liniară a coloanelor j_1, j_2, \dots, j_r . Această combinație se citește din matricea scară redusă. De exemplu dacă în forma scară redusă o astfel de coloană este:

$$\begin{bmatrix} s_{1j} \\ s_{2j} \\ \vdots \\ s_{kj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, k \leq r$$

atunci ea se exprimă în funcție de coloanele j_1, j_2, \dots, j_k astfel:

$$\begin{bmatrix} s_{1j} \\ s_{2j} \\ \vdots \\ s_{kj} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = s_{1j} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + s_{2j} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + s_{kj} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Exemplul 9. Să se studieze dependența sau independența sistemului de vectori din \mathbb{R}^5 :

$$v_1 = (1, 2, 4, -2, 5)^T, v_2 = (2, 3, 0, 1, -2)^T, v_3 = (4, 5, -8, 7, -16)^T$$

și în caz de dependență sa se determine relația/relațiile dintre vectori.

Matricea asociată $A = [v_1|v_2|v_3]$ este:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 5 \\ 4 & 0 & -8 \\ -2 & 1 & 7 \\ 5 & -2 & -16 \end{bmatrix}$$

Forma scară redusă:

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

Din S_A^0 rezultă că rangul matricii A este 2 deci diferit de numărul de vectori. Prin urmare cei trei vectori sunt liniar dependenți și $v_3 = -2v_1 + 3v_2$, pentru că în S_A^0 avem: coloana 3 = $-2 \times$ coloana 1 + $3 \times$ coloana 2.

3.4 Baze într-un spațiu vectorial

În cursul de algebră liniară studiem spații vectoriale care pot fi ”construite” pornind de la un număr finit de vectori, e_1, e_2, \dots, e_n , liniar independenți. Cu alte cuvinte studiem ca model un spațiu vectorial V/\mathbb{K} ce constă din mulțimea vectorilor v ce se exprimă ca o combinație liniară a vectorilor e_1, e_2, \dots, e_n .

Definiția 3.4.1 Un sistem ordonat de vectori, $\mathcal{B} = (e_1, e_2, \dots, e_n)$, din spațiul vectorial V/\mathbb{K} , formează o bază în V dacă:

B1. Vectorii sunt liniar independenți.

B2. Orice alt vector v din spațiul vectorial V se exprimă ca o combinație liniară a vectorilor e_1, e_2, \dots, e_n :

$$v = x_1e_1 + x_2e_2 + \dots + x_ne_n, \quad x_1, x_2, \dots, x_n \in \mathbb{K} \quad (3.8)$$

Cu alte cuvinte sistemul de vectori, (e_1, e_2, \dots, e_n) , este un sistem maximal de vectori liniar independenți în V , adică orice sistem (v, e_1, \dots, e_n) , cu $v \neq e_i, i = \overline{1, n}$, este liniar dependent.

Propoziția 3.4.1 Având fixată o bază, \mathcal{B} , în spațiul vectorial V/\mathbb{K} , orice vector $v \in V$ se exprimă ca o combinație liniară unică a vectorilor din bază.

Demonstrație: Presupunem că vectorul v admite două exprimări în baza \mathcal{B} :

$$v = x_1e_1 + x_2e_2 + \dots + x_ne_n \quad \text{și} \quad v = y_1e_1 + y_2e_2 + \dots + y_ne_n$$

Deoarece $v - v = \theta$, rezultă scăzând cele două expresii ale vectorului v că:

$$(x_1 - y_1)e_1 + (x_2 - y_2)e_2 + \cdots + (x_n - y_n)e_n = \theta$$

ceea ce înseamnă că o combinație liniară a vectorilor e_1, e_2, \dots, e_n dă vectorul nul. Dar cum vectorii bazei sunt liniar independenți rezultă că:

$$x_1 - y_1 = 0, x_2 - y_2 = 0, \dots, x_n - y_n = 0$$

adică $x_i = y_i, \forall i = \overline{1, n}$, și deci exprimarea vectorului v în baza \mathcal{B} este unică. \square

Scalarii x_1, x_2, \dots, x_n din exprimarea vectorului v în funcție de vectorii bazei \mathcal{B} , se numesc **coordonatele vectorului** v în baza \mathcal{B} .

Numele de bază este sugestiv, deoarece vectorii ei constituie fundamentul, "bază" pe care se construiește întreg spațiul. Cunoscând vectorii bazei, prin combinații liniare ale acestora, se "construiește" orice alt vector din spațiu.

Exemplul 10. Sistemul de vectori $\mathcal{B} = (e_1 = (1, 0, 0)^T, e_2 = (0, 1, 0)^T, e_3 = (0, 0, 1)^T)$ constituie o bază în spațiul vectorial \mathbb{R}^3 .

B1. Să arătăm că vectorii sunt liniar independenți. Matricea asociată: $A = [e_1 | e_2 | e_3]$ este:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Deoarece $\det(A) = 1$, rangul matricii este 3, deci egal cu numărul de vectori. Prin urmare conform criteriului practic, vectorii e_1, e_2, e_3 sunt liniar independenți.

B2. Să arătăm că orice alt vector $v = (x_1, x_2, x_3)^T$ din \mathbb{R}^3 se exprimă ca o combinație liniară a vectorilor e_1, e_2, e_3 :

$$\begin{aligned} v = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \\ &= x_1 e_1 + x_2 e_2 + x_3 e_3 \end{aligned}$$

Rezultă că coordonatele unui vector $v = (x_1, x_2, x_3)^T$ din \mathbb{R}^3 , în baza \mathcal{B} , sunt chiar numerele reale ce definesc tripletul v . De exemplu coordonatele vectorului $v = (-5, 4, 1)^T$ în baza \mathcal{B} sunt $-5, 4, 1$, adică $v = -5e_1 + 4e_2 + 1e_3$. Această bază se numește *baza canonică* sau *baza standard* din \mathbb{R}^3 .

În mod analog:

Exemplul 11. Se arată că în spațiul vectorial \mathbb{R}^n/\mathbb{R} , sistemul de vectori:

$$\mathcal{B} = (e_1 = (1, 0, 0, \dots, 0)^T, \dots, e_i = (0, 0, \dots, \underbrace{1}_i, \dots, 0)^T, \dots, e_n = (0, 0, 0, \dots, 1)^T)$$

constituie o bază. Această bază se numește *baza canonică* din \mathbb{R}^n/R .