Introdução à Álgebra Linear

Profa. Patrícia Hernandes Baptistelli
Departamento de Matemática - UEM

Maringá - 2018

Capítulo 1

Matrizes e Sistemas Lineares

Neste capítulo abordamos um dos mais importantes conceitos da Álgebra Linear: matrizes e sistemas lineares. Destacamos o conceito de escalonamento de matrizes, que nos permite determinar se uma matriz é inversível e, quando isso ocorre, nos dá um método efetivo para o cálculo de sua inversa. Além disso, o processo de escalonamento permite determinar se um sistema linear possui solução ou não.

1.1 Matrizes

Nesta seção, vamos apresentar os diversos tipos de matrizes. Ao longo deste texto, $\mathbb{K}$ denota o conjunto $\mathbb{R}$ ou $\mathbb{C}$.

Definição 1.1. *Sejam $m, n \in \mathbb{Z}$. Uma **matriz** A **de ordem** $m \times n$ (lê-se m por n) é um arranjo de $m.n$ elementos $a_{ij} \in \mathbb{K}$, com $1 \leq i \leq m$ e $1 \leq j \leq n$, dispostos em m linhas e n colunas, representado por*

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}.$$

Denotaremos a matriz acima por $A = (a_{ij})_{m \times n}$ ou simplesmente $A = (a_{ij})$, onde i indica a linha e j indica a coluna em que se localiza a_{ij}.

Usaremos letras maiúsculas para denotar matrizes. Cada número a_{ij} é chamado **elemento** ou **entrada** da matriz. As entradas de uma matriz podem ser números reais, números complexos, funções ou ainda outras matrizes. No decorrer desse curso, cada entrada a_{ij} será um número real ou complexo e denotaremos por $\mathcal{M}_{m \times n}(\mathbb{K})$ o conjunto de todas as matrizes de ordem $m \times n$ com entradas em $\mathbb{K}$.

Exemplo 1.2. $\begin{pmatrix} 1 & 0 \\ -5 & \frac{2}{3} \\ 5i-2 & 10 \end{pmatrix}_{3\times 2}$, $(4)_{1\times 1}$, $\begin{pmatrix} 13 & 3 & -4 & 27 \\ 1 & 1 & 0 & \frac{4}{5} \end{pmatrix}_{2\times 4}$, $(i \quad 0)_{1\times 2}$.

Dizemos que duas matrizes $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{p\times q}$ são iguais se $m = p$, $n = q$ e seus elementos correspondentes são iguais, ou seja, $a_{ij} = b_{ij}$ para todo i, j. Algumas matrizes recebem um nome especial por apresentarem propriedades específicas:

1. **Matriz Nula**: Uma matriz $A = (a_{ij})_{m\times n}$ é dita nula se $a_{ij} = 0$ para todo $1 \le i \le m$ e $1 \le j \le n$.

2. **Matriz Linha**: Uma matriz $A = (a_{ij})_{m\times n}$ é chamada matriz linha se $m = 1$, ou seja, se o número de linhas da matriz é igual a 1.

3. **Matriz Coluna**: Uma matriz $A = (a_{ij})_{m\times n}$ é chamada matriz coluna se $n = 1$, ou seja, se o número de colunas da matriz é igual a 1.

4. **Matriz Quadrada**: Uma matriz $A = (a_{ij})_{m\times n}$ é chamada quadrada se $m = n$, ou seja, se o número de linhas e o número de colunas de A coincidem. Neste caso, dizemos simplesmente que A tem **ordem** n e definimos a **diagonal principal** de A como os elementos da matriz em que $i = j$, ou seja, os elementos a_{ii}.

$$A = \begin{pmatrix} \mathbf{a_{11}} & a_{12} & \dots & a_{1n} \\ a_{21} & \mathbf{a_{22}} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & \mathbf{a_{nn}} \end{pmatrix}_{n\times n}.$$

   A **diagoanal secundária** é formada pelos elementos a_{ij} da matriz tais que $i + j = n + 1$.

5. **Matriz Diagonal**: Uma matriz quadrada $A = (a_{ij})$ é dita diagonal se $a_{ij} = 0$ para $i \ne j$, ou seja, os elementos que não estão na diagonal principal são nulos.

   Exemplos: $\begin{pmatrix} 7i & 0 \\ 0 & \frac{2}{3} \end{pmatrix}_{2\times 2}$, $\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -\frac{16}{250} \end{pmatrix}_{4\times 4}$.

6. **Matriz Identidade**: Uma matriz $A = (a_{ij})$ é chamada identidade se ela for uma matriz diagonal tal que $a_{ii} = 1$ para todo i, ou seja, todos os elementos da diagonal principal são iguais a 1 e os de fora dela são nulos.

Exemplo: $I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{4\times 4}$.

7. **Matriz Triangular**: Uma matriz quadrada $A = (a_{ij})$ é chamada triangular superior (respectivamente, triangular inferior) se todos os elementos abaixo (respectivamente, acima) da diagonal principal forem nulos, ou seja, $a_{ij} = 0$ para todo $i > j$ (respectivamente, $a_{ij} = 0$ para todo $i < j$).

   Exemplos: $\begin{pmatrix} 1 & 3 \\ 0 & 7-2i \end{pmatrix}_{2\times 2}$, $\begin{pmatrix} 15 & 0 & 0 \\ 9 & 2 & 0 \\ \frac{1}{4} & 3 & -1 \end{pmatrix}_{3\times 3}$.

   Note que as únicas matrizes que são simultaneamente triangular superior e triangular inferior são as matrizes diagonais.

8. **Matriz Simétrica**: Uma matriz quadrada $A = (a_{ij})$ é simétrica se $a_{ij} = a_{ji}$ para todo $i \neq j$.

   Exemplo: $\begin{pmatrix} -5i & 9 & 4 \\ 9 & 2 & 0 \\ 4 & 0 & 18 \end{pmatrix}_{3\times 3}$.

9. **Matriz Antissimétrica**: Uma matriz quadrada $A = (a_{ij})$ é dita antissimétrica se $a_{ij} = -a_{ji}$ para todo i, j. Neste caso, os elementos da diagonal principal são nulos, já que $a_{ii} = -a_{ii}$ se, e somente se, $a_{ii} = 0$.

   Exemplo: $\begin{pmatrix} 0 & -91 & 3 \\ 91 & 0 & -1+i \\ -3 & 1-i & 0 \end{pmatrix}_{3\times 3}$.

1.2 Operações com Matrizes

Nesta seção vamos introduzir as operações usuais com matrizes, como soma, multiplicação por escalar e produto, bem como o conceito de matrizes inversível e transposta.

1.2.1 Soma de Matrizes

Sejam $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$ matrizes de mesma ordem. A soma da matriz A com B, denotada por $A + B$, é a matriz cujos elementos são a soma dos elementos correspondentes de A e B, ou seja,

$$A + B := (a_{ij} + b_{ij})_{m\times n}.$$

Exemplo: $\begin{pmatrix} 1 & -1 \\ 3 & 0 \\ 8 & 4 \end{pmatrix}_{3\times 2} + \begin{pmatrix} 3 & 5 \\ i & -2 \\ -1 & 1 \end{pmatrix}_{3\times 2} = \begin{pmatrix} 4 & 4 \\ 3+i & -2 \\ 7 & 5 \end{pmatrix}_{3\times 2}.$

Note, por exemplo, que não é possível realizar a soma das matrizes

$$A = \begin{pmatrix} 1 & -1 \\ 3 & 0 \\ 8 & 4 \end{pmatrix}_{3\times 2} \quad \text{e} \quad B = \begin{pmatrix} 3 & 5 \\ i & -2 \end{pmatrix}_{2\times 2},$$

pois o número de linhas de A é diferente do número de linhas de B.

Proposição 1.3. *Para quaisquer matrizes A, B e C de ordem $m \times n$, temos:*

(i) Comutatividade: $A + B = B + A$;

(ii) Associatividade: $A + (B + C) = (A + B) + C$;

(iii) Elemento neutro: $A + 0_{m\times n} = A$, onde $0_{m\times n}$ denota a matriz nula de ordem $m \times n$. Dizemos que $0_{m\times n}$ é o elemento neutro com respeito à soma de matrizes.

Demonstração: Sejam $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$ e $C = (c_{ij})_{m\times n}$. Então temos:

(i) $A + B = (a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = (b_{ij}) + (a_{ij}) = B + A$, com a terceira igualdade seguindo da comutatividade dos números reais e complexos.

(ii) $A + (B + C) := (a_{ij}) + \big((b_{ij}) + (c_{ij})\big) := (a_{ij}) + (b_{ij} + c_{ij}) := (a_{ij} + b_{ij} + c_{ij}) = \big((a_{ij}) + (b_{ij})\big) + (c_{ij}) = (A + B) + C$, com a terceira igualdade seguindo da associatividade dos números reais e complexos.

(iii) $A + 0_{m\times n} = (a_{ij}) + (0) := (a_{ij} + 0) = (a_{ij}) = A$, uma vez que 0 é o elemento neutro nos números reais e complexos com respeito à soma. ■

1.2.2 Multiplicação de Matriz por Escalar

Sejam $A = (a_{ij})_{m\times n}$ e $k \in \mathbb{K}$ um escalar. Definimos o produto de A pelo escalar k, denotado por kA, como sendo a matriz cujos elementos são o produto de k por cada a_{ij}, ou seja,

$$kA := (ka_{ij})_{m\times n}.$$

Exemplo: $i\begin{pmatrix} 1 & -1 & i & \sqrt{3} & 0 \\ 0 & 0 & 1+5i & \pi & 24 \end{pmatrix}_{2\times 5} = \begin{pmatrix} i & -i & -1 & i\sqrt{3} & 0 \\ 0 & 0 & i-5 & i\pi & 24i \end{pmatrix}_{2\times 5}$.

Proposição 1.4. *Para quaisquer matrizes A e B de ordem $m \times n$ e quaisquer $k, k_1, k_2 \in \mathbb{K}$, valem os seguintes:*

(i) Distributividade com respeito à soma de matrizes: $k(A+B) = kA + kB$;

(ii) Distributividade com respeito à soma de escalares: $(k_1+k_2)A = k_1A + k_2A$;

(iii) $0A = 0_{m\times n}$, onde $0 \in \mathbb{K}$ e $0_{m\times n}$ denota a matriz nula de ordem $m \times n$.

(iv) $k_1(k_2A) = (k_1k_2)A$.

Demonstração: Sejam $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$ e $k, k_1, k_2 \in \mathbb{K}$. Então temos:

(i) $k(A+B) = k(a_{ij}+b_{ij}) := \big(k(a_{ij}+b_{ij})\big) = (ka_{ij}+kb_{ij}) = (ka_{ij})+(kb_{ij}) := kA+kB$, com a terceira igualdade seguindo da distributividade válida nos números reais e complexos.

(ii) $(k_1+k_2)A := ((k_1+k_2)a_{ij}) = (k_1a_{ij}+k_2a_{ij}) = (k_1a_{ij})+(k_2a_{ij}) = k_1A+k_2A$, com a segunda igualdade seguindo também da associatividade dos números reais e complexos.

(iii) $0A := (0a_{ij}) = (0) = 0_{m\times n}$, com a segunda igualdade válida pois o produto do escalar zero com qualquer número real ou complexo é nulo.

(iv) $k_1(k_2A) := k_1(k_2a_{ij}) := (k_1k_2a_{ij}) = (k_1k_2)A$. ■

1.2.3 Produto de Matrizes

Sejam $A = (a_{ij})_{m\times n}$ e $B = (b_{jk})_{n\times p}$ matrizes cujo número de colunas da primeira coincide com o número de linhas da segunda. Definimos o produto de A por B, denotado por AB, como sendo a matriz $C = (c_{ik})_{m\times p}$ tal que para cada $i \in \{1, \cdots, m\}$ e $k \in \{1, \cdots, p\}$,

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}.$$

Exemplo:

$$\begin{pmatrix} 1 & -1 & i \\ 2 & 1 & 0 \end{pmatrix}_{2\times 3} \cdot \begin{pmatrix} i & 1 & -1 & \sqrt{3} \\ 0 & x & i & 1 \\ 0 & 0 & 1 & 8 \end{pmatrix}_{3\times 4} = \begin{pmatrix} i & 1-x & -1 & \sqrt{3}-1+8i \\ 2i & 2+x & -2+i & 2\sqrt{3}+1 \end{pmatrix}_{2\times 4}.$$

Proposição 1.5. *Sejam A, B e C matrizes arbitrárias com as respectivas ordens para que cada produto abaixo esteja definido. Então temos:*

(i) $AI = A$ *e* $IA = A$, *onde* I *é a matriz identidade;*

(ii) $A(B + C) = AB + AC$ *e* $(A + B)C = AC + BC$*;*

(iii) $(AB)C = A(BC)$*;*

(iv) $OA = O$ *e* $AO = O$, *onde* O *é a matriz nula.*

Demonstração:

(i) Sejam $A = (a_{ij})_{m\times n}$ e $I = (b_{jk})_{n\times n}$, com $b_{jk} = 1$ se $j = k$ e $b_{jk} = 0$ se $j \neq k$. Supondo $AI = (c_{ik})_{m\times n}$, então por definição temos

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = a_{ik}b_{kk} = a_{ik},$$

para todo $1 \leq i \leq m$ e $1 \leq k \leq n$. Portanto,

$$AI = (c_{ik})_{m\times n} = (a_{ik})_{m\times n} = A.$$

Analogamente podemos provar que $IA = A$, considerando $I = (b_{jk})_{m\times m}$.

(ii) Sejam $A = (a_{ij})_{m\times n}$, $B = (b_{jk})_{n\times p}$ e $C = (c_{jk})_{n\times p}$. Então $B + C = (b_{jk} + c_{jk})_{n\times p}$ e, considerando $A(B + C) = (d_{ik})_{m\times p}$, temos por definição

$$d_{ik} = \sum_{j=1}^{n} a_{ij}(b_{jk} + c_{jk}) = \sum_{j=1}^{n}(a_{ij}b_{jk} + a_{ij}c_{jk}) = \sum_{j=1}^{n} a_{ij}b_{jk} + \sum_{j=1}^{n} a_{ij}c_{jk}.$$

Logo, $A(B+C) = AB+AC$. A demonstração de que $(A+B)C = AC+BC$, com as ordens apropriadas, pode ser feita de modo análogo e será deixada como exercício.

(iii) Deixado como exercício.

(iv) Sejam $A = (a_{ij})_{m\times n}$ e $O = (b_{jk})_{n\times p}$, com $b_{jk} = 0$ para todo j, k. Supondo $AO = (c_{ik})_{m\times p}$, então por definição temos

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} = \sum_{j=1}^{n} a_{ij}0 = 0,$$

para todo $1 \leq i \leq m$ e $1 \leq k \leq p$. Portanto, $AO = O$, onde a matriz nula que aparece no lado direito da igualdade tem ordem $m \times p$. Analogamente podemos provar que $OA = O$, considerando as ordens apropriadas. ■

Observação 1.6. *O produto de matrizes* ***não é comutativo****, pois em geral $AB \neq BA$. De fato, mesmo que AB esteja bem definida, o produto BA pode não estar. E mesmo que AB e BA estejam bem definidas, pode não ocorrer a igualdade.*

Definição 1.7. *Se A e B são duas matrizes tais que $AB = BA$, então dizemos que A* ***comuta*** *com B ou que A e B* ***comutam****.*

Exemplo 1.8. Sejam

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \ B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \text{ e } C = \begin{pmatrix} 3 & -1 \\ \pi & 7 \end{pmatrix}.$$

Então A e B comutam, mas A e C não comutam:

$$AB = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = BA, \quad AC = \begin{pmatrix} 3+2\pi & 13 \\ -6+\pi & 9 \end{pmatrix} \neq \begin{pmatrix} 5 & 5 \\ \pi - 14 & 2\pi + 7 \end{pmatrix} = CA.$$

1.2.4 Matriz Inversa

Daqui em diante, denotaremos por I_n a matriz identidade de ordem n.

Definição 1.9. *Seja A uma matriz quadrada de ordem n. Dizemos que A é* ***inversível*** *se existir uma matriz B tal que $AB = BA = I_n$. Neste caso, B é chamada de* ***inversa*** *da matriz A e é denotada por $B = A^{-1}$.*

A nomenclatura e a notação introduzidas na definição anterior fazem sentido, já que se A for inversível, sua inversa é única. De fato, sejam B e C matrizes tais que $AC = CA = I_n$ e $AB = BA = I_n$. Então

$$C = I_nC = (BA)C = B(AC) = BI_n = B.$$

Além disso, veja que se A^{-1} existir, ela deve ser uma matriz quadrada de mesma ordem de A.

Exemplo 1.10. Considere as matrizes A e B do exemplo anterior. Vimos que A e B comutam, no entanto $AB = BA \neq I_2$, ou seja, B não é a inversa da matriz A, nem A é a inversa da matriz B. Agora, tome

$$D = \begin{pmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}.$$

Cálculos simples nos mostram que $AD = DA = I_2$, ou seja, a matriz D é a inversa da matriz A e vice-versa. Em símbolos, $A^{-1} = D$ e $D^{-1} = A$.

A proposição abaixo nos garante que se B é uma matriz tal que $AB = I_n$, então B é a inversa de A (o mesmo vale se tivermos a igualdade $BA = I_n$).

Proposição 1.11. *Seja A uma matriz quadrada de ordem n. Se existir uma matriz B tal que $AB = I_n$, então $BA = I_n$.*

Demonstração: Será feita mais adiante, pois ainda não temos ferramentas suficientes para tal. ■

Proposição 1.12. *Sejam A e B matrizes inversíveis. Então*

(i) A^{-1} é inversível, com $(A^{-1})^{-1} = A$;

(ii) AB é inversível, com $(AB)^{-1} = B^{-1}A^{-1}$.

Demonstração:

(i) Sendo $B = A^{-1}$, devemos provar que $B^{-1} = A$. Pela Proposição 1.11, basta mostrar que $BA = I_n$. Mas isso é claro, pois

$$BA = A^{-1}A = I_n.$$

(ii) Pela Proposição 1.11, basta provar que $AB(B^{-1}A^{-1}) = I_n$, o que de fato ocorre:

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I_n. \qquad ■$$

Em geral, o cálculo da inversa de uma matriz é um processo trabalhoso. Veremos, mais adiante, alguns métodos que nos auxiliam nesse processo.

1.2.5 Matriz Transposta

Definição 1.13. *Definimos a transposta de uma matriz $A = (a_{ij})_{m\times n}$, denotada por A^t, como a matriz cujas linhas são as colunas de A e cujas colunas são as linhas de A, isto é, $A^t = (a_{ji})_{n\times m}$.*

Exemplo: Se $A = \begin{pmatrix} 1 & -1 & i \\ 0 & 1 & 0 \end{pmatrix}_{2\times 3}$, então $A^t = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ i & 0 \end{pmatrix}_{3\times 2}$.

Proposição 1.14. *Para quaisquer matrizes A e B de ordem $m \times n$ e $k \in \mathbb{K}$, temos:*

(i) $(A + B)^t = A^t + B^t$;

(ii) $(kA)^t = kA^t$;

(iii) $(AB)^t = B^tA^t$;

(iv) $(A^t)^t = A$;

(v) $A^t = A$ *se, e somente se,* A *é simétrica;*

(vi) $A^t = -A$ *se, e somente se,* A *é antissimétrica;*

(vii) *Se* A *é inversível, então* A^t *também é com inversa* $(A^t)^{-1} = (A^{-1})^t$.

Demonstração: Considere $A = (a_{ij})_{m\times n}$.

(i) Se $B = (b_{ij})_{m\times n}$, então $A + B = (c_{ij})$, onde $c_{ij} := a_{ij} + b_{ij}$. Logo,

$$(A+B)^t := (c_{ji}) = (a_{ji} + b_{ji}) := (a_{ji}) + (b_{ji}) = A^t + B^t.$$

(ii) Por definição, $kA = (c_{ij})$ com $c_{ij} := ka_{ij}$. Portanto,

$$(kA)^t := (c_{ji}) = (ka_{ji}) = k(a_{ji}) = kA^t.$$

(iii) Se $B = (b_{jk})_{n\times p}$, então $AB = (c_{ik})_{m\times p}$, com

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}, \quad \forall\ 1 \leq i \leq m,\ 1 \leq k \leq p.$$

Agora, $A^t = (d_{ji})_{n\times m}$ e $B^t = (e_{kj})_{p\times n}$, onde $d_{ji} = a_{ij}$ e $e_{kj} = b_{jk}$, de modo que $B^tA^t = (f_{ki})$, com

$$f_{ki} = \sum_{j=1}^{n} e_{kj}d_{ji} = \sum_{j=1}^{n} b_{jk}a_{ij} = \sum_{j=1}^{n} a_{ij}b_{jk} = c_{ik}.$$

Portanto, $B^tA^t = (AB)^t$, uma vez que o elemento da linha k e coluna i de B^tA^t é o elemento da linha i e coluna k de AB.

(iv) Como $A^t = (a_{ji})$, então $(A^t)^t = (a_{ji})^t = (a_{ij}) = A$.

(v) Primeiramente observe que se $(a_{ji})_{n\times m} = A^t = A = (a_{ij})_{m\times n}$, então $m = n$, ou seja, A é uma matriz quadrada de ordem n. Além disso, $A^t = A$ se, e somente se, $a_{ji} = a_{ij}$, para todo $1 \leq i, j \leq n$, o ocorre se, e somente se, A é simétrica.

(vi) Assim como no item anterior, se $A^t = -A$, então A é uma matriz quadrada de ordem n. Além disso, $A^t = -A$ se, e somente se, $a_{ji} = -a_{ij}$, para todo $1 \leq i, j \leq n$, o ocorre se, e somente se, A é antissimétrica.

(vii) Pela propriedade (iii) temos que

$$A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t = I_n.$$

Segue, pela Proposição 1.11, que $(A^{-1})^t A^t = I_n$, o que implica que a matriz inversa de A^t é $(A^{-1})^t$. ■

Definição 1.15. *Seja $A = (a_{ij})_{n\times n}$ uma matriz quadrada de ordem n. Definimos o* **traço** *de A, denotado por $tr(A)$, como a soma dos elementos da diagonal principal de A, ou seja,*

$$tr(A) = \sum_{i=1}^{n} a_{ii}.$$

1.3 Escalonamento de Matrizes

O processo de escalonamento de matrizes é um dos mais úteis no estudo de matrizes e de sistemas lineares. Mencionamos anteriormente que determinar a inversa de uma matriz é uma tarefa complicada. Todavia através do processo de escalonamento de uma matriz temos um método sistemático (mas não menos trabalhoso) para determinar sua inversa. Esse processo também nos auxilia na determinação da solução de um sistema linear, já que nos fornece um outro sistema mais simples com a propriedade de possuir o mesmo conjunto solução do sistema inicial.

Para o que segue, denotemos por L_i a linha i de uma matriz. Vamos começar apresentando as três operações sobre as linhas de uma matriz, chamadas *operações elementares*. São elas:

1. Permutar a linha L_i pela linha L_j, com $j \neq i$ ($L_i \leftrightarrow L_j$).

2. Multiplicar a linha L_i por um escalar $k \in \mathbb{K}$ não nulo ($L_i \to kL_i$).

3. Substituir a linha L_i pela linha L_i mais k vezes a linha L_j, com $k \in \mathbb{K}$ não nulo e $j \neq i$ ($L_i \to L_i + kL_j$).

Tais operações elementares também podem ser realizadas sobre as colunas de uma matriz ao invés das linhas.

Exemplo 1.16.

$$A = \begin{pmatrix} 1 & -1 \\ 4 & 3 \\ 2 & 5 \end{pmatrix} \underset{L_2 \leftrightarrow L_3}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 2 & 5 \\ 4 & 3 \end{pmatrix} \underset{L_1 \to 2L_1}{\longrightarrow} \begin{pmatrix} 2 & -2 \\ 2 & 5 \\ 4 & 3 \end{pmatrix} \underset{L_3 \to L_3 - \frac{1}{2}L_1}{\longrightarrow} \begin{pmatrix} 2 & -2 \\ 2 & 5 \\ 3 & 4 \end{pmatrix} = B.$$

Definição 1.17. *Sejam A e B matrizes de ordem $m \times n$. Dizemos que B é* ***linha equivalente*** *à matriz A se B for obtida de A por meio de um número finito de operações elementares sobre as linhas de A. Vamos denotar a linha equivalência entre A e B por $A \sim B$.*

Definição 1.18. *1. Uma matriz $m \times n$ está na* ***forma escalonada por linhas*** *se:*

(a) Toda linha nula, se houver, está abaixo de todas as linhas não nulas;

(b) Cada primeiro elemento não nulo de uma linha está à direita do primeiro elemento não nulo da linha anterior. Em outras palavras, se as linhas $L_1, \ldots, L_r$ são não nulas e se o primeiro elemento não nulo da linha L_i ocorre na coluna k_i, então $k_1 < k_2 < \cdots < k_r$.

2. Uma matriz escalonada por linhas é ***linha reduzida à forma escada****, ou está na forma escada, se:*

(c) O primeiro elemento não nulo de uma linha não nula é 1;

(d) O primeiro elemento não nulo de uma linha é o único elemento não nulo em sua coluna.

Veja que em uma matriz escalonada por linhas o número de zeros que precedem o primeiro elemento não nulo de uma linha aumenta a cada linha, até que sobrem apenas as linhas nulas, se houverem. O mesmo ocorre com uma matriz linha reduzida à forma escada, já que também está na forma escalonada.

Exemplo 1.19. 1. As matrizes A e B do Exemplo 1.16 são linha equivalentes, ou seja, $A \sim B$. Vamos reduzir a matriz A a uma matriz linha reduzida à forma escada:

$$A = \begin{pmatrix} 1 & -1 \\ 4 & 3 \\ 2 & 5 \end{pmatrix} \underset{L_2 \to \frac{1}{4}L_2}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 1 & \frac{3}{4} \\ 2 & 5 \end{pmatrix} \underset{L_3 \to \frac{1}{2}L_3}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 1 & \frac{3}{4} \\ 1 & \frac{5}{2} \end{pmatrix} \underset{L_2 \to L_2 - L_1}{\longrightarrow}$$

$$\begin{pmatrix} 1 & -1 \\ 0 & \frac{7}{4} \\ 1 & \frac{5}{2} \end{pmatrix} \underset{L_3 \to L_3 - L_1}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 0 & \frac{7}{4} \\ 0 & \frac{7}{2} \end{pmatrix} \underset{L_2 \to \frac{4}{7}L_2}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 0 & \frac{7}{2} \end{pmatrix} \underset{L_3 \to \frac{2}{7}L_3}{\longrightarrow}$$

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \underset{L_3 \to L_3 - L_2}{\longrightarrow} \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \underset{L_1 \to L_1 + L_2}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = C.$$

A matriz C é uma matriz linha reduzida à forma escada tal que $A \sim C$. A matriz

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

também é linha equivalente à matriz A, mas não é uma matriz linha reduzida à forma escada, embora esteja na forma escalonada.

2. A matriz abaixo não está na forma escada, pois falha o item (d) da definição anterior.

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

3. A matriz abaixo não está na forma escada, pois falham os itens (b) e (c) da definição anterior.

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

4. A matriz abaixo está na forma escada.

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

Teorema 1.20. *Toda matriz A é linha equivalente a uma única matriz linha reduzida à forma escada.*

Demonstração: Suponha $A = (a_{ij})_{m \times n}$ uma matriz não nula (se A for nula, já está reduzida à forma escada). Então existe um menor inteiro positivo k_1 tal que $a_{ik_1} \neq 0$ para alguma linha L_i não nula. Permute tal linha, se necessário, de modo que esta seja a primeira linha da matriz A, ou seja, $L_i \leftrightarrow L_1$. Depois, realize as seguintes operações elementares nessa ordem:

1. $L_1 \to \frac{1}{a_{1k_1}} L_1$, o que é possível pois $a_{1k_1} \neq 0$;

2. $L_j \to L_j - a_{jk_1} L_1$, para todo $2 \leq j \leq m$.

Assim, obtemos uma matriz $B = (b_{ij})_{m \times n}$ cujo primeiro elemento não nulo da primeira linha é 1 (ou seja, $b_{1k_1} = 1$) e todos os demais elementos abaixo de b_{1k_1} são iguais a zero. O elemento b_{1k_1} é chamado pivô da linha L_1. Se as linhas L_i da matriz B, com $2 \leq i \leq m$, forem todas nulas, B já está linha reduzida à forma escada. Se não, existe um menor inteiro positivo k_2 tal que $b_{ik_2} \neq 0$ para alguma linha L_i não nula, com $2 \leq i \leq m$. Permute tal linha, se necessário, de modo que esta seja a segunda linha da matriz B, ou seja, $L_i \leftrightarrow L_2$. Depois, realize as seguintes operações elementares nessa ordem:

1. $L_2 \to \frac{1}{b_{2k_2}} L_2$, o que é possível pois $b_{2k_2} \neq 0$;

2. $L_j \to L_j - b_{jk_2}L_2$, para todo $3 \leq j \leq m$.

Observe que $k_1 < k_2$. Assim, temos agora uma nova matriz C cujas colunas anteriores a k_2 permaneceram como na matriz B, o primeiro elemento não nulo da segunda linha é $c_{2k_2} = 1$ (esse é o pivô da linha L_2) e todos os demais elementos abaixo de c_{2k_2} são nulos. Repetindo esse processo, obteremos uma matriz $M = (m_{ij})_{m\times n}$ com $k_1 < k_2 < \cdots < k_r$ que está na forma escalonada. Para que o primeiro elemento não nulo de cada linha L_i com $r \leq i \leq m$ seja 1, basta fazer $L_i \to \frac{1}{m_{ik_i}}L_i$. Vamos supor que M já esteja nesse formato.

Resta somente que cada pivô m_{ik_i} seja o único elemento não nulo em sua coluna. Para isso, comece pelo último pivô, o elemento $m_{rk_r} = 1$, anulando todos os elementos acima dele (os que estão abaixo já são nulos). As operações elementares correspondentes a essa etapa são:

$$L_j \to L_j - a_{jk_r}L_r, \text{ para todo } 1 \leq j \leq r-1.$$

Repita o processo com todos os pivôs m_{ik_i} de baixo para cima até (e inclusive) o pivô m_{2k_2}, ou seja, para cada i fixado, com i variando de $r-1$ até 2, as operações elementares correspondentes a essa etapa são:

$$L_j \to L_j - a_{jk_i}L_i, \text{ para todo } 1 \leq j \leq i-1.$$

Como acima do primeiro pivô $m_{1k_1} = 1$ não há elementos, o processo se completa. Obtemos assim uma matriz N linha equivalente à matriz A que está linha reduzida à forma escada.

Provemos agora a unicidade da matriz N. Para isso, suponha que A seja linha equivalente a duas matrizes N e P, ambas linha reduzidas à forma escada. Logo, $A \sim N$ e $A \sim P$. Como as operações com linhas são reversíveis, temos que $N \sim P$. Entretanto, N e P estão na forma escada, o que implica que $N = P$, pois nenhuma das três operações elementares com linhas pode ser efetuada numa matriz na forma escada sem que ela perca esta condição. ■

Em particular, a demonstração do teorema acima nos mostra que qualquer matriz A é linha equivalente a uma matriz na forma escalonada, que não é única. Veremos no Teorema 1.41 da Seção 1.6 que a identidade é a única matriz na forma escada que é linha equivalente a uma matriz inversível.

Exemplo 1.21. 1. Considere a matriz

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

Vamos seguir os passos descritos na demonstração do teorema anterior a fim de obter a única matriz na forma escada linha equivalente a A.

O menor inteiro positivo k_1 tal que $a_{ik_1} \neq 0$ é $k_1 = 1$. Neste caso podemos escolher $i = 1$ ou $i = 3$. Escolho $i = 1$. Logo $a_{1k_1} = -1$ e temos:

$$A = \begin{pmatrix} -1 & 4 & 2 \\ 0 & 1 & -3 \\ -1 & 1 & 1 \end{pmatrix} \underset{L_1 \to -L_1}{\longrightarrow} \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ -1 & 1 & 1 \end{pmatrix} \underset{L_3 \to L_3 + L_1}{\longrightarrow} \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ 0 & -3 & -1 \end{pmatrix} = B.$$

Assim, obtemos uma matriz B cujo primeiro elemento não nulo da primeira linha é 1 e todos os demais elementos abaixo de b_{1k_1} são iguais a zero. Agora, o menor inteiro positivo k_2 tal que $b_{ik_2} \neq 0$ é $k_2 = 2$. Observe que $k_1 < k_2$. Podemos escolher $i = 2$ ou $i = 3$. Considerando $i = 2$ temos $b_{2k_2} = 1$ e, portanto,

$$B = \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ 0 & -3 & -1 \end{pmatrix} \underset{L_3 \to L_3 + 3L_2}{\longrightarrow} \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & -10 \end{pmatrix} = C.$$

Obtemos então uma matriz C escalonada cujas colunas anteriores a $k_2 = 2$ permaneceram como na matriz B, o primeiro elemento não nulo da segunda linha é $c_{2k_2} = 1$ e todos os demais elementos abaixo de c_{2k_2} são nulos. Resta agora reduzir C à forma escada. Para isso, faça:

$$C = \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & -10 \end{pmatrix} \underset{L_3 \to \frac{1}{-10} L_3}{\longrightarrow} \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \underset{L_2 \to L_2 + 3L_3}{\longrightarrow} \begin{pmatrix} 1 & -4 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\underset{L_1 \to L_1 + 2L_3}{\longrightarrow} \begin{pmatrix} 1 & -4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underset{L_1 \to L_1 + 4L_2}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Considere a matriz A do item anterior. Realizando a operação $L_2 \to L_2 - L_1$ seguida de $L_3 \to L_3 - 3L_1$ nas linhas de A, obtemos a matriz

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

Realizando essas mesmas operações na matriz identidade I_3 obtemos a matriz

$$E = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix},$$

que pode ser escrita como o produto $E = E_2E_1$, onde

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}.$$

Observe que E_1 e E_2 são as matrizes obtidas pelas operações $L_2 \to L_2 - L_1$ e $L_3 \to L_3 - 3L_1$ nas linhas de I_3, respectivamente. Além disso,

$$EA = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 4 & 2 \\ 0 & 1 & -3 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 4 & 2 \\ 1 & -3 & 5 \\ 2 & -11 & -5 \end{pmatrix} = B.$$

3. Considere a matriz

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

Realizando primeiro a operação $L_1 \to -L_1$ e depois $L_2 \to L_2 - 4L_1$ nas linhas de A obtemos, respectivamente, as matrizes

$$B = \begin{pmatrix} 1 & -1 & -3 \\ 4 & 2 & 0 \end{pmatrix} \quad \text{e} \quad C = \begin{pmatrix} 1 & -1 & -3 \\ 0 & 6 & 12 \end{pmatrix}.$$

Aplicando as mesmas operações na matriz identidade I_2 de ordem 2 obtemos, respectivamente, as matrizes

$$E_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{e} \quad E_2 = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix},$$

que multiplicadas pela matriz A nos fornece $(E_2E_1)A = E_2(E_1A) = E_2B = C$. Observe que a matriz

$$E = E_2E_1 = \begin{pmatrix} -1 & 0 \\ 4 & 1 \end{pmatrix}$$

é obtida aplicando primeiro $L_1 \to -L_1$ e depois $L_2 \to L_2 - 4L_1$ nas linhas de I_2.

O fato que observamos nos exemplos anteriores não é uma coincidência. Como veremos a seguir, aplicar uma operação elementar nas linhas de uma matriz A de ordem $m \times n$ é o mesmo que aplicar essa operação elementar na matriz identidade de ordem m e, em seguida, multiplicar essa matriz por A.

Definição 1.22. *Uma* ***matriz elementar*** *E é qualquer matriz obtida da matriz identidade pela aplicação de uma das operações elementares.*

Proposição 1.23. *Seja A uma matriz de ordem $m \times n$. Se A' é a matriz obtida de A pela aplicação de uma operação elementar "e" sobre linhas, então $A' = EA$, onde $E = e(I_m)$ é a matriz elementar correspondente obtida da matriz identidade de ordem m. No caso de uma operação elementar sobre colunas, vale que $A' = AE$.*

Corolário 1.24. *Se B é linha equivalente a uma matriz A, então $B = E_r \dots E_2 E_1 A$, onde $E_1, \dots, E_r$ são as matrizes elementares correspondentes a cada o operação elementar necessária para obter B.*

Corolário 1.25. *Toda matriz elementar $E = e(I_m)$ é inversível e sua inversa também é uma matriz elementar, obtida aplicando a operação inversa de e nas linhas de I_m.*

operação elementar		operação inversa
$L_i \leftrightarrow L_j$	$\longrightarrow$	$L_j \leftrightarrow L_i$
$L_i \to kL_i$	$\longrightarrow$	$L_i \to \frac{1}{k}L_i$
$L_i \to L_i + kL_j$	$\longrightarrow$	$L_i \to L_i - kL_j$

1.4 Resolução de Sistemas Lineares via escalonamento

Nesta seção vamos determinar um método prático, via escalonamento de matrizes, para obter as soluções de um sistema linear. Comecemos com as seguintes definições:

Definição 1.26. *Uma **equação linear** nas variáveis (ou incógnitas) $x_1, x_2, \dots, x_n$ é uma equação do tipo*

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

*onde $a_i, b \in \mathbb{K}$, para todo $1 \leq i \leq n$. Cada constante a_i é chamada **coeficiente** da variável x_i e a constante b é chamada **termo independente**.*

Uma **solução** de uma equação linear é uma $n-$upla $(x_1, x_2, \dots, x_n)$ que a satisfaça. Por exemplo, a tripla $(2, 1, -2)$ é uma solução da equação linear $2x_1 + 3x_2 - x_3 = 9$, pois $2.2 + 3.1 - (-2) = 9$. Uma equação linear em 2 ou mais variáveis possuem infinitas soluções. Geometricamente podemos ver isso para $n = 2$ e $n = 3$, já que uma equação em 2 variáveis representa uma reta e em 3 variáveis representa um plano.

Definição 1.27. *Um **sistema de equações lineares**, ou simplesmente um **sistema linear**, com m equações e n incógnitas é um conjunto de equações do tipo*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{1.1}$$

com $a_{ij}, b_i \in \mathbb{K}$, $1 \leq i \leq m$ e $1 \leq j \leq n$. As constantes a_{ij} são os coeficientes das incógnitas x_j, e b_i são os termos independentes.

Por exemplo,

$$\begin{cases} x+y+z &= 3 \\ 2x+2y+z &= 1 \\ -x+y-z &= 2 \end{cases}$$

é um sistema linear com 3 equações e 3 incógnitas x, y e z.

Definição 1.28. *Uma* ***solução*** *de um sistema linear da forma (1.1) é uma n-upla* $(x_1, x_2, \ldots, x_n)$, *com* $x_i \in \mathbb{K}$, *satisfazendo simultaneamente as* m *equações do sistema.*

Para exemplificar, considere o sistema linear

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

É fácil ver que $(x_1, x_2, x_3) = (7, 6, -10)$ é solução desse sistema, uma vez que satisfaz todas as 3 equações. Podemos nos perguntar: tal sistema admite mais soluções? Se sim, quais são elas? Como mencionamos anteriormente, nosso propósito nesta seção é apresentar um método que nos permita determinar se um sistema linear possui ou não solução e, caso as tenha, determinar quais são. Comecemos definindo matriz ampliada de um sistema.

Ao sistema (1.1) podemos associar uma equação matricial da forma

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

Assim, se denotarmos por C as matriz dos coeficientes, por X a matriz das variáveis e por B a matriz dos termos independentes, podemos escrever (1.1) como $CX = B$.

Definição 1.29. *A matriz*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}_{m\times(n+1)} \tag{1.2}$$

é denominada ***matriz ampliada*** *ou* ***matriz aumentada*** *do sistema (1.1). Por simplicidade, vamos eventualmente denotar* $A = (C \vdots B)$, *onde* C *é a matriz dos coeficientes e* B *a matriz dos termos independentes do sistema.*

O teorema abaixo nos apresenta um modo de determinar as soluções de um sistema linear (caso existam) por meio de outro sistema linear mais simples e obtido a partir do sistema original. O teorema garante que o conjunto solução, sendo ele vazio ou não, de um sistema linear não se altera ao realizarmos operações elementares em sua matriz ampliada.

TEOREMA 1.30. *Sistemas lineares com matrizes ampliadas linha equivalentes possuem o mesmo conjunto solução.*

DEMONSTRAÇÃO: Considere os sistemas lineares (I) e (II) cujas formas matriciais são, respectivamente,

$$CX = B \quad \text{e} \quad C'X = B',$$

com as matrizes dos coeficientes C, C' de ordem $m \times n$ e as matrizes dos termos independentes B, B' de ordem $m \times 1$. Sejam $A = (C \vdots B)$ e $A' = (C' \vdots B')$ suas respectivas matrizes ampliadas, que por hipótese são linha equivalentes. Pelo Corolário 1.24, podemos escrever $A' = MA$, onde $M = E_r \dots E_1$ é um produto de matrizes elementares de ordem m e, portanto, é inversível (Corolário 1.25). Ainda,

$$(C' \vdots B') = A' = MA = M(C \vdots B) = (MC \vdots MB),$$

onde MC e MB têm ordens $m \times n$ e $m \times 1$, respectivamente. Portanto $C' = MC$ e $B' = MB$. Além disso, como M é inversível

$$CX = B \;\Leftrightarrow\; M(CX) = MB \;\Leftrightarrow (MC)X = MB \;\Leftrightarrow\; C'X = B',$$

ou seja, se X for solução do sistema (I) então X será solução do sistema (II) e vice-versa.

■

Portanto, encontrar as soluções de um sistema linear se reduz a um problema matricial e o processo de escalonamento de matrizes nos fornece um método sistemático para a obtenção de tais soluções:

Método de resolução de sistemas via escalonamento: Dado o sistema $CX = B$, reduza sua matriz ampliada $A = (C \vdots B)$ à forma escada $A' = (C' \vdots B')$. Resolva o sistema associado $C'X = B'$, que é um sistema mais fácil de ser resolvido e cujas soluções coincidem com as soluções do sistema inicial.

EXEMPLO 1.31. Vamos verificar se os sistemas abaixo possuem ou não solução.

$$(a)\left\{\begin{array}{rcl} x+y+z &=& 3\\ y+2z &=& 2\\ x-z &=& 1 \end{array}\right. \qquad (b)\left\{\begin{array}{rcl} x+y+z &=& 3\\ 2y+z &=& 2\\ y+2z &=& 2 \end{array}\right. \qquad (c)\left\{\begin{array}{rcl} x+y+z &=& -10\\ 2x+y+z &=& -20\\ y+z &=& -40 \end{array}\right.$$

No caso (a), a matriz ampliada do sistema e sua forma escada são, respectivamente,

$$A = \begin{pmatrix} 1 & 1 & 1 & 3\\ 0 & 1 & 2 & 2\\ 1 & 0 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & 1\\ 0 & 1 & 2 & 2\\ 0 & 0 & 0 & 0 \end{pmatrix} = A'.$$

O sistema associado a A' é

$$\begin{cases} x - z &= 1 \\ y + 2z &= 2 \end{cases},$$

o qual possui infinitas soluções da forma $(x, y, z) = (1 + \lambda, 2 - 2\lambda, \lambda)$, com $\lambda \in \mathbb{K}$, sendo essas as soluções do sistema inicial.
No caso (b), a matriz ampliada do sistema e sua forma escada são, respectivamente,

$$A = \begin{pmatrix} 1 & 1 & 3 & 3 \\ 0 & 2 & 1 & 2 \\ 0 & 1 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & \frac{5}{3} \\ 0 & 1 & 0 & \frac{2}{3} \\ 0 & 0 & 1 & \frac{2}{3} \end{pmatrix} = A'.$$

Portanto, o sistema associado a A' possui uma única solução $(x, y, z) = (\frac{5}{3}, \frac{2}{3}, \frac{2}{3})$, sendo essa a solução do sistema inicial.
No caso (c), a matriz ampliada do sistema e sua forma escada são, respectivamente,

$$A = \begin{pmatrix} 1 & 1 & 1 & -10 \\ 2 & 1 & 1 & -20 \\ 0 & 1 & 1 & -40 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = A',$$

com o sistema associado a A' dado por

$$\begin{cases} x &= 0 \\ y + z &= 0 \\ 0x + 0y + 0z &= 1 \end{cases}.$$

Portanto, o sistema inicial não possui solução.

No próximo teorema veremos que um sistema de equações lineares pode admitir uma única solução, infinitas soluções ou nenhuma solução. Para isso, precisamos da seguinte definição:

Definição 1.32. *Seja A uma matriz de ordem $m \times n$. O posto de A, denotado por p_A, é o número de linhas não nulas de A quando reduzida à forma escalonada ou à forma escada. A nulidade de A, denotada por n_A, é a diferença $n - p_A$.*

Equivalentemente, o posto de uma matriz A também pode ser definido como o número de linhas (ou de colunas) linearmente independentes de A. Dessa forma, é possível mostrar que a nulidade de uma matriz nunca é negativa, já que $p_A \leq min\{m, n\}$.

Teorema 1.33. **(Teorema do Posto)** *Considere um sistema de equações lineares com m equações e n variáveis, sendo A sua matriz ampliada e C a matriz dos coeficientes. Então o sistema admite solução se, e somente se, $p_A = p_C$ (ou seja, o posto de A e C coincidem). Ainda,* *(i) Se $p_A = p_C = n$ então o sistema terá solução única;*

(ii) Se $p_A = p_C = r < n$ então o sistema terá infinitas soluções. Podemos escolher $n_C = n - r$ variáveis livres, e as outras r variáveis serão dadas em função dessas.

Demonstração: Considere A na forma (1.2) e suponha que o sistema admita solução. Se $p_A \neq p_C$, então $p_A > p_C$ (pois p_A nunca é menor que p_C), ou seja, a matriz A reduzida à forma escada deve ter pelo menos uma linha da forma

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & c_k \end{pmatrix},$$

com $c_k \neq 0$. Logo, o sistema associado à matriz reduzida, cujas soluções coincidem com as do sistema inicial, tem uma equação da forma

$$0x_1 + 0x_2 + \ldots + 0x_n = c_k \neq 0$$

e, portanto, não admite solução. Essa contradição nos garante que $p_A = p_C$.

Suponha agora que $p_A = p_C$ e observe que ambos os itens (i) e (ii) provam a recíproca do teorema. Portanto, vamos considerar os seguintes casos:

(i) $p_A = p_C = n$. Nesse caso, $m \geq n$ (pois caso contrário $p_C < n$) e a matriz ampliada do sistema pode ser reduzida à forma escada

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & 0 & \cdots & 0 & c_2 \\ 0 & 0 & 1 & \cdots & 0 & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & c_n \end{pmatrix}_{n\times(n+1)} \quad \text{ou} \quad \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & 0 & \cdots & 0 & c_2 \\ 0 & 0 & 1 & \cdots & 0 & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & c_n \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}_{m\times(n+1)}.$$

Em ambos os casos, a solução do sistema é $x_1 = c_1,\ x_2 = c_2,\ \ldots,\ x_n = c_n$, ou seja, o sistema admite uma única solução.

(ii) $p_A = p_C = r < n$. Nesse caso temos várias possibilidades para a matriz na forma escada linha equivalente a A, denotada por A'. A menos de uma permutação das variáveis podemos supor primeiramente que A' é da forma

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & a_{1r+1} & \cdots & a_{1n} & c_1 \\ 0 & 1 & 0 & \cdots & 0 & a_{2r+1} & \cdots & a_{2n} & c_2 \\ 0 & 0 & 1 & \cdots & 0 & a_{3r+1} & \cdots & a_{3n} & c_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & a_{rr+1} & \cdots & a_{rn} & c_r \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{m\times(n+1)},$$

cujo sistema associado, com n variáveis e $r < n$ equações, é da forma

$$\begin{cases} x_1 + a_{1r+1}x_{r+1} + \ldots + a_{1n}x_n = c_1 \\ x_2 + a_{2r+1}x_{r+1} + \ldots + a_{2n}x_n = c_2 \\ x_3 + a_{3r+1}x_{r+1} + \ldots + a_{3n}x_n = c_3 \\ \vdots \\ x_r + a_{rr+1}x_{r+1} + \ldots + a_{rn}x_n = c_r \end{cases}.$$

Portanto, o sistema admite infinitas soluções com $n_C = n - r$ variáveis livres, a saber $x_{r+1}, \ldots, x_n$. Variando as variáveis livres em $\mathbb{K}$ obtemos

$$\begin{cases} x_1 = c_1 - a_{1r+1}x_{r+1} - \ldots - a_{1n}x_n \\ x_2 = c_2 - a_{2r+1}x_{r+1} - \ldots - a_{2n}x_n \\ x_3 = c_3 - a_{3r+1}x_{r+1} - \ldots - a_{3n}x_n \\ \vdots \\ x_r = c_r - a_{rr+1}x_{r+1} - \ldots - a_{rn}x_n \end{cases}.$$

A matriz A' também pode ser da forma

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & a_{1r+2} & \cdots & a_{1n} & c_1 \\ 0 & 0 & 1 & \cdots & 0 & a_{2r+2} & \cdots & a_{2n} & c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{rr+2} & \cdots & a_{rn} & c_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{m\times(n+1)},$$

cujo sistema associado, com $n-1$ variáveis e $r < n$ equações, admite infinitas soluções, a saber

$$\begin{cases} x_2 = c_1 - a_{1r+2}x_{r+2} - \ldots - a_{1n}x_n \\ \vdots \\ x_{r+1} = c_r - a_{rr+2}x_{r+2} - \ldots - a_{rn}x_n, \end{cases}$$

sendo $x_1, x_{r+2}, \ldots, x_n$ as $n_C = n - r$ variáveis livres. Procedendo sucessivamente podemos analisar as outras possibilidades, a menos de uma permutação das variáveis, e concluir que o sistema possui infinitas soluções. ■

Um sistema de equações lineares é dito **possível e determinado (SPD)** se o sistema possuir uma única solução, é dito **possível e indeterminado (SPI)** se admitir infinitas soluções e é dito **impossível (SI)** se não admitir solução.

EXERCÍCIO 1.34. *Verifique se os sistemas do Exemplo 1.31 são SPD, SPI ou SI analisando apenas o posto das respectivas matrizes dos coeficientes e ampliadas.*

DEFINIÇÃO 1.35. *Um sistema de equações lineares é dito* ***homogêneo*** *se todos os termos independentes são iguais a zero, ou seja,*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases}$$

com $a_{ij}, b_i \in \mathbb{K}$, $1 \leq i \leq m$ e $1 \leq j \leq n$.

Observe que $(x_1, \ldots, x_n) = (0, \ldots, 0)$ é sempre solução de um sistema homogêneo, chamada **solução nula ou trivial**. Portanto, um sistema homogêneo é sempre SPD ou SPI, e no primeiro caso sua única solução é a trivial.

1.5 Determinante

O determinante de uma matriz quadrada é um número real ou complexo associado a essa matriz e que depende apenas de seus elementos. Nesta seção, vamos definir o determinante de uma matriz de ordem n utilizando o desenvolvimento de Laplace, uma fórmula recursiva que recai em determinantes de matrizes de ordem inferior a n. Como veremos na próxima seção, o determinante de uma matriz nos diz se ela é inversível ou não.

DEFINIÇÃO 1.36. *O* ***determinante*** *de uma matriz $A = (a_{ij})_{n\times n}$, denotado por $\det(A)$ ou por $|A|$, é um número real ou complexo definido como segue:*

(a) Se $n = 1$, então $\det(A) = a_{11}$.

(b) Se $n > 1$, então fixada uma linha L_i de A temos

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}), \tag{1.3}$$

onde A_{ij} é a submatriz de ordem $n-1$ obtida de A retirando-se a linha L_i e a coluna C_j.

A igualdade (1.3) é chamada de **Desenvolvimento de Laplace** e também pode ser definida fixando-se uma coluna de A ao invés de uma linha, ou seja, variando $i = 1, \ldots, n$ ao invés de j.

Exemplo 1.37. 1. Sejam

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Por Laplace, fixando a primeira linha de A e de B temos, respectivamente,

$$\det(A) = (-1)^{1+1} a_{11} |(a_{22})| + (-1)^{1+2} a_{12} |(a_{21})| = a_{11}a_{22} - a_{12}a_{21}$$

e

$$\begin{aligned} \det(B) &= (-1)^{1+1} b_{11} \begin{vmatrix} b_{22} & b_{23} \\ b_{32} & b_{33} \end{vmatrix} + (-1)^{1+2} b_{12} \begin{vmatrix} b_{21} & b_{23} \\ b_{31} & b_{33} \end{vmatrix} + (-1)^{1+3} b_{13} \begin{vmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix} \\ &= b_{11}(b_{22}b_{33} - b_{23}b_{32}) - b_{12}(b_{21}b_{33} - b_{23}b_{31}) + b_{13}(b_{21}b_{32} - b_{22}b_{31}) \\ &= b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} - (b_{11}b_{23}b_{32} + b_{12}b_{21}b_{33} + b_{13}b_{22}b_{31}). \end{aligned}$$

Existe um esquema de memorização para o cálculo do determinante de uma matriz de ordem 3 conhecido como **Regra de Sarrus**, devido ao matemático francês Pierre Frederic Sarrus. Ele consiste em repetir as duas primeiras colunas à direita da matriz B. Em seguida, os elementos da diagonal principal são multiplicados. Esse processo deve ser feito também com as diagonais que estão à direita da diagonal principal para que seja possível somar os produtos dessas três diagonais, resultando em um número p. O mesmo deve ser realizado com a diagonal secundária e as demais diagonais à sua direita, subtraindo os produtos encontrados do número p, obtido anteriormente. O número resultante é o determinante da matriz B.

Exemplo 1.38. Usando o desenvolvimento de Laplace, vamos calcular o determinante da matriz

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

Fixando a terceira linha da matriz, obtemos

$$\det(A) = (-1)^{3+1} 2 \begin{vmatrix} -1 & 0 & 3 \\ 4 & 3 & 1 \\ -3 & 7 & 2 \end{vmatrix} + (-1)^{3+2} 5 \begin{vmatrix} 1 & 0 & 3 \\ 0 & 3 & 1 \\ 1 & 7 & 2 \end{vmatrix} = 2(112) - 5(-10) = 274.$$

Proposição 1.39. *Sejam A e B matrizes quadradas de ordem n.*

(a) Se A possui uma linha nula, então $\det(A) = 0$*;*

(b) Se B é uma matriz obtida de A permutando-se a posição de duas linhas de A ($L_i \leftrightarrow L_j$), então $\det(B) = -\det(A)$*;*

(c) Se A possui duas linhas iguais, então $\det(A) = 0$*;*

(d) Se B é uma matriz obtida de A multiplicando uma linha de A por um escalar $k \in \mathbb{K}$ não nulo ($L_i \to kL_i$), então $\det(B) = k\det(A)$*;*

(e) Se B é uma matriz obtida de A somando-se a uma linha de A um múltiplo não nulo de outra linha ($L_i \to L_i + kL_r$), então $\det(B) = \det(A)$*;*

(f) $\det(kA) = k^n \det(A)$, *para todo escalar* $k \in \mathbb{K}$;

(g) $\det(A^t) = \det(A)$;

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

Demonstração:

(a) Basta usar (1.3) fixando justamente a linha nula.

(b) Pode ser feita usando uma definição alternativa de determinante por meio de permutações.

(c) Suponha que A possua duas linhas L_i e L_j iguais e seja B a matriz obtida de A permutando-se L_i e L_j. Então $B = A$, de modo que $\det(B) = \det(A)$. Entretanto, pelo item (b) o determinante troca de sinal quando permutamos duas linhas da matriz, ou seja, $\det(B) = -\det(A)$. Portanto, $\det(A) = 0$.

(d) Seja B a matriz obtida de A após a aplicação da operação elementar $L_i \to kL_i$. Fixando essa linha (a linha kL_i) no desenvolvimento de Laplace, temos

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j} k a_{ij} \det(B_{ij}),$$

onde B_{ij} é a submatriz obtida de B retirando-se a linha kL_i e a coluna j, ou seja, B_{ij} não possui a linha kL_i qualquer que seja $j = 1, \ldots, n$. Desse modo concluímos que $B_{ij} = A_{ij}$, onde A_{ij} é a submatriz obtida de A retirando-se a linha i e a coluna j. Portanto,

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j} k a_{ij} \det(A_{ij}) = k \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) = k \det(A).$$

(e) Seja B a matriz obtida de A pela operação elementar $L_i \to L_i + kL_r$. Fixando essa linha (a linha $L_i + kL_r$) no desenvolvimento de Laplace, temos

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j}(a_{ij} + ka_{rj}) \det(B_{ij}),$$

onde B_{ij} é a submatriz obtida de B retirando-se a linha $L_i + kL_r$ e a coluna j. Pelo mesmo argumento anterior concluímos que B_{ij} coincide com a submatriz A_{ij} obtida de A retirando-se a linha i e a coluna j. Portanto,

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) + k \sum_{j=1}^{n} (-1)^{i+j} a_{rj} \det(A_{ij}) = \det(A) + k \det(C),$$

onde

$$C = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{r1} & \dots & a_{rj} & \dots & a_{rn} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{r1} & \dots & a_{rj} & \dots & a_{rn} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

e a última igualdade segue por (1.3) fixando a i-ésima linha de C. Como C possui duas linhas iguais temos que $\det(C) = 0$. Logo, $\det(B) = \det(A)$.

(f) Temos que kA é a matriz obtida de A multiplicando suas n linhas por um escalar k. Segue do item (d) que se multiplicarmos uma linha da matriz por k o determinante também é multiplicado por k. Aplicando esse resultado nas n linhas de kA temos que $\det(kA) = k^n \det(A)$.

(g) Como o determinante também pode ser calculado fixando-se uma coluna ao invés de uma linha, o resultado segue.

(h) Será demonstrado nos Teoremas 1.45 e 1.51. ■

Proposição 1.40. *Se $A = (a_{ij})_{n\times n}$ é uma matriz triangular, então* $\det(A) = \prod_{i=1}^{n} a_{ii}$.

Demonstração: Suponhamos A uma matriz triangular superior. Vamos fazer a prova por indução sobre a ordem n da matriz. Se A tem ordem $n = 2$, ou seja,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix},$$

então $\det(A) = a_{11}a_{22} - 0a_{12} = a_{11}a_{22}$.

Agora suponha que A tenha ordem $n > 2$ e que toda matriz triangular superior de ordem $n-1$ tenha como determinante o produto dos elementos de sua diagonal principal. Fixando a última linha de A no desenvolvimento de Laplace, obtemos

$$\det(A) = (-1)^{n+n} a_{nn} \det(A_{nn}) = a_{nn} \det(A_{nn}),$$

onde A_{nn} é a matriz obtida de A retirando-se a última linha e a última coluna. Claramente A_{nn} é uma matriz triangular superior de ordem $(n-1)$. Pela hipótese de indução $\det(A_{nn}) = \prod_{i=1}^{n-1} a_{ii}$. Portanto,

$$\det(A) = a_{nn} \prod_{i=1}^{n-1} a_{ii} = \prod_{i=1}^{n} a_{ii}.$$

A prova para uma matriz triangular inferior é análoga, agora fixando a primeira linha no cálculo do determinante. ■

1.6 Inversão de Matrizes

Nosso objetivo nesta seção é apresentar dois métodos práticos para o cálculo da inversa de uma matriz inversível, o primeiro deles usando escalonamento de matrizes e o segundo usando matriz adjunta. O próximo teorema nos afirma, entre outras coisas, que ou existem operações elementares que transformam A em I_n ou A não é inversível.

Teorema 1.41. *Seja A uma matriz quadrada de ordem n. As afirmações abaixo são equivalentes:*

(i) A é inversível;

(ii) A é linha equivalente à matriz identidade I_n;

(iii) A é um produto finito de matrizes elementares.

Demonstração: $(i) \Rightarrow (ii)$ Suponhamos que A seja inversível com inversa A^{-1}, ou seja, $AA^{-1} = A^{-1}A = I_n$. Então o sistema linear homogêneo $AX = 0$ possui uma única solução

$$X = I_n X = (A^{-1}A)X = A^{-1}(AX) = A^{-1}0 = 0.$$

Por outro lado, seja A' a matriz na forma escada linha equivalente à matriz A. Pelo Teorema 1.30, os sitemas $AX = 0$ e $A'X = 0$ possuem o mesmo conjunto solução, ou seja, $A'X = 0$ admite uma única solução, de modo que $p_{A'} = n$. Como A' é uma matriz quadrada na forma escada, temos que $A' = I_n$. Portanto, A é linha equivalente a I_n.

$(ii) \Rightarrow (iii)$ Suponhamos que A seja linha equivalente a I_n. Pelo Corolário 1.24, existem matrizes elementares $E_1, \dots, E_r$ tais que $E_r \dots E_1 A = I_n$. Como cada E_i é inversível temos que $E_r \dots E_1$ é inversível e, portanto,

$$A = (E_r \dots E_1)^{-1} I_n = E_1^{-1} \dots E_r^{-1}.$$

Como a inversa de uma matriz elementar é também uma matriz elementar, o resultado segue.

$(iii) \Rightarrow (i)$ Seja $A = E_1 \dots E_s$, onde $E_1, \dots, E_s$ são matrizes elementares. Como cada E_i é inversível, segue que A também é inversível. ■

Corolário 1.42. *Se A é uma matriz inversível, então sua inversa A^{-1} é obtida a partir da matriz identidade I_n aplicando-se a mesma sequência de operações elementares necessárias para reduzir A a I_n.*

Demonstração: Pelo teorema anterior, se A é inversível então $A \sim I_n$ e, portanto, existem matrizes elementares $E_1, \dots, E_r$ tais que $E_r \dots E_1 A = I_n$. Pela unicidade da inversa segue que $A^{-1} = E_r \dots E_1 = E_r \dots E_1 I_n$. ■

Método de obtenção da inversa via escalonamento: Posicione as matrizes A e I_n lado a lado, de modo a formar a matriz $(A \vdots I_n)$. As operações elementares nessa matriz produzem operações idênticas em A e I_n. Se A for inversível, então $(A \vdots I_n) \sim (I_n \vdots A^{-1})$, sob as mesmas operações necessárias para reduzir A a I_n.

Exercício 1.43. *Determine, se possível, a inversa das matrizes abaixo por escalonamento de matrizes.*

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & 3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

Resposta: $A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix}, \quad B^{-1} =$

Neste momento estamos aptos a provar parcialmente o item (h) da Proposição 1.39, apenas no caso em que A e B são matrizes inversíveis. Comecemos com o seguinte lema:

Lema 1.44. *Se E for uma matriz elementar de ordem n e A uma matriz qualquer de mesma ordem, então* $\det(EA) = \det(E)\det(A)$.

Demonstração: Os determinantes nos itens abaixo seguem da Proposição 1.39. Analisemos todos os tipos de matrizes elementares :

(i) Se E é obtida de I_n por uma permutação de linhas segue que $\det(E) = -1$. Logo, $\det(EA) = -\det(A) = (-1)\det(A) = \det(E)\det(A)$;

(ii) Se E é obtida de I_n pela multiplicação de uma das linhas por uma constante não nula k, então $\det(E) = k$. Segue que $\det(EA) = k\det(A) = \det(E)\det(A)$;

(iii) Se E é obtida de I_n pela substituição da i-ésima linha por ela própria mais k vezes a j-ésima linha, então $\det(E) = \det(I) = 1$. Logo, $\det(EA) = \det(A) = 1\det(A) = \det(E)\det(A)$. ■

Teorema 1.45. *Sejam A e B matrizes inversíveis. Então*

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

Demonstração: Pelo Teorema 1.41, como A é inversível, A é um produto de matrizes elementares $E_1, \ldots, E_k$. Logo, $\det(A) = \det(E_1 \ldots E_k)$ e $\det(AB) = \det(E_1 \ldots E_k B)$. Aplicando o lema anterior sucessivas vezes obtemos que

$$\det(A) = \det(E_1)\det(E_2 \ldots E_k) = \ldots = \det(E_1)\det(E_2)\ldots\det(E_k) \qquad \text{e}$$

$$\det(AB) = \det(E_1)\det(E_2 \ldots E_k B) = \ldots = \det(E_1)\det(E_2)\ldots\det(E_k)\det(B),$$

donde segue o resultado. ■

A prova para o caso em que A ou B não é inversível será feita no Teorema 1.51.

Uma outra maneira de calcularmos a inversa de uma matriz é através da matriz adjunta.

Definição 1.46. *Seja $A = (a_{ij})_{n\times n}$ uma matriz de ordem n. Para cada $1 \leq i, j \leq n$, definimos o **cofator** da entrada a_{ij} como o escalar*

$$a'_{ij} = (-1)^{i+j}\det(A_{ij}),$$

*onde A_{ij} é a submatriz de ordem $n-1$ obtida retirando-se a i-ésima linha e a j-ésima coluna de A. A **matriz cofatora** (ou **matriz dos cofatores**) de A, denotada por $Cof(A)$, é a matriz de ordem n cujas entradas são os cofatores a'_{ij}. A **matriz adjunta** de A, denotada por $Adj(A)$, é definida como a transposta de $Cof(A)$.*

Exemplo 1.47. Considere novamente a matriz

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Os cofatores de cada a_{ij} são: $a'_{11} = (-1)^{1+1}\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$, $a'_{12} = (-1)^{1+2}\begin{vmatrix} 3 & 0 \\ 1 & 1 \end{vmatrix} = -3$, $a'_{13} = (-1)^{1+3}\begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} = 2$, $a'_{21} = (-1)^{2+1}\begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} = 0$, $a'_{22} = (-1)^{2+2}\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1$, $a'_{23} = (-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1$, $a'_{31} = (-1)^{3+1}\begin{vmatrix} 2 & 2 \\ 1 & 0 \end{vmatrix} = -2$, $a'_{32} = (-1)^{3+2}\begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} = 6$, $a'_{33} = (-1)^{3+3}\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} = -5$. Por definição, $\mathrm{Cof}(A) = (a'_{ij})$ e $Adj(A) = \mathrm{Cof}(A)^t$, ou seja,

$$\mathrm{Cof}(A) = \begin{pmatrix} 1 & -3 & 2 \\ 0 & -1 & 1 \\ -2 & 6 & -5 \end{pmatrix} \quad \text{e} \quad Adj(A) = \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix}.$$

Proposição 1.48. *Para toda matriz A de ordem n,*

$$A(Adj(A)) = \det(A)I_n.$$

Demonstração: Denotemos $A = (a_{ij})_{n\times n}$ e $\mathrm{Cof}(A) = (a'_{kj})_{n\times n}$. Pela Definição 1.46, $a'_{kj} = (-1)^{k+j}\det(A_{kj})$, onde A_{kj} é a submatriz de ordem $n-1$ obtida retirando-se a k-ésima linha e a j-ésima coluna de A. Como $Adj(A) = \mathrm{Cof}(A)^t$ temos

$$A(Adj(A)) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} a'_{11} & \dots & a'_{n1} \\ \vdots & \dots & \vdots \\ a'_{1n} & \dots & a'_{nn} \end{pmatrix},$$

de modo que se denotarmos $A(Adj(A)) = (d_{ik})_{n\times n}$, teremos

$$d_{ik} = a_{i1}a'_{k1} + a_{i2}a'_{k2} + \ldots + a_{in}a'_{kn}, \quad \forall\ 1 \le i, k \le n.$$

Se $i = k$, então

$$d_{ii} = \sum_{j=1}^{n} a_{ij}a'_{ij} = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det(A_{ij}) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) = \det(A),$$

onde a última igualdade segue pelo desenvolvimento de Laplace. Se $i \neq k$, então

$$d_{ik} = \sum_{j=1}^{n} a_{ij}a'_{kj} = \sum_{j=1}^{n} a_{ij}(-1)^{k+j}\det(A_{kj}) = \sum_{j=1}^{n}(-1)^{k+j}a_{ij}\det(A_{kj}) = \det(D),$$

onde D é a matriz obtida de A substituindo a linha L_k pela linha L_i. Observe que a última igualdade da equação acima segue pelo desenvolvimento de Laplace, fixando a k-ésima linha da matriz D. Por outro lado, como D possui duas linhas iguais, temos $\det(D) = 0$, ou seja, $d_{ik} = 0$ sempre que $i \neq k$. Portanto,

$$A(Adj(A)) = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \dots & \vdots \\ d_{n1} & \dots & d_{nn} \end{pmatrix} = \begin{pmatrix} \det(A) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \det(A) \end{pmatrix} = \det(A)I_n.$$

De modo análogo, é possível provar que $(Adj(A))A = \det(A)I_n$. ■

Teorema 1.49. *Uma matriz quadrada A é inversível se, e somente se,* $\det(A) \neq 0$. *Nesse caso,*

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

Demonstração: Suponha que A seja inversível, ou seja, existe uma matriz A^{-1} tal que $AA^{-1} = I_n$. Pelo Teorema 1.45,

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}),$$

o que implica em $\det(A) \neq 0$. Reciprocamente, suponha que $\det(A) \neq 0$. Pela proposição anterior,

$$A\frac{1}{\det(A)} Adj(A) = I_n,$$

o que por definição implica que A é inversível e, pela unicidade da inversa,

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

■

Exemplo 1.50. Considere a matriz A do Exemplo 1.47, cuja matriz adjunta é

$$Adj(A) = \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix}.$$

Como $\det(A) = -1$, temos que

$$A^{-1} = \frac{1}{\det(A)} Adj(A) = \frac{1}{-1} \begin{pmatrix} 1 & 0 & -2 \\ -3 & -1 & 6 \\ 2 & 1 & -5 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 2 \\ 3 & 1 & -6 \\ -2 & -1 & 5 \end{pmatrix},$$

que coincide com a inversa obtida no Exercício 1.43.

Agora estamos preparados para finalizar a prova do item (h) da Proposição 1.39.

Teorema 1.51. *Para quaisquer matrizes quadradas A e B,*

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

Demonstração: A demonstração para o caso em que A e B são inversíveis foi feita no Teorema 1.45. Suponhamos que A não seja inversível, ou seja, $\det(A) = 0$. Se AB fosse inversível, existiria uma matriz C tal que $A(BC) = (AB)C = I_n$. Nesse caso, por definição, A seria inversível, o que é uma contradição. Portanto, AB não é inversível e, pelo Teorema 1.49, $\det(AB) = 0$. Como $\det(A) = 0$, o resultado segue. O mesmo vale se supormos que B não é inversível. ■

1.7 Regra de Cramer

A determinação da inversa de uma matriz via matriz adjunta nos fornece um outro método de resolução de sistemas lineares, denominado **regra de Cramer**, que só se aplica a sistemas cujo número de equações coincide com o número de incógnitas. Por esse método, a solução de um sistema linear quadrado é dada em termos dos determinantes de matrizes específicas.

Comecemos considerando um sistema de n equações e n incógnitas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}, \tag{1.4}$$

que pode ser escrito matricialmente como $CX = B$, onde C é a matriz dos coeficientes, B é a matriz dos termos independentes e X é a matriz das variáveis. Seja C_i a matriz obtida de C substituindo-se a i-ésima coluna de C pela matriz coluna B. Então temos o seguinte:

Teorema 1.52. **(Regra de Cramer)** *O sistema (1.4) admite solução única se, e somente se,* $\det(C) \neq 0$. *Nesse caso, a solução do sistema é dada por*

$$x_1 = \frac{\det(C_1)}{\det(C)}, \quad x_2 = \frac{\det(C_2)}{\det(C)}, \quad \ldots, \quad x_n = \frac{\det(C_n)}{\det(C)}.$$

Demonstração: Suponha que o sistema $CX = B$ admita uma única solução e seja $A = (C \vdots B)$ a matriz ampliada do sistema. Seja $A' = (C' \vdots B')$ a matriz na forma escada linha equivalente à matriz ampliada A. Pelo Teorema 1.30, o sistema $C'X = B'$ também possui uma única solução e, portanto, $p_{C'} = n$. Como C' está na forma escada, temos necessariamente que $C' = I_n$. Logo, $C \sim C' = I_n$. Pelo Teorema 1.41, a matriz C é inversível e, portanto, $\det(C) \neq 0$.

Reciprocamente, se $\det(C) \neq 0$, então C é inversível e, nesse caso,

$$X = I_n X = (C^{-1}C)X = C^{-1}(CX) = C^{-1}B,$$

sendo essa a única solução do sistema $CX = B$.

Nosso objetivo agora é usarmos a igualdade $X = C^{-1}B$ para obtermos a solução explícita do sistema. Como $C^{-1} = \dfrac{1}{\det(C)} Adj(C)$, podemos escrever

$$X = \frac{1}{\det(C)} (\mathrm{Cof}(C))^t B. \tag{1.5}$$

Denotando $\mathrm{Cof}(C) = (c_{ij})$, podemos reescrever (1.5) como

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(C)} \begin{pmatrix} c_{11} & \dots & c_{n1} \\ \vdots & \dots & \vdots \\ c_{1n} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

de onde segue que

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(C)} \begin{pmatrix} c_{11}b_1 + \ldots + c_{n1}b_n \\ \vdots \quad \dots \quad \vdots \\ c_{1n}b_1 + \ldots + c_{nn}b_n \end{pmatrix}. \tag{1.6}$$

Lembremos que, por definição de matriz cofatora, $c_{ij} = (-1)^{i+j}\det(C_{ij})$, onde C_{ij} é a submatriz de ordem $n-1$ obtida de C retirando-se sua i-ésima linha e j-ésima coluna, para todo $1 \leq i, j \leq n$. Agora, vamos calcular o determinante da matriz C_j obtida substituindo-se a j-ésima coluna de C pela matriz coluna B, ou seja,

$$C_j = \begin{pmatrix} a_{11} & \dots & a_{1j-1} & b_1 & a_{1j+1} & \dots & a_{1n} \\ \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nj-1} & b_n & a_{nj+1} & \dots & a_{nn} \end{pmatrix}.$$

Observe que retirando-se a i-ésima linha e a j-ésima coluna de C_j obteremos exatamente a matriz C_{ij} definida acima. Portanto, aplicando o desenvolvimento de Laplace fixando a coluna j da matriz C_j, obtemos que

$$\det(C_j) = \sum_{i=1}^{n} (-1)^{i+j} b_i \det(C_{ij}) = \sum_{i=1}^{n} c_{ij} b_i = c_{1j}b_1 + \ldots + c_{nj}b_n,$$

para todo $1 \leq j \leq n$. Por (1.6), segue que

$$x_i = \frac{c_{1i}b_1 + \ldots + c_{ni}b_n}{\det(C)} = \frac{\det(C_i)}{\det(C)}$$

para todo $i = 1, \ldots, n$. ■

Corolário 1.53. *Um sistema homogêneo $AX = 0$ possui infinitas soluções se, e somente se,* $\det(A) = 0$.

Demonstração: Primeiro lembremos que um sistema homogêneo sempre possui solução. Pelo teorema anterior, o sistema $AX = 0$ possui somente a solução $X = 0$ se, e somente se, $\det(A) \neq 0$. Portanto, o sistema possui outras soluções além da trivial se, e somente se, $\det(A) = 0$. ■

Exemplo 1.54. Considere novamente o sistema linear

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

apresentado na Seção 1.4. Vimos anteriormente que $(x_1, x_2, x_3) = (7, 6, -10)$ é uma solução desse sistema e nos questionamos se tal sistema admitia mais soluções. Vamos tentar resolvê-lo usando a Regra de Cramer. Primeiramente, calculamos o determinante da matriz dos coeficientes

$$C = \begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 4 \\ 1 & -3 & -2 \end{pmatrix},$$

obtendo $\det(C) = 1 \neq 0$. Pelo Teorema 1.52 o sistema acima admite uma única solução dada por

$$x_1 = \frac{\det(C_1)}{\det(C)}, \quad x_2 = \frac{\det(C_2)}{\det(C)}, \quad x_3 = \frac{\det(C_3)}{\det(C)},$$

onde

$$C_1 = \begin{pmatrix} 1 & 4 & 3 \\ 4 & 5 & 4 \\ 9 & -3 & -2 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 4 & 4 \\ 1 & 9 & -2 \end{pmatrix}, C_3 = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 5 & 4 \\ 1 & -3 & 9 \end{pmatrix}.$$

Portanto,

$$x_1 = \det(C_1) = -10 + 144 - 36 - 135 + 32 + 12 = 7,$$

$$x_2 = \det(C_2) = -8 + 4 + 54 - 12 + 4 - 36 = 6,$$

$$x_3 = \det(C_3) = 45 + 16 - 6 - 5 - 72 + 12 = -10.$$

Capítulo 2

Espaços Vetoriais

Neste capítulo vamos explorar os conceitos relacionados a espaços vetoriais, também conhecidos como espaços lineares. Veremos que o conceito de vetor é muito mais amplo do que o abordado na disciplina de geometria analítica, mas agora sem a interpretação geométrica possível lá. A ideia é caracterizar os vários conjuntos mais abstratos que possuem uma estrutura algébrica parecida com a dos conjuntos de vetores no plano e no espaço, e essa abordagem nos permite fazer uma análise sistemática de todos esses casos.

2.1 Espaços e Subespaços Vetoriais

Para o que segue, seja V um conjunto não vazio munido de duas operações binárias

$$\begin{array}{rcl} + : V \times V & \to & V \\ (u,\ v) & \mapsto & u+v \end{array} \qquad\qquad \begin{array}{rcl} \cdot : \mathbb{K} \times V & \to & V \\ (k,\ v) & \mapsto & kv \end{array},$$

onde $\mathbb{K}$ é um corpo qualquer. Para nossos propósitos, assumimos $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Definição 2.1. *Dizemos que um conjunto não vazio V munido das operações "$+$" e "$\cdot$" é um **espaço vetorial** sobre $\mathbb{K}$ (ou um $\mathbb{K}$-espaço vetorial) se, para quaisquer $u, v, w \in V$ e $k, k_1, k_2 \in \mathbb{K}$, tivermos:*

1. *$(u+v)+w = u+(v+w)$;*

2. *$u+v = v+u$;*

3. *Existe $\bar{0} \in V$ tal que $u+\bar{0} = u$. O elemento $\bar{0}$ é único e é chamado de **vetor nulo** de V.*

4. *Existe $u' \in V$ tal que $u+u' = \bar{0}$. O elemento u', que é único para cada $u \in V$, é chamado de **vetor oposto** de u e será denotado por $-u$;*

5. *$k(u+v) = ku+kv$;*

6. $(k_1 + k_2)u = k_1u + k_2u$*;*

7. $k_1(k_2u) = k_1k_2u$*;*

8. $1u = u$*.*

Todos os elementos de um espaço vetorial são denominados **vetores** e os elementos de $\mathbb{K}$ são denominados **escalares**. A seguir, vamos apresentar alguns exemplos de espaços vetoriais.

EXEMPLO 2.2.

1. $\mathbb{R}^2 = \{(x, y);\ x, y \in \mathbb{R}\}$ munido das operações usuais

$$(x, y) + (z, w) := (x + z, y + w) \quad \text{e} \quad k(x, y) := (kx, ky),\ k \in \mathbb{R}$$

é um espaço vetorial sobre $\mathbb{R}$.

2. $\mathbb{R}^2$ munido das operações

$$(x, y) + (z, w) := (x + z, 0) \quad \text{e} \quad k(x, y) := (kx, ky),\ k \in \mathbb{R}$$

não é um espaço vetorial sobre $\mathbb{R}$, pois não possui vetor nulo. De fato, seja $\bar{0} = (a, b)$ tal que $(x, y) + (a, b) = (x, y)$. Por definição, $(x, y) + (a, b) = (x + a, 0)$, ou seja, $(x + a, 0) = (x, y)$, o que é um absurdo já que y é arbitrário.

3. Mais geralmente, $\mathbb{R}^n = \{(x_1, \ldots, x_n);\ x_1, \ldots, x_n \in \mathbb{R}\}$ munido das operações usuais

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

$$k(x_1, \ldots, x_n) := (kx_1, \ldots, kx_n),\ k \in \mathbb{R}$$

é um espaço vetorial sobre $\mathbb{R}$.

4. Seja $\mathbb{C}^n = \{(z_1, \ldots, z_n);\ z_1, \ldots, z_n \in \mathbb{C}\}$ munido das operações

$$(z_1, \ldots, z_n) + (w_1, \ldots, w_n) := (z_1 + w_1, \ldots, z_n + w_n)$$

$$k(z_1, \ldots, z_n) := (kz_1, \ldots, kz_n),\ k \in \mathbb{C}.$$

Então $\mathbb{C}^n$ é um espaço vetorial sobre $\mathbb{C}$. Observe que podemos considerar $k \in \mathbb{R}$. Nesse caso, $\mathbb{C}^n$ também é um espaço vetorial sobre $\mathbb{R}$.

5. O conjunto $\mathcal{M}_{m\times n}(\mathbb{K})$ das matrizes de ordem $m \times n$ com entradas em $\mathbb{K}$, munido das operações usuais de soma de matrizes e produto de uma matriz por um escalar $k \in \mathbb{K}$, é um espaço vetorial sobre $\mathbb{K}$.

6. Seja X um conjunto não vazio e considere o conjunto $\mathcal{F}(X)$ formado por todas as funções $f : X \to \mathbb{K}$. Note que $\mathcal{F}(X)$ é não vazio, já que X é não vazio. Dados $f, g \in \mathcal{F}(X)$ e $k \in \mathbb{K}$, defina $f + g : X \to \mathbb{K}$ e $kf : X \to \mathbb{K}$ como

$$(f+g)(x) := f(x) + g(x) \quad \text{e} \quad (kf)(x) := kf(x).$$

   Com essas operações temos que a função nula definida por $\bar{0}(x) = 0$ para todo $x \in X$ é o vetor nulo de $\mathcal{F}(X)$. Além disso, dada $f \in \mathcal{F}(X)$, a função $-f \in \mathcal{F}(X)$ definida por $(-f)(x) = -f(x)$ é o vetor oposto de f. Também é fácil verificar as outras condições para que $\mathcal{F}(X)$ seja um espaço vetorial sobre $\mathbb{K}$. Enfatizamos aqui que X é um conjunto qualquer e que a estrutura de espaço vetorial de $\mathcal{F}(X)$ depende somente das operações de $\mathbb{K}$.

7. Seja $\mathcal{P}(\mathbb{K})$ o conjunto de todos os polinômios

$$p(x) = a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0,$$

   com coeficientes $a_i \in \mathbb{K}$ e qualquer grau $n \geq 0$. Sejam $n, m \in \mathbb{N}$, com $n \geq m$, $k \in \mathbb{K}$ e $p(x), q(x) \in \mathcal{P}(\mathbb{K})$ dados por

$$p(x) = a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0 \quad \text{e} \quad q(x) = b_m x^m + \ldots + b_2 x^2 + b_1 x + b_0.$$

   Defina as operações

$$p(x)+q(x) := a_n x^n + \ldots + a_{m+1} x^{m+1} + (a_m + b_m) x^m + \ldots + (a_1 + b_1) x + (a_0 + b_0),$$
$$kp(x) := ka_n x^n + \ldots + ka_2 x^2 + ka_1 x + ka_0.$$

   Sob essas operações, $\mathcal{P}(\mathbb{K})$ é um espaço vetorial sobre $\mathbb{K}$.

Exercício 2.3. *Mostre que se V é um espaço vetorial sobre K, então $-v = -1v$ para todo $v \in V$.*

Podemos agora nos perguntar se todos os subconjuntos S de um espaço vetorial V herdam sua estrutura de espaço vetorial, satisfazendo todas as oito propriedades da Definição 2.1. Erroneamente, a resposta para essa questão parece ser positiva, uma vez que se as oito propriedades da Definição 2.1 são válidas em V, então são válidas em $S \subseteq V$. No entanto, a resposta é negativa e para compreendê-la introduzimos o seguinte:

Definição 2.4. *Seja V um espaço vetorial sobre um corpo $\mathbb{K}$. Um* ***subespaço vetorial*** *de V é um subconjunto não vazio de V que é um espaço vetorial sobre $\mathbb{K}$ com as operações de adição e multiplicação por escalar induzidas de V.*

O resultado a seguir nos dá um método para verificarmos quando um subconjunto de um espaço vetorial é um subespaço vetorial, sem a necessidade de checarmos as oito propriedades da Definição 2.1.

PROPOSIÇÃO 2.5. *Um subconjunto não vazio S de V é um subespaço vetorial de V se, e somente se, para todo $u, v \in S$ e $k \in \mathbb{K}$ tivermos $ku + v \in S$.*

DEMONSTRAÇÃO: Claramente se S é um subespaço vetorial de V, então dados $u, v \in S$ e $k \in \mathbb{K}$ temos $u + v \in S$ e $ku \in S$. Logo, $ku + v \in S$.

Reciprocamente, suponhamos que S seja um subconjunto não vazio de V tal que $ku + v \in S$ para todo $u, v \in S$ e $k \in \mathbb{K}$. Como S é não vazio, existe $v_0 \in S$. Logo, $\bar{0} = -v_0 + v_0 = -1v_0 + v_0 \in S$, por hipótese. Também por hipótese, se $u, v \in S$ e $k \in \mathbb{K}$, então $u+v = 1u+v \in S$ e $ku = ku+\bar{0} \in S$. Portanto, as operações "$+$" e "$\cdot$" estão bem definidas em S. Claramente, as oito propriedades da Definição 2.1 são satisfeitas para os elementos de S, pois são satisfeitas para todos os elementos de V e $S \subseteq V$. Logo, S é um subespaço vetorial de V. ■

Pela demonstração da proposição anterior, todo subespaço de um espaço vetorial V contém o vetor nulo de V. Deste modo, todo espaço vetorial admite pelo menos dois subespaços, chamados **triviais**: V e $\{\bar{0}\}$. A seguir, exploraremos alguns exemplos utilizando a Proposição 2.5.

EXEMPLO 2.6.

1. $S = \{(x, x);\ x \in \mathbb{R}\}$ é um subespaço vetorial de $\mathbb{R}^2$. De fato, primeiro observe que $\bar{0} = (0, 0) \in S$, o que implica que S é não vazio. Além disso, se $u = (x, x), v = (y, y) \in S$ e $k \in \mathbb{K}$ então
$$ku + v = k(x, x) + (y, y) = (kx + y, kx + y) \in S.$$

2. $S = \{(x, y, z) \in \mathbb{R}^3;\ x = 0 \text{ e } y \geq z\}$ não é um subespaço vetorial de $\mathbb{R}^3$ com as operações usuais. De fato, considere $u \in S$ e $k \in \mathbb{R}$. Então, $u = (0, y, z)$, com $y, z \in \mathbb{R}$ e $y \geq z$. Se $k > 0$ então $ky \geq kz$. Entretanto, se $k < 0$ teremos $ky \leq kz$, o que implica que
$$ku = k(0, y, z) = (0, ky, kz) \notin S.$$
Logo, S não é um subespaço vetorial de $\mathbb{R}^3$.

3. $S = \{(x, y, z) \in \mathbb{R}^3;\ x = y + 1\}$ não é um subespaço vetorial de $\mathbb{R}^3$, pois $\bar{0} = (0, 0, 0) \notin S$.

4. $S = \{X \in \mathcal{M}_{n\times 1}(\mathbb{R});\ AX = 0 \text{ com } A \in \mathcal{M}_{n\times n}(\mathbb{R})\}$ é um subespaço vetorial de $\mathcal{M}_{n\times 1}(\mathbb{R})$, ou seja, o conjunto das soluções de um sistema homogêneo determina um subespaço vetorial do espaço das matrizes $\mathcal{M}_{n\times 1}(\mathbb{R})$.

   De fato, como o sistema $AX = 0$ é um sistema homogêneo segue que S é não vazio, pois a solução nula $X = 0$ é uma solução do sistema. Logo, $\bar{0} \in S$. Agora, dados $X, Y \in S$ e $k \in \mathbb{R}$ temos que
$$A(kX + Y) = A(kX) + AY = kAX + AY = k0 + 0 = 0.$$

Logo, $kX + Y \in S$, ou seja, S é um subespaço vetorial de $\mathcal{M}_{n\times 1}(\mathbb{R})$. Isso implica, entre outras coisas, que a soma e a multiplicação por escalar de quaisquer duas soluções de um sistema homogêneo é também uma solução desse sistema.

5. O conjunto $S = \{A \in \mathcal{M}_{n\times n}(\mathbb{K});\ A = A^t\}$ das matrizes simétricas de ordem n com entradas em $\mathbb{K}$ é um subespaço de $\mathcal{M}_{n\times n}(\mathbb{K})$. De fato, S é não vazio pois a matriz nula $\bar{0}$ de ordem n é uma matriz simétrica. Logo, $\bar{0} \in S$. Além disso, dados $A, B \in S$ e $k \in \mathbb{K}$ temos que

$$(kA + B)^t = (kA)^t + B^t = kA^t + B^t = kA + B,$$

ou seja, $kA + B \in S$. Portanto, S é um subespaço vetorial de $\mathcal{M}_{n\times n}(\mathbb{K})$.

6. O conjunto das funções contínuas $f : X \to \mathbb{K}$ é um subespaço do espaço vetorial $\mathcal{F}(X)$, definido no Exemplo 2.2. Isso ocorre pois a soma de quaisquer duas funções contínuas é ainda uma função contínua, o mesmo valendo para o produto de uma função contínua por um escalar $k \in \mathbb{K}$.

7. Seja $n \in \mathbb{N}$ fixado e considere o subconjunto

$$\mathcal{P}_n(\mathbb{K}) = \{a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0;\ a_i \in \mathbb{K}\} \subset \mathcal{P}(\mathbb{K})$$

constituído por todos os polinômios de grau $\leq n$. Então $\mathcal{P}_n(\mathbb{K})$ é um subespaço vetorial de $\mathcal{P}(\mathbb{K})$. Com efeito, se $k \in \mathbb{K}$ e $p(x), q(x) \in \mathcal{P}_n(\mathbb{K})$ são polinômios de grau $m \leq n$ e $r \leq n$, respectivamente, então

$$\text{grau } (kp(x) + q(x)) \leq \max\{\text{grau } kp(x), \text{ grau } q(x)\} = \max\{m, r\} \leq n.$$

(No caso em que $k = 0$ o grau do polinômio $kp(x)$ não está definido, podendo ser considerado qualquer grau). Portanto, $kp(x) + q(x) \in \mathcal{P}_n(\mathbb{K})$, como desejado.

OBSERVAÇÃO 2.7. *O conjunto dos polinômios de grau n **não** é um subespaço vetorial de $\mathcal{P}(\mathbb{K})$, pois a soma de dois polinômios de grau n pode ter grau estritamente menor que n.*

2.2 Espaços Gerados

Vamos introduzir nessa seção um conceito chave desta disciplina, o de geradores de um espaço vetorial. A grosso modo, definimos um conjunto gerador de um espaço como um conjunto de vetores capaz de "produzir" todos os vetores desse espaço. Para compreender melhor esse conceito, precisamos da definição de combinação linear. Para nossos propósitos, vamos considerar somente conjuntos geradores por um número finito de vetores.

DEFINIÇÃO 2.8. *Sejam V um $\mathbb{K}$-espaço vetorial e $v_1, \ldots, v_n \in V$. Um vetor $u \in V$ é dito uma* ***combinação linear*** *dos vetores $v_1, \ldots, v_n$ se existirem escalares $k_1, \ldots, k_n \in \mathbb{K}$ tais que*

$$u = k_1 v_1 + \cdots + k_n v_n.$$

Os escalares $k_1, \ldots, k_n$ são chamados de ***coeficientes*** *da combinação linear.*

EXEMPLO 2.9.

1. O vetor $u = (-1, -1) \in \mathbb{R}^2$ é combinação linear dos vetores $v_1 = (1, 2)$ e $v_2 = (3, 5)$. De fato, suponha $(-1, -1) = k_1(1, 2) + k_2(3, 5)$. Logo, $k_1 + 3k_2 = -1$ e $2k_1 + 5k_2 = -1$, o qual admite solução $k_1 = 2$ e $k_2 = -1$. Portanto,

$$(-1, -1) = 2(1, 2) + (-1)(3, 5).$$

2. Todos os vetores do plano Oxz podem ser escritos como combinação linear de $(1, 0, 0)$ e $(0, 0, 1)$. De fato,

$$(x, 0, z) = x(1, 0, 0) + z(0, 0, 1).$$

Portanto $\{(x, 0, z);\ x, z \in \mathbb{R}\}$, que é um subespaço vetorial de $\mathbb{R}^3$, é o conjunto de todos os vetores escritos como combinação linear de $(1, 0, 0)$ e $(0, 0, 1)$.

3. A matriz $\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R})$ é escrita como combinação linear de quatro vetores de $\mathcal{M}_{2\times 2}(\mathbb{R})$:

$$\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} = 2\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 4\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

4. O polinômio $p(x) = x^3 + 2x^2 - x + 1$ é combinação linear dos vetores $p_1(x) = x^3$, $p_2(x) = x^2$, $p_3(x) = x$ e $p_4(x) = 1$:

$$p(x) = p_1(x) + 2p_2(x) - p_3(x) + p_4(x).$$

Observe que $p(x)$ não é combinação linear dos vetores $q_1(x) = x^3 - x^2$, $q_2(x) = x^2 - x$ e $q_3(x) = x - 1$.

DEFINIÇÃO 2.10. *Sejam $v_1, \ldots, v_n$ vetores de um espaço vetorial V. Definimos o* ***subespaço vetorial gerado por*** *$v_1, \ldots, v_n$, cuja denotação é $[v_1, \ldots, v_n]$, como o conjunto de todas as combinações lineares de $v_1, \ldots, v_n$, ou seja,*

$$[v_1, \ldots, v_n] = \{\alpha_1 v_1 + \ldots + \alpha_n v_n \in V;\ \alpha_1, \ldots, \alpha_n \in \mathbb{K}\}.$$

Os vetores $v_1, \ldots, v_n$ são chamados de ***geradores*** *de $[v_1, \ldots, v_n]$.*

Exercício 2.11. *Mostre que $S = [v_1, \ldots, v_n]$ é de fato um subespaço vetorial de V.*

Observe que a quantidade de vetores que formam o conjunto gerador $\{v_1, \ldots, v_n\}$ é limitada, mas a quantidade de combinações lineares que determinam $[v_1, \ldots, v_n]$ é ilimitada. Neste caso, dizemos que $S = [v_1, \ldots, v_n]$ é **finitamente gerado**.

Exemplo 2.12. 1. $[(1,3,2),(-1,0,1)] = \{k_1(1,3,2) + k_2(-1,0,1);\ k_1, k_2 \in \mathbb{R}\}$
$= \{(k_1 - k_2, 3k_1, 2k_1 + k_2);\ k_1, k_2 \in \mathbb{R}\}.$

2. $[1, x, x^2] = \{\alpha 1 + \beta x + \gamma x^2;\ \alpha, \beta, \gamma \in \mathbb{R}\} = \mathcal{P}_2(\mathbb{R})$.

3. $[(1,0),(0,1)] = \{\alpha(1,0) + \beta(0,1);\ \alpha, \beta \in \mathbb{R}\} = \{(\alpha, \beta);\ \alpha, \beta \in \mathbb{R}\} = \mathbb{R}^2$.

4. $\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] = \left\{x \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + z \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + w \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix};\ x, y, z, w \in \mathbb{R}\right\} = \left\{\begin{pmatrix} x & y \\ z & w \end{pmatrix};\ x, y, z, w \in \mathbb{R}\right\} = \mathcal{M}_{2\times 2}(\mathbb{R})$.

5. Vamos encontrar um conjunto de geradores para o subespaço vetorial de $\mathbb{R}^4$ dado por $U = \{(x,y,z,t) \in \mathbb{R}^4;\ x - y + t + z = 0\}$. Observe que

$$(x,y,z,t) = (x, x+t+z, z, t) = x(1,1,0,0) + z(0,1,1,0) + t(0,1,0,1).$$

Logo, $U = [(1,1,0,0),(0,1,1,0),(0,1,0,1)]$

2.3 Independência Linear

Vimos na seção anterior que cada vetor em um espaço vetorial pode ser "construído" a partir dos elementos de um conjunto gerador usando-se apenas as operações de soma e multiplicação por um escalar. Neste caso, seria conveniente encontrar um conjunto gerador "mínimo" do espaço. Por "mínimo" queremos dizer um conjunto gerador sem vetores desnecessários, ou seja, no qual todos os vetores são necessários para se gerar o espaço todo. Para encontrar um conjunto gerador mínimo, é preciso considerar como os vetores no conjunto dependem um do outro.

Definição 2.13. *Um conjunto de vetores $\{v_1, \ldots, v_n\}$ de um espaço vetorial V é dito **linearmente independente** (ou LI) se nenhum vetor deste conjunto é combinação linear dos demais. Caso contrário, dizemos que o conjunto é **linearmente dependente** (ou LD).*

Exemplo 2.14. 1. Um conjunto com dois vetores $\{u, v\}$ de um espaço vetorial V é LD se, e somente se, um dos vetores é combinação linear do outro, ou seja, existe $\alpha \in \mathbb{R}$ tal que $u = \alpha v$ (ou $v = \alpha u$). Portanto, $\{u, v\}$ é LI se, e somente se, u e v não são múltiplos.

2. Pelo item anterior, o conjunto $\{(1,0),(0,1)\}$ é LI pois não é possível escrever um deles como múltiplo do outro. No entanto, o conjunto $\{(1,0),(0,1),(1,1)\}$ é LD pois
$$(1,1) = 1(1,0) + 1(0,1).$$

Proposição 2.15. *Seja V um $\mathbb{K}$-espaço vetorial.*

1. *Todo conjunto $L \subset V$ que contém o vetor nulo é um conjunto LD.*

2. *Todo conjunto que contém um conjunto LD é também LD.*

3. *Todo subconjunto de um conjunto LI é ainda LI.*

4. *Um conjunto S de vetores é LI se, e somente se, todo subconjunto finito de S é LI.*

Demonstração: *1*. Observe que o vetor nulo é uma combinação linear de quaisquer vetores $v_1, \ldots, v_n$ de L, pois $\bar{0} = 0v_1 + \ldots + 0v_n$.

2. Se S é um conjunto que contém um subconjunto S' que é LD, então existe $v \in S'$ (em particular em S) que é uma combinação linear de outros vetores de S' e, portanto, de S.

3. Seja S um conjunto LI que contém um subconjunto S'. Se S' fosse LD, então pelo item anterior S também seria LD, o que contraria a hipótese. Logo, S' é LI.

4. A ida é válida pelo item *3*. A recíproca será provada por contrapositva. Suponha que S seja LD. Logo, existe algum vetor $v \in S$ que é uma combinação linear finita de certos vetores $v_1, \ldots, v_n \in S$. Logo, o subconjunto $S' = \{u, v_1, \ldots, v_n\} \subset S$ é LD, ou seja, S admite um subconjunto finito que é LD. ■

O resultado abaixo nos apresenta um método para verificarmos se um conjunto é linearmente independente ou não.

Proposição 2.16. *Um conjunto $\{v_1, \ldots, v_n\}$ de um espaço vetorial V é LI se, e somente se, toda combinação linear trivial*
$$a_1v_1 + \ldots + a_nv_n = \bar{0},$$
com $a_i \in \mathbb{K}$, implicar que $a_1 = \ldots = a_n = 0$.

Demonstração: Vamos provar as duas implicações por contrapositiva. Suponha que exista um escalar $a_i \in \mathbb{K}$ não nulo satisfazendo $a_1v_1 + \ldots + a_iv_i + \ldots + a_nv_n = \bar{0}$. Logo,
$$v_i = -\frac{a_1}{a_i}v_1 - \ldots - \frac{a_{i-1}}{a_i}v_{i-1} - \frac{a_{i+1}}{a_i}v_{i+1} - \ldots - \frac{a_n}{a_i}v_n,$$
ou seja, v_i é uma combinação linear dos demais vetores e, por definição, o conjunto $\{v_1, \ldots, v_n\}$ é LD.

Para a recíproca, suponha que $\{v_1, \ldots, v_n\}$ é LD. Assim existe pelo menos um v_i que é combinação linear dos demais vetores, ou seja, existem $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n \in \mathbb{K}$ tais que

$$v_i = a_1v_1 + \ldots + a_{i-1}v_{i-1} + a_{i+1}v_{i+1} + \ldots + a_nv_n.$$

Neste caso,

$$a_1v_1 + \ldots + a_{i-1}v_{i-1} - v_i + a_{i+1}v_{i+1} + \ldots + a_nv_n = \bar{0},$$

ou seja, existe uma combinação linear nula dos vetores $v_1, \ldots, v_n$ com pelo menos um escalar não nulo (a saber, $a_i = -1$). ■

Exemplo 2.17.

1. O conjunto $\{(1, 3, 2), (-1, 0, 1)\}$ é LI. De fato, considere a combinação linear

$$\alpha(1, 3, 2) + \beta(-1, 0, 1) = \bar{0}.$$

Então $(\alpha - \beta, 3\alpha, 2\alpha + \beta) = (0, 0, 0)$, donde segue que $\alpha = \beta = 0$.

2. O conjunto $\{(1, 0), (0, 1)\}$ é LI, pois se

$$\alpha(1, 0) + \beta(0, 1) = (\alpha, \beta) = \bar{0},$$

então $\alpha = \beta = 0$.

3. O conjunto $\{(1, 1, 0, 0), (0, 1, 1, 0), (0, 1, 0, 1), (2, 4, 2, 0)\}$ é LD, pois existe uma combinação linear nula desses elementos em que os coeficientes não são todos nulos. De fato, para cada $\alpha \in \mathbb{R}$ temos

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

onde apenas o terceiro coeficiente da combinação é nulo se $\alpha \neq 0$.

4. O conjunto $\left\{\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}\right\}$ é LD, pois para cada $\alpha \in \mathbb{R}$ temos

$$\alpha\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \alpha\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} - \alpha\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} - \alpha\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

onde todos os coeficientes da combinação podem ser não nulos.

5. O conjunto $\{1, x, \ldots, x^{n-1}, x^n\}$ em $\mathcal{P}_n(\mathbb{K})$ é LI. De fato, considere uma combinação linear qualquer

$$a_01 + a_1x + \ldots + a_nx^n = \bar{0}.$$

Como $\bar{0} = 0 + 0x + \ldots + 0x^n$, por igualdade de polinômios segue que $a_0 = \ldots = a_n = 0$.

6. O conjunto $\{1, x, \ldots, x^n, \ldots\}$ em $\mathcal{P}(\mathbb{K})$ é LI. Esse fato segue diretamente do item *4.* da Proposição 2.15, já que todo conjunto finito formado por monômios de graus distintos é LI.

Proposição 2.18. *Sejam $v_1, \dots, v_n \in \mathbb{K}^n$. O conjunto $\{v_1, \dots, v_n\}$ é LI se, e somente se, o determinante da matriz cujas colunas são formadas pelos vetores v_i é não nulo.*

Demonstração: Escreva, primeiramente, cada vetor v_i como $v_i = (a_{1i}, \dots, a_{ni}) \in \mathbb{K}^n$, para $1 \leq i \leq n$. Queremos mostrar que $\{v_1, \dots, v_n\}$ é LI se, e somente se, o determinante da matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

é não nulo. Para isso, vamos usar a Proposição 2.16. Considere a combinação linear nula $x_1v_1 + \ldots + x_nv_n = \bar{0}$ com coeficientes $x_1, \dots, x_n \in \mathbb{K}$, ou seja

$$x_1(a_{11}, \dots, a_{n1}) + \ldots + x_n(a_{1n}, \dots, a_{nn}) = (0, \dots, 0).$$

Determinar os valores de $x_1, \dots, x_n$ para os quais tal igualdade ocorra é equivalente a resolver o sistema linear homogêneo com n equações e n variáveis

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n & = & 0 \\ \vdots & & \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n & = & 0 \end{cases},$$

que pode ser escrito matricialmente como $AX = 0$, onde A é a matriz dos coeficientes, X é a matriz das variáveis e 0 é a matriz nula. Pela Proposicão 2.16, $\{v_1, \dots, v_n\}$ é LI se, e somente se, o sistema $AX = 0$ admite solução única $(x_1, \dots, x_n) = (0, \dots, 0)$. E pela regra de Cramer, $AX = 0$ admite solução única se, e somente se, $\det(A) \neq 0$. ■

2.4 Bases e Dimensão

Na seção anterior começamos nossa busca por um conjunto gerador que possua um número "mínimo" de vetores, ou seja, um conjunto LI. E essa busca é fundamental pois se um conjunto de vetores gerar um espaço vetorial V e for LD, então a representação de todo $v \in V$ em termos desses vetores não é única. Para se ter unicidade, o conjunto gerador deve ser também LI e, quando isso ocorre, damos a ele o nome de base. Com a base de um espaço vetorial é possível compreender toda a sua estrutura algébrica.

Definição 2.19. *Sejam V um $\mathbb{K}$-espaço vetorial e $\mathcal{B} \subset V$ um conjunto não vazio. Dizemos que $\mathcal{B}$ é uma **base** de V se $\mathcal{B}$ for LI e $[\mathcal{B}] = V$, ou seja, o conjunto gerado por $\mathcal{B}$ é todo o espaço.*

Pela definição acima se $\mathcal{B} = \{v_1, \dots, v_n\}$ for uma base de um espaço vetorial V, então qualquer vetor $v \in V$ se escreverá de maneira única como combinação linear de $v_1, \dots, v_n$.

De fato, primeiramente como $[\mathcal{B}] = V$ temos que $v \in [\mathcal{B}]$, ou seja, existem escalares $k_1, \ldots, k_n$ tais que $v = k_1v_1 + \cdots + k_nv_n$. Suponha agora que existam $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tais que $v = \lambda_1v_1 + \cdots + \lambda_nv_n$. Neste caso,

$$k_1v_1 + \cdots + k_nv_n = \lambda_1v_1 + \cdots + \lambda_nv_n,$$

o que implica que $(k_1 - \lambda_1)v_1 + \ldots + (k_n - \lambda_n)v_n = \bar{0}$. Como $\mathcal{B}$ é um conjunto LI, segue que $k_i - \lambda_i = 0$ para todo $i = 1, \ldots, n$ e, portanto, $k_i = \lambda_i$. Assim, os escalares da combinação $v = k_1v_1 + \cdots + k_nv_n$ são únicos.

Exemplo 2.20. 1. Por convenção, o espaço vetorial nulo $V = \{\bar{0}\}$ tem como base o conjunto vazio.

2. O conjunto $\{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$ constitui uma base de $\mathbb{R}^n$. De fato, dado $(x_1, \ldots, x_n) \in \mathbb{R}^n$ temos

$$(x_1, \ldots, x_n) = x_1(1, 0, \ldots, 0) + \ldots + x_n(0, 0, \ldots, 1) \in [(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)].$$

Logo, $\mathbb{R}^n \subseteq [(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)]$. Como $[(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)] \subseteq \mathbb{R}^n$, segue a igualdade. Além disso, se $\alpha_1(1, 0, \ldots, 0) + \ldots + \alpha_n(0, 0, \ldots, 1) = \bar{0}$, então $(\alpha_1, \ldots, \alpha_n) = (0, \ldots, 0)$, implicando que $\alpha_i = 0$ para todo $1 \leq i \leq n$.

Deste modo, com apenas n vetores podemos encontrar qualquer vetor no espaço euclidiano n-dimensional.

3. O conjunto $\{1, x, \ldots, x^{n-1}, x^n\}$ é uma base de $\mathcal{P}_n(\mathbb{K})$. De fato, claramente

$$[1, x, \ldots, x^n] = \{\alpha_0 1 + \alpha_1 x + \ldots + \alpha_n x^n;\ \alpha_0, \ldots, \alpha_n \in \mathbb{K}\} = \mathcal{P}_n(\mathbb{K}).$$

Além disso, mostramos no Exemplo 2.17 que $\{1, x, \ldots, x^n\}$ é LI e, portanto, esse conjunto é uma base de $\mathcal{P}_n(\mathbb{K})$.

4. Considere as matrizes $A_{ij} = (a_{kl})_{n\times n}$ definidas como

$$a_{kl} = \begin{cases} 1, & \text{se } k = i \text{ e } l = j \\ 0, & \text{caso contrário} \end{cases}.$$

O conjunto $\{A_{ij} = (a_{kl}) \in \mathcal{M}_{n\times n}(\mathbb{K}); 1 \leq i, j \leq n\}$ é uma base de $\mathcal{M}_{n\times n}(\mathbb{K})$.

Todas as bases apresentadas nos itens anteriores são chamadas de **bases canônicas**.

5. O conjunto $\{(1, 2), (-1, 0), (2, 3)\}$ gera $\mathbb{R}^2$, entretanto não é LI. De fato, obviamente $[(1, 2), (-1, 0), (2, 3)] \subseteq \mathbb{R}^2$, pois cada vetor pertence a $\mathbb{R}^2$, que é um $\mathbb{R}$-espaço vetorial. Agora, dado $(x, y) \in \mathbb{R}^2$ temos

$$(x, y) = (-6x + 2y)(1, 2) + x(-1, 0) + (4x - y)(2, 3) \in [(1, 2), (-1, 0), (2, 3)].$$

Logo, $\mathbb{R}^2 \subseteq [(1,2),(-1,0),(2,3)]$, de onde segue que $\mathbb{R}^2 = [(1,2),(-1,0),(2,3)]$. No entanto,

$$(1,2) = \frac{1}{3}(-1,0) + \frac{2}{3}(2,3),$$

o que implica que $\{(1,2),(-1,0),(2,3)\}$ é LD. Portanto $\{(1,2),(-1,0),(2,3)\}$ não é uma base de $\mathbb{R}^2$.

6. O conjunto $\{(1,2),(2,3)\}$ constitui uma base de $\mathbb{R}^2$. De fato, dado $(x,y) \in \mathbb{R}^2$ temos

$$(x,y) = (-3x+2y)(1,2) + (2x-y)(2,3) \in [(1,2),(2,3)].$$

Logo, $\mathbb{R}^2 = [(1,2),(2,3)]$. Agora, considere a matriz $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ cujas colunas são formadas pelos vetores $(1,2)$ e $(2,3)$. Como $\det(A) \neq 0$ segue, pela Proposição 2.18, que $\{(1,2),(2,3)\}$ é LI.

Observe pelos itens 2. e 6. que os conjuntos $\{(1,0),(0,1)\}$ e $\{(1,2),(2,3)\}$ são ambos bases de $\mathbb{R}^2$. Em geral, a base de um espaço vetorial não é única. O próximo resultado nos mostra que um espaço vetorial finitamente gerado por um conjunto de vetores não nulos sempre admite uma base. Mais do que isso, sempre podemos extrair desse conjunto gerador uma base para V.

Proposição 2.21. *Todo $\mathbb{K}$-espaço vetorial V gerado por um conjunto finito de vetores não nulos possui uma base.*

Demonstração: Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ um conjunto formado por vetores não nulos e suponha que $V = [\mathcal{B}]$. A ideia da demonstração é extrair de $\mathcal{B}$ uma base de V. Com efeito, se $\mathcal{B}$ for LI, então por definição $\mathcal{B}$ já é uma base de V. Se $\mathcal{B}$ for LD, então existe pelo menos um vetor de $\mathcal{B}$ que é combinação linear dos demais. Sem perda de generalidade, suponha que v_n seja tal elemento (reordenando o conjunto se necessário), ou seja, existem escalares $k_1, \ldots, k_{n-1}$ tais que

$$v_n = k_1 v_1 + \ldots + k_{n-1} v_{n-1}.$$

Então

$$\begin{aligned} V &= [v_1, \ldots, v_n] = \{\alpha_1 v_1 + \ldots + \alpha_{n-1} v_{n-1} + \alpha_n v_n;\ \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} \\ &= \{\alpha_1 v_1 + \ldots + \alpha_{n-1} v_{n-1} + \alpha_n (k_1 v_1 + \ldots + k_{n-1} v_{n-1});\ \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} \\ &= \{(\alpha_1 + k_1 \alpha_n) v_1 + \ldots + (\alpha_{n-1} + k_{n-1} \alpha_n) v_{n-1};\ \alpha_1, \ldots, \alpha_n \in \mathbb{K}\} \\ &= \{\beta_1 v_1 + \ldots + \beta_{n-1} v_{n-1};\ \beta_1, \ldots, \beta_{n-1} \in \mathbb{K}\} \\ &= [v_1, \ldots, v_{n-1}]. \end{aligned}$$

Portanto, $V = [\mathcal{B}_1]$, onde $\mathcal{B}_1 = \{v_1, \ldots, v_{n-1}\}$. Se $\mathcal{B}_1$ for LI, então $\mathcal{B}_1$ é uma base de V, por definição. Se $\mathcal{B}_1$ for LD, existe um vetor de $\mathcal{B}_1$ que é combinação linear dos demais.

Analogamente ao que foi feito antes, podemos retirar tal vetor do conjunto obtendo um novo conjunto $\mathcal{B}_2$ que ainda gera V. Se $\mathcal{B}_2$ for LI, obtemos o resultado. Se não, repetimos este processo um número finito de vezes até encontrar um subconjunto de $\mathcal{B}$ formado por vetores LI que geram V (se reduzirmos o conjunto a um único vetor, esse conjunto é linearmente independente pela Proposição 2.16, já que $\mathcal{B}$ foi definido como um conjunto de vetores não nulos). O subconjunto resultante é uma base para V. ■

Proposição 2.22. *Se um $\mathbb{K}$-espaço vetorial V for gerado por vetores $v_1, \ldots, v_n$, então qualquer subconjunto de V com mais de n vetores é LD.*

Demonstração: Por hipótese, $V = [v_1, \ldots, v_n]$. Pela proposição anterior, podemos extrair de $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base $\mathcal{B}' = \{v_1, \ldots, v_r\}$ para V, com $r \leq n$ (reoordenando $\mathcal{B}$ se necessário). Sejam $w_1, \ldots, w_m \in V$ com $m > n$. Como $v_1, \ldots, v_r$ geram V temos

$$\begin{cases} w_1 = \alpha_{11}v_1 + \alpha_{12}v_2 + \ldots + \alpha_{1r}v_r \\ \quad\vdots \\ w_m = \alpha_{m1}v_1 + \alpha_{m2}v_2 + \ldots + \alpha_{mr}v_r. \end{cases}$$

Para mostrar que $\{w_1, \ldots, w_m\}$ é LD devemos encontrar escalares $x_1, \ldots, x_m \in \mathbb{K}$ não todos nulos tais que $x_1w_1 + \ldots, x_mw_m = \bar{0}$. Se $x_1, \ldots, x_m$ são escalares quaisquer, então

$$\begin{aligned} \bar{0} &= x_1w_1 + \ldots + x_mw_m \\ &= x_1(\alpha_{11}v_1 + \alpha_{12}v_2 + \ldots + \alpha_{1r}v_r) + \ldots + x_m(\alpha_{m1}v_1 + \alpha_{m2}v_2 + \ldots + \alpha_{mr}v_r) \\ &= (\alpha_{11}x_1 + \ldots + \alpha_{m1}x_m)v_1 + \ldots + (\alpha_{1r}x_1 + \ldots + \alpha_{mr}x_m)v_r. \end{aligned}$$

Como $\{v_1, \ldots, v_r\}$ é LI temos que

$$\begin{cases} \alpha_{11}x_1 + \alpha_{21}x_2 + \ldots + \alpha_{m1}x_m = 0 \\ \alpha_{12}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{m2}x_m = 0 \\ \qquad\qquad\vdots \\ \alpha_{1r}x_1 + \alpha_{2r}x_2 + \ldots + \alpha_{mr}x_m = 0 \end{cases},$$

que é um sistema linear homogêneo com r equações e m variáveis. Como $r < m$, então o posto da matriz dos coeficientes satisfaz $p_C \leq r < m$, implicando que o sistema é possível e indeterminado. Logo, o sistema possui solução $(x_1, \ldots, x_m) \neq (0, \ldots, 0)$. Em outras palavras, existem escalares $x_1, \ldots, x_m$ não todos nulos tais que $x_1w_1 + \ldots, x_mw_m = \bar{0}$. Portanto, $\{w_1, \ldots, w_m\}$ é LD. ■

Corolário 2.23. *Toda base de um espaço vetorial finitamente gerado possui o mesmo número de vetores.*

Demonstração: Sejam $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{w_1, \ldots, w_m\}$ bases de um espaço vetorial V. Como $\mathcal{B}$ gera V e $\mathcal{B}'$ é LI, segue da proposição anterior que $m \leq n$. Do mesmo modo, como $\mathcal{B}'$ gera V e $\mathcal{B}$ é LI, segue da proposição anterior que $n \leq m$. Portanto, $n = m$. ■

Portanto, podemos ter infinitas bases para V, porém cada uma delas terá o mesmo número de vetores. Esse fato nos leva a definir a dimensão de um espaço vetorial como sendo o número de vetores de uma base para V.

Definição 2.24. *Seja V um $\mathbb{K}$-espaço vetorial. Dizemos que V possui dimensão finita se V possuir uma base com um número finito de elementos. Neste caso, se $\mathcal{B} = \{v_1, \ldots, v_n\}$ é uma base de V, então a* ***dimensão de** V **sobre*** *$\mathbb{K}$, denotada por $\dim_{\mathbb{K}} V$, é igual a n. Se $V = \{\bar{0}\}$, então a dimensão de V sobre $\mathbb{K}$ é igual a 0.*

Exemplo 2.25. De acordo com o Exemplo 2.20 temos o seguinte:

1. A dimensão de $\mathbb{R}^n$ sobre $\mathbb{R}$ é igual a n. Em símbolos $\dim_{\mathbb{R}} \mathbb{R}^n = n$.

2. A dimensão de $\mathcal{P}_n(\mathbb{K})$ sobre $\mathbb{K}$ é igual a $n+1$. Em símbolos $\dim_{\mathbb{K}} \mathcal{P}_n(\mathbb{K}) = n+1$.

3. A dimensão de $\mathcal{M}_{n\times n}(\mathbb{K})$ sobre $\mathbb{K}$ é igual a n^2. Em símbolos $\dim_{\mathbb{K}} \mathcal{M}_{n\times n}(\mathbb{K}) = n^2$.

4. A dimensão de $\mathbb{C}^n$ sobre $\mathbb{C}$ é igual a n, e sobre $\mathbb{R}$ é igual a $2n$. Em símbolos

   $$\dim_{\mathbb{C}} \mathbb{C}^n = n \qquad \text{e} \qquad \dim_{\mathbb{R}} \mathbb{C}^n = 2n.$$

   Para verificar isso, basta mostrar que o conjunto

   $$\{(1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$$

   constitui uma base de $\mathbb{C}^n$ sobre $\mathbb{C}$ e o conjunto

   $$\{(1,0,\ldots,0),\ldots,(0,\ldots,0,1),(i,0,\ldots,0),\ldots,(0,\ldots,0,i)\}$$

   constitui uma base de $\mathbb{C}^n$ sobre $\mathbb{R}$.

Agora que sabemos da importância de uma base para um espaço vetorial, como podemos construir uma base para V selecionando vetores desse espaço? Vimos na Proposição 2.21 que a partir de um conjunto finito de geradores de V podemos sempre extrair uma base. Agora, dado um subconjunto S de vetores em V, como construir uma base para V acrescentando vetores em S?

Teorema 2.26. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita. Qualquer subconjunto de vetores LI de V pode ser completado de modo a formar uma base de V.*

Demonstração: Suponha que $\dim_{\mathbb{K}} V = n$ e seja $S = \{v_1, \ldots, v_r\}$ um conjunto LI contido em V. Pela Proposição 2.22, temos $r \leq n$. Se $V = [S]$, então S forma uma base de V e, neste caso, $r = n$. Se $V \neq [S]$, então existe um vetor $v_{r+1} \in V$ tal que $v_{r+1} \notin [S]$, ou seja, v_{r+1} não é combinação linear de $v_1, \ldots, v_r$. Por definição, $S' = \{v_1, \ldots, v_r, v_{r+1}\}$ é LI. Se $V = [S']$, então S' é a base procurada. Caso contrário, existe $v_{r+2} \in V$ tal que $v_{r+2} \notin [v_1, \ldots, v_r, v_{r+1}]$. Assim $S'' = \{v_1, \ldots, v_{r+1}, v_{r+2}\}$ é LI, e pode ou não ser uma base de V. Como $\dim_{\mathbb{K}} V = n$, não podemos ter um conjunto com mais do que n vetores linearmente independentes (Proposição 2.22) e, portanto, esse processo termina após um número finito de passos. Assim, é possível obter uma base $\{v_1, \ldots, v_r, \ldots, v_n\}$ para V. ■

Corolário 2.27. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão n. Então um subconjunto $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$ é LI se, e somente se, $\mathcal{B}$ gera V.*

Demonstração: Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ um subconjunto LI de V. Se $V \neq [\mathcal{B}]$, então pelo teorema anterior podemos completar $\mathcal{B}$ a fim de obter uma base de V. Mas isso é um absurdo, pois $\dim_{\mathbb{K}} V = n$. Portanto, $\mathcal{B}$ gera V.

Por outro lado, suponha que $V = [\mathcal{B}]$, onde $\mathcal{B} = \{v_1, \ldots, v_n\}$. Se $\mathcal{B}$ for LD, então pela Proposição 2.21 podemos extrair de $\mathcal{B}$ uma base de V, cujo número de vetores é menor do que n, o que é um absurdo pois $\dim_{\mathbb{K}} V = n$. Portanto, $\mathcal{B}$ é LI. ■

Portanto, conhecida a dimensão de um espaço vetorial V, o problema de determinar uma base para tal espaço se resume ao problema de encontrar um conjunto gerador para V ou um subconjunto LI de V cujo número de vetores coincide com a dimensão de V.

2.5 Operações com Subespaços Vetoriais

A seguir definimos a união e a interseção de subespaços vetoriais como uma adaptação da união e intersecção de conjuntos. Uma outra operação importante entre subespaços vetoriais é a soma, diferente da união pois, em geral, a união de dois subespaços não é um subespaço.

2.5.1 Soma e Interseção

Definição 2.28. *Sejam S_1 e S_2 subespaços de um $\mathbb{K}$-espaço vetorial V. A interseção de S_1 e S_2, denotada por $S_1 \cap S_2$, é o conjunto de todos os vetores $v \in V$ tais que $v \in S_1$ e $v \in S_2$. A união de S_1 e S_2, denotada por $S_1 \cup S_2$, é o conjunto de todos os vetores $v \in V$ tais que $v \in S_1$ ou $v \in S_2$. Em símbolos,*

$$S_1 \cap S_2 = \{v \in V;\ v \in S_1 \ \textit{e}\ v \in S_2\} \quad \textit{e} \quad S_1 \cup S_2 = \{v \in V;\ v \in S_1 \ \textit{ou}\ v \in S_2\}.$$

Proposição 2.29. *Seja V um $\mathbb{K}$-espaço vetorial. A interseção de uma coleção arbitrária de subespaços vetoriais de V é um subespaço vetorial de V.*

Demonstração: Seja $\{S_i\}_{i\in J}$ uma coleção arbitrária de subespaços vetoriais de V. Provemos que $S = \bigcap_{i\in J} S_i$ é um subespaço vetorial de V. De fato, note que $\bar{0} \in S$ pois $\bar{0} \in S_i$ para todo $i \in J$, onde $\bar{0}$ é o vetor nulo de V. Logo, $S \neq \emptyset$. Agora, sejam $u, v \in S$ e $k \in \mathbb{K}$. Como $u, v \in S_i$ para todo $i \in J$ e como todos os S_i são subespaços de V segue que $ku+v \in S_i$ para todo $i \in J$. Logo, $ku+v \in S$. Pela Proposição 2.5, S é um subespaço de V. ■

Em geral, a união de subespaços vetoriais não é um subespaço vetorial. Considere, por exemplo, os subespaços de $\mathbb{R}^2$:

$$S_1 = \{(x, y) \in \mathbb{R}^2; y = 0\} \quad \text{e} \quad S_2 = \{(x, y) \in \mathbb{R}^2; x = 0\}.$$

Tome $(1, 0) \in S_1$ e $(0, 1) \in S_2$. Observe que $(1, 0) + (0, 1) = (1, 1)$, que não pertence a S_1 e não pertence a S_2. Logo, $(1, 0) + (0, 1) \notin S_1 \cup S_2$, implicando que $S_1 \cup S_2$ não é um subespaço vetorial de $\mathbb{R}^2$. Ao contrário do que ocorre com a união de dois subepaços, a soma de dois subespaços sempre será um subespaço vetorial.

Definição 2.30. *Sejam S_1 e S_2 subespaços de um $\mathbb{K}$-espaço vetorial V. Definimos a soma de S_1 e S_2, denotada por $S_1 + S_2$, como o conjunto*

$$S_1 + S_2 = \{v \in V;\ v = v_1 + v_2,\ \textit{com}\ v_1 \in S_1, v_2 \in S_2\}.$$

Proposição 2.31. *Se S_1 e S_2 são subespaços de um $\mathbb{K}$-espaço vetorial V, então $S_1 + S_2$ também é um subespaço vetorial de V*

Demonstração: Sejam S_1 e S_2 subespaços vetoriais de V. Primeiramente, $S_1+S_2 \neq \emptyset$ pois $\bar{0} = \bar{0} + \bar{0} \in S_1 + S_2$. Sejam $u, v \in S_1 + S_2$ e $k \in \mathbb{K}$. Por definição, $u = u_1 + u_2$ e $v = v_1 + v_2$, com $u_1, v_1 \in S_1$ e $u_2, v_2 \in S_2$. Logo,

$$ku + v = k(u_1 + u_2) + (v_1 + v_2) = (ku_1 + v_1) + (ku_2 + v_2).$$

Como S_1 e S_2 são subespaços de V, temos que $ku_1 + v_1 \in S_1$ e $ku_2 + v_2 \in S_2$. Logo, $ku + v = w_1 + w_2 \in S_1 + S_2$, implicando pela Proposição 2.5 que $S_1 + S_2$ é um subespaço vetorial de V. ■

Observe que se V é um espaço vetorial de dimensão finita, então qualquer subespaço S de V também tem dimensão finita e $\dim_{\mathbb{K}} S \leq \dim_{\mathbb{K}} V$. De fato, se $S = \{\bar{0}\}$ então não há nada a mostrar, pois $\dim_{\mathbb{K}} V \geq 0 = \dim_{\mathbb{K}} S$. Do mesmo modo, se $S = V$, então claramente $\dim_{\mathbb{K}} S = \dim_{\mathbb{K}} V$. Suponhamos então que $\{\bar{0}\} \neq S \neq V$ e considere $\mathcal{B} = \{w_1, \ldots, w_n\}$

uma base de S. Como $S \neq V$, existe $v \in V$ tal que $v \notin S$, logo $v \notin [w_1, \dots, w_n]$. Como $\mathcal{B}$ é LI, segue que $\mathcal{C} = \{w_1, \dots, w_n, v\}$ também é LI. Pela Proposição 2.16, podemos completar $\mathcal{C}$ a fim de obter uma base para V. Logo, $\dim_{\mathbb{K}} V \geq n+1 > \dim_{\mathbb{K}} S$, como queríamos. Temos também o seguinte:

TEOREMA 2.32. *Se S_1, S_2 são subespaços de um $\mathbb{K}$-espaço vetorial V de dimensão finita, então*

$$\dim_{\mathbb{K}}(S_1 + S_2) = \dim_{\mathbb{K}} S_1 + \dim_{\mathbb{K}} S_2 - \dim_{\mathbb{K}}(S_1 \cap S_2).$$

DEMONSTRAÇÃO: Veja [3, Proposição 2.4.5], pp. 60-62. ■

EXEMPLO 2.33. 1. Considere $V = \mathbb{R}^3$ e seus subespaços

$$S_1 = \{(x, y, z) \in \mathbb{R}^3;\ y = z = 0\} \quad \text{e} \quad S_2 = \{(x, y, z) \in \mathbb{R}^3;\ y = x + z\}.$$

Observe que $\mathcal{B}_1 = \{(1, 0, 0)\}$ e $\mathcal{B}_2 = \{(1, 1, 0), (0, 1, 1)\}$ são bases de S_1 e S_2, respectivamente. De fato, é fácil ver que $S_1 = [(1, 0, 0)]$ e $S_2 = [(1, 1, 0), (0, 1, 1)]$, pois

$$(x, 0, 0) = x(1, 0, 0) \quad \text{e} \quad (x, x+z, z) = x(1, 1, 0) + z(0, 1, 1), \quad x, z \in \mathbb{R}.$$

Além disso, $\mathcal{B}_1$ e $\mathcal{B}_2$ são LI. Logo, $\dim_{\mathbb{R}} S_1 = 1$ e $\dim_{\mathbb{R}} S_2 = 2$. Como $S_1 \cap S_2 = \{(0, 0, 0)\}$ segue que $\dim_{\mathbb{R}} S_1 \cap S_2 = 0$. Pelo teorema anterior, $\dim_{\mathbb{R}} S_1 + S_2 = 3$. Como $S_1 + S_2$ é um subespaço vetorial de $\mathbb{R}^3$, então $S_1 + S_2 = \mathbb{R}^3$.

Vamos explorar mais um pouco nossa última afirmação. Claramente $S_1 + S_2 \subset \mathbb{R}^3$, por definição. Agora, dado $v = (x, y, z) \in \mathbb{R}^3$ temos

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

Logo, $v = v_1 + v_2 \in S_1 + S_2$, pois $v_1 = (x - y + z)(1, 0, 0) \in [(1, 0, 0)] = S_1$ e $v_2 = (y - z)(1, 1, 0) + z(0, 1, 1) \in [(1, 1, 0), (0, 1, 1)] = S_2$. Como v é arbitrário, $\mathbb{R}^3 \subset S_1 + S_2$, donde segue a igualdade. Portanto, mesmo que não saibamos a priori a dimensão de $S_1 + S_2$, é possível mostrar que $S_1 + S_2 = \mathbb{R}^3$. Como neste caso $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ gera $S_1 + S_2$, cuja dimensão é 3, segue pelo Corolário 2.27 que $\mathcal{B}$ é uma base de $S_1 + S_2 = \mathbb{R}^3$.

2. Sejam

$$W_1 = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}; a, b \in \mathbb{R} \right\} \quad \text{e} \quad W_2 = \left\{ \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}; c, d \in \mathbb{R} \right\}$$

subespaços de $\mathcal{M}_{2\times 2}(\mathbb{R})$. Observe que $W_1 \cap W_2 = \{\bar{0}\}$, onde $\bar{0}$ representa a matriz nula de ordem 2. Além disso,

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \right\} \quad \text{e} \quad \mathcal{B}_2 = \left\{ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

são bases de W_1 e W_2, respectivamente. Então, pelo teorema anterior, temos que $\dim_{\mathbb{R}} W_1 + W_2 = 2 + 2 - 0 = 4$. Portanto, $W_1 + W_2 = \mathcal{M}_{2\times 2}(\mathbb{R})$.

3. Sejam $S_1 = \{(x, y, z, t) \in \mathbb{R}^4; x - y - t = 0\}$ e $S_2 = \{(0, 0, z, 0);\ z \in \mathbb{R}\}$ subespaços de $\mathbb{R}^4$. Facilmente vemos que $\mathcal{B}_2 = \{(0, 0, 1, 0)\}$ é uma base de S_2. Como

$$(y + t, y, z, t) = y(1, 1, 0, 0) + z(0, 0, 1, 0) + t(1, 0, 0, 1),$$

então $\mathcal{B}_1 = \{(1, 1, 0, 0), (0, 0, 1, 0), (1, 0, 0, 1)\}$ é uma base de S_1 (uma vez que também é LI). Logo, $\dim_{\mathbb{R}} S_1 = 3$ e $\dim_{\mathbb{R}} S_2 = 1$. Como $S_2 \subset S_1$, temos $S_1 \cap S_2 = S_2$, donde segue que $\dim_{\mathbb{R}} S_1 \cap S_2 = \dim_{\mathbb{R}} S_2 = 1$. Pelo teorema anterior, $\dim_{\mathbb{R}} S_1 + S_2 = 3 + 1 - 1 = 3$. Logo, $S_1 + S_2 \neq \mathbb{R}^4$. Para conhecer a estrutura de $S_1 + S_2$, precisamos determinar sua base. Observe que $v = v_1 + v_2 \in S_1 + S_2$ se, e somente se,

$$v = k_1(1, 1, 0, 0) + k_2(0, 0, 1, 0) + k_3(1, 0, 0, 1) + k_4(0, 0, 1, 0),$$

para $k_1, k_2, k_3, k_4 \in \mathbb{R}$. Portanto, se denotarmos $v = (x, y, z, t)$ temos $x = k_1 + k_3$, $y = k_1$, $z = k_2 + k_4$ e $t = k_3$, ou seja, $x - y - t = 0$. Logo, $S_1 + S_2 \subset S_1$. Como ambos os subespaços têm a mesma dimensão, $S_1 + S_2 = S_1$ e, portanto, $\mathcal{B}_1$ é uma base de $S_1 + S_2$ (observe que $\mathcal{B}_1 \cup \mathcal{B}_2 = \mathcal{B}_1$).

Observe que em todos os exemplos acima, uma base para a soma de dois subespaços coincide com a união de suas bases. Essa é exatamente a ideia para determinar uma base para a soma de subespaços, desde que essa união seja sempre LI.

2.5.2 Somas diretas

Dado um $\mathbb{K}$-espaço vetorial V, muitas vezes é conveniente escrever seus elementos como uma soma de elementos de dois ou mais subespaços. Nem sempre essa soma é única, mas existem casos em que garantimos a unicidade. A fim de entender melhor esse conceito, definimos:

Definição 2.34. *Sejam V um $\mathbb{K}$-espaço vetorial e W_1, W_2 dois subespaços vetoriais de V. Diremos que a soma $W_1 + W_2$ é direta se $W_1 \cap W_2 = \{\bar{0}\}$. Neste caso, escrevemos $W_1 \oplus W_2$.*

Proposição 2.35. *Sejam V um $\mathbb{K}$-espaço vetorial e W_1, W_2 dois subespaços de V. Então $V = W_1 \oplus W_2$ se, e somente se, cada $v \in V$ se escreve de maneira única como $v = w_1 + w_2$, com $w_i \in W_i, i = 1, 2$.*

Demonstração: Suponhamos que $V = W_1 \oplus W_2$. Por definição, todo $v \in V$ pertence ao subespaço $W_1 + W_2$, onde $W_1 \cap W_2 = \{\bar{0}\}$, isto é, existem $w_1 \in W_1$ e $w_2 \in W_2$ tais que $v = w_1 + w_2$. Suponha que $v = w_1 + w_2 = u_1 + u_2$, para $u_i \in W_i$, $i = 1, 2$. Isso implica que $w_1 - u_1 = u_2 - w_2$. Como W_1 e W_2 são subespaços de V, temos $w_1 - u_1 \in W_1$ e $u_2 - w_2 \in W_2$. Então $w_1 - u_1 = u_2 - w_2 \in W_1 \cap W_2 = \{\bar{0}\}$, implicando que $w_1 = u_1$ e $w_2 = u_2$. Reciprocamente, suponha que cada $v \in V$ se escreva de maneira única como $v = w_1 + w_2$, com $w_i \in W_i$, $i = 1, 2$. Então $v \in W_1 + W_2$ e, portanto, $V = W_1 + W_2$. Resta mostrar que $W_1 \cap W_2 = \{\bar{0}\}$. Para isso, tome $w \in W_1 \cap W_2$. Como $w \in V$, ele pode ser escrito como $w = w + \bar{0} \in W_1 + W_2$ se considerarmos $w \in W_1$ e como $w = \bar{0} + w \in W_1 + W_2$ se considerarmos $w \in W_2$. Pela hipótese de unicidade, $w = \bar{0}$, o que nos garante que $W_1 \cap W_2 = \{\bar{0}\}$. ■

Exemplo 2.36. Seja $V = \mathcal{M}_{n\times n}(\mathbb{K})$ o espaço vetorial das matrizes quadradas de ordem n com entradas em $\mathbb{K}$. Considere

$$W_1 = \{A \in V; A^t = A\} \quad \text{e} \quad W_2 = \{A \in V; A^t = -A\}$$

os subespaços das matrizes simétricas e antissimétricas, respectivamente. Vamos mostrar que $V = W_1 \oplus W_2$. De fato, dada $A \in V$, podemos escrever

$$A = \frac{A + A^t}{2} + \frac{A - A^t}{2}.$$

Agora $A_1 = \dfrac{A + A^t}{2} \in W_1$ e $A_2 = \dfrac{A - A^t}{2} \in W_2$, pois

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

Portanto, $A = A_1 + A_2 \in W_1 + W_2$. Como a única matriz que é simétrica e antissimétrica simultaneamente é a matriz nula, segue que $W_1 \cap W_2 = \{\bar{0}\}$ e , portanto, $V = W_1 \oplus W_2$.

Esse exemplo nos mostra que toda matriz quadrada se escreve de maneira única como a soma de uma matriz simétrica com uma matriz antissimétrica.

2.6 Matriz Mudança de Base

Muitas vezes pode ser conveniente trocarmos a base do espaço em que estamos trabalhando de modo a facilitar nossos cálculos. Essa mudança de base sempre pode ser feita por meio de uma matriz, chamada **matriz mudança de base**, e para introduzí-la começaremos com o conceito de cooordenadas de um vetor com relação a uma base.

Definição 2.37. *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita e $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base de V. Fixemos a ordem dos elementos de $\mathcal{B}$ e a chamemos de **base ordenada** de V. Dado $v \in V$, existem escalares $k_1, \ldots, k_n \in \mathbb{K}$ únicos tais que*

$$v = k_1 v_1 + k_2 v_2 + \ldots + k_n v_n.$$

Pela unicidade, tais escalares são chamados ***coordenadas de*** v ***em relação à base*** $\mathcal{B}$ *e denotamos*

$$[v]_{\mathcal{B}} = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.$$

Sejam agora $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{B}' = \{u_1, \ldots, u_n\}$ bases ordenadas de um mesmo espaço vetorial V. Dado um vetor $v \in V$, podemos escrevê-lo nas bases $\mathcal{B}$ e $\mathcal{B}'$ respectivamente como

$$v = k_1 v_1 + k_2 v_2 + \ldots + k_n v_n \tag{2.1}$$

e

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n, \tag{2.2}$$

com $k_1, \ldots, k_n, \lambda_1, \ldots, \lambda_n \in \mathbb{K}$. Vejamos como podemos relacionar as coordenadas de v em relação a essas duas bases. Como $\mathcal{B}'$ também é uma base de V, podemos escrever os vetores v_i como combinação linear dos vetores da base $\mathcal{B}'$:

$$\begin{cases} v_1 = a_{11}u_1 + a_{21}u_2 + \ldots + a_{n1}u_n \\ v_2 = a_{12}u_1 + a_{22}u_2 + \ldots + a_{n2}u_n \\ \qquad\vdots \\ v_n = a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{nn}u_n. \end{cases} \tag{2.3}$$

Substituindo em (2.1) temos

$$\begin{aligned} v &= k_1 v_1 + k_2 v_2 \ldots + k_n v_n \\ &= k_1(a_{11}u_1 + a_{21}u_2 + \ldots + a_{n1}u_n) + k_2(a_{12}u_1 + a_{22}u_2 + \ldots + a_{n2}u_n) + \ldots + \\ &\quad k_n(a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{nn}u_n) \\ &= (a_{11}k_1 + a_{12}k_2 + \ldots + a_{1n}k_n)u_1 + (a_{21}k_1 + a_{22}k_2 + \ldots + a_{2n}k_n)u_2 + \ldots + \\ &\quad (a_{n1}k_1 + a_{n2}k_2 + \ldots + a_{nn}k_n)u_n. \end{aligned}$$

Como as coordenadas em relação a uma base fixada são únicas, por (2.2) temos

$$\lambda_i = a_{i1}k_1 + a_{i2}k_2 + \ldots + a_{in}k_n,$$

para todo $1 \leq i \leq n$. Isso significa que as coordenadas de v na base $\mathcal{B}'$ são dadas por

$$[v]_{\mathcal{B}'} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} a_{11}k_1 + a_{12}k_2 + \ldots + a_{1n}k_n \\ a_{21}k_1 + a_{22}k_2 + \ldots + a_{2n}k_n \\ \vdots \\ a_{n1}k_1 + a_{n2}k_2 + \ldots + a_{nn}k_n \end{pmatrix},$$

que na forma matricial torna-se

$$[v]_{\mathcal{B}'} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A[v]_{\mathcal{B}}$$

para $A = (a_{ij})_{n\times n}$. De acordo com (2.3), as colunas da matriz A são formadas pelas coordenadas de cada $v_i \in \mathcal{B}$ em relação à base $\mathcal{B}'$. Por esse motivo, a matriz A é chamada **matriz de mudança da base $\mathcal{B}$ para a base $\mathcal{B}'$** e é denotada por $[I]^{\mathcal{B}}_{\mathcal{B}'}$. Com isso provamos o seguinte teorema:

Teorema 2.38. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita. Sejam $\mathcal{B}$ e $\mathcal{B}'$ duas bases ordenadas de V. Para todo $v \in V$, temos*

$$[v]_{\mathcal{B}'} = [I]^{\mathcal{B}}_{\mathcal{B}'}[v]_{\mathcal{B}},$$

onde a matriz de mudança de base $[I]^{\mathcal{B}}_{\mathcal{B}'}$ é inversível e $([I]^{\mathcal{B}}_{\mathcal{B}'})^{-1} = [I]^{\mathcal{B}'}_{\mathcal{B}}$.

Demonstração: Resta provar que $[I]^{\mathcal{B}}_{\mathcal{B}'}$ é inversível e que sua inversa é a matriz de mudança da base $\mathcal{B}'$ para a base $\mathcal{B}$. De fato, suponhamos que $\dim_{\mathbb{K}} V = n$. Como $\mathcal{B}$ e $\mathcal{B}'$ são bases arbitrárias, temos

$$[v]_{\mathcal{B}'} = [I]^{\mathcal{B}}_{\mathcal{B}'}[v]_{\mathcal{B}} \quad \text{e} \quad [v]_{\mathcal{B}} = [I]^{\mathcal{B}'}_{\mathcal{B}}[v]_{\mathcal{B}'},$$

ou seja,

$$[v]_{\mathcal{B}'} = [I]^{\mathcal{B}}_{\mathcal{B}'}[v]_{\mathcal{B}} = [I]^{\mathcal{B}}_{\mathcal{B}'}[I]^{\mathcal{B}'}_{\mathcal{B}}[v]_{\mathcal{B}'}.$$

Como essa igualdade ocorre para todo $v \in V$, segue que $[I]^{\mathcal{B}}_{\mathcal{B}'}[I]^{\mathcal{B}'}_{\mathcal{B}} = I_n$ onde I_n é a matriz identidade de ordem n. Por definição, $[I]^{\mathcal{B}}_{\mathcal{B}'}$ é inversível com inversa

$$([I]^{\mathcal{B}}_{\mathcal{B}'})^{-1} = [I]^{\mathcal{B}'}_{\mathcal{B}}.$$

∎

Exemplo 2.39. Sejam $\mathcal{B} = \{(1,-2),(3,-4)\}$ e $\mathcal{B}' = \{(1,3),(2,1)\}$ bases de $\mathbb{R}^2$. Determinemos $[I]^{\mathcal{B}}_{\mathcal{B}'}$. Para isso, escreva

$$(1,-2) = a(1,3) + b(2,1) \quad \text{e} \quad (3,-4) = c(1,3) + d(2,1).$$

Cálculos diretos nos mostram que $a = -1, b = 1, c = -\dfrac{11}{5}$ e $d = \dfrac{13}{5}$. Portanto,

$$[I]^{\mathcal{B}}_{\mathcal{B}'} = \begin{pmatrix} -1 & -\frac{11}{5} \\ 1 & \frac{13}{5} \end{pmatrix}$$

é a matriz mudança de base da base $\mathcal{B}$ para a base $\mathcal{B}'$. Como todo $v = (x, y) \in \mathbb{R}^2$ pode ser escrito como

$$(x, y) = \frac{-4x - 3y}{2}(1, -2) + \frac{2x + y}{2}(3, -4)$$

temos que as coordenadas de v em relação à base $\mathcal{B}$ é dada por

$$[v]_{\mathcal{B}} = \begin{pmatrix} \dfrac{-4x - 3y}{2} \\ \dfrac{2x + y}{2} \end{pmatrix}.$$

Pelo Teorema 2.38 segue que as coordenadas de v em relação à base $\mathcal{B}'$ é dada por

$$[v]_{\mathcal{B}'} = [I]^{\mathcal{B}}_{\mathcal{B}'}[v]_{\mathcal{B}} = \begin{pmatrix} -1 & -\dfrac{11}{5} \\ 1 & \dfrac{13}{5} \end{pmatrix} \begin{pmatrix} \dfrac{-4x - 3y}{2} \\ \dfrac{2x + y}{2} \end{pmatrix} = \begin{pmatrix} \dfrac{-x + 2y}{5} \\ \dfrac{3x - y}{5} \end{pmatrix},$$

o que quer dizer que todo $v = (x, y) \in \mathbb{R}^2$ pode ser escrito como

$$(x, y) = \frac{-x + 2y}{5}(1, 3) + \frac{3x - y}{5}(2, 1).$$

Capítulo 3

Transformações Lineares

Neste capítulo vamos explorar o conceito de transformação linear, que é um tipo particular de função entre dois espaços vetoriais que preserva as operações de adição vetorial e multiplicação por escalar. Uma transformação linear também pode ser chamada de operador linear no caso em que o domínio e contradomínio coincidem. O conceito de autovalores e autovetores de uma transformação linear também é de interesse em diversas áreas da matemática.

Em todo o texto, V e W são $\mathbb{K}$-espaços vetoriais de dimensão finita.

3.1 Transformações e Operadores lineares

DEFINIÇÃO 3.1. *Uma aplicação $T : V \to W$ é dita uma **transformação linear** se para todo $u, v \in V$ e $k \in \mathbb{K}$ tivermos $T(ku + v) = kT(u) + T(v)$. Essa condição pode ser reescrita separadamente como:*

(i) $T(u + v) = T(u) + T(v)$;

(ii) $T(ku) = kT(u)$.

*Se $V = W$ chamamos T de **operador linear**.*

EXEMPLO 3.2.

1. A aplicação $T : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x, y) = (2x - y,\ x,\ 4y - x)$ é uma transformação linear.

2. A aplicação $T : V \to W$ definida por $T(v) = \bar{0}_W$, para todo $v \in V$ e para $\bar{0}_W$ sendo o vetor nulo de W, é uma transformação linear chamada **transformação nula**.

3. A aplicação $I : V \to V$ dada por $I(v) = v$ é um operador linear chamado **operador identidade**.

4. A aplicação $T : \mathcal{P}_2(\mathbb{K}) \to \mathcal{M}_{2\times 2}(\mathbb{K})$ dada por $T(a_0 + a_1x + a_2x^2) = \begin{pmatrix} a_0 & a_1 \\ a_1 & a_2 \end{pmatrix}$ é uma transformação linear.

5. A aplicação $L : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $L(x, y) = (x, y + 2, 0)$ não é uma transformação linear, assim como $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (y, x^2, y)$ não é uma transformação linear.

Daqui em diante, denotaremos por $\bar{0}_V$ e $\bar{0}_W$ os vetores nulos de V e W, respectivamente. Observe que se $T : V \to W$ é uma transformação linear então $T(\bar{0}_V) = \bar{0}_W$. De fato,

$$T(\bar{0}_V) = T(\bar{0}_V + \bar{0}_V) = T(\bar{0}_V) + T(\bar{0}_V),$$

donde concluímos que $T(\bar{0}_V) = \bar{0}_W$, já que o vetor nulo de W é único.

Portanto, se $T(\bar{0}_V) \neq \bar{0}_W$ então T não pode ser uma transformação linear. Entretanto a recíproca da afirmação acima não é verdadeira. De fato, considere $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x^2, y)$. Veja que $T(0, 0) = (0, 0)$, entretanto T não é linear. De fato, para todo $k \neq 1$ temos

$$T(k(x, y)) = T(kx, ky) = ((kx)^2, ky) = (k^2x^2, ky) = k(kx^2, y) \neq k(x^2, y) = kT(x, y).$$

Uma transformação linear $T : V \to W$ é completamente determinada se soubermos o que ela faz nos vetores de uma base de V. De fato, dada $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base de V, todo $v \in V$ é escrito como $v = k_1v_1 + \ldots + k_nv_n$ com $k_i \in \mathbb{K}$. Neste caso, como T é linear, temos

$$T(v) = T(k_1v_1 + \ldots + k_nv_n) = k_1T(v_1) + \ldots + k_nT(v_n).$$

Ou seja, se soubermos a imagem dos vetores de uma base de V por uma aplicação T, é possível determinar $T(v)$ para todo $v \in V$. Vamos ver como isso funciona no próximo exemplo.

Exemplo 3.3. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear tal que $T(2, 3) = (-1, 5)$ e $T(0, 1) = (2, 1)$. Como $\mathcal{B} = \{(2, 3), (0, 1)\}$ é uma base de $\mathbb{R}^2$ temos que todo vetor $v = (x, y) \in \mathbb{R}^2$ é escrito como combinação linear de $(2, 3)$ e $(0, 1)$. Em particular,*

$$(x, y) = \frac{x}{2}(2, 3) + \frac{2y - 3x}{2}(0, 1).$$

Logo,

$$T(x, y) = T\left(\frac{x}{2}(2, 3) + \frac{2y - 3x}{2}(0, 1)\right) = \frac{x}{2}T(2, 3) + \frac{2y - 3x}{2}T(0, 1) = \frac{x}{2}(-1, 5) + \frac{2y - 3x}{2}(2, 1)$$

e, portanto,

$$T(x, y) = \left(\frac{-7x + 4y}{2}, x + y\right).$$

Algumas transformações lineares $T : \mathbb{R}^2 \to \mathbb{R}^2$ recebem nomes especiais:

1. **Reflexão em torno do eixo** Ox: $T(x, y) = (x, -y)$.
2. **Reflexão em torno do eixo** Oy: $T(x, y) = (-x, y)$.
3. **Reflexão em torno da origem**: $T(x, y) = (-x, -y)$.
4. **Dilatação**: $T(x, y) = k(x, y), \quad k \in \mathbb{R}, \ |k| > 1$.
5. **Contração**: $T(x, y) = k(x, y), \quad k \in \mathbb{R}, |k| < 1$.
6. **Rotação**: $T(x, y) = (x \cos\theta - y \sin\theta, x \sin\theta + y \cos\theta)$, com $0 \leq \theta \leq 2\pi$.

OBSERVAÇÃO 3.4. *Observe que a translação $T(x, y) = (x + a, y + b)$ não é transformação linear se $ab \neq 0$, pois neste caso $T(0, 0) \neq (0, 0)$.*

TEOREMA 3.5. *Sejam V um $\mathbb{K}$-espaço vetorial de dimensão finita com uma base $\mathcal{B} = \{v_1, \ldots, v_n\}$ ordenada. Sejam W um $\mathbb{K}$-espaço vetorial e $w_1, \ldots, w_n$ vetores arbitrários de W. Então existe uma única transformação linear $T : V \to W$ tal que $T(v_j) = w_j$ para $j = 1, \ldots, n$.*

DEMONSTRAÇÃO: Por hipótese, dado $v \in V$ existem únicos escalares $k_1, \ldots, k_n$ tais que $v = k_1v_1 + \ldots + k_nv_n$. Definimos $T(v) = k_1w_1 + \ldots + k_nw_n$. Primeiramente, T está bem definida, pois se $v = w \in V$ temos $T(v) = T(w) \in W$. Por definição, $T(v_j) = w_j$ para todo $j = 1, \ldots, n$. Note também que T é linear. Para ver isso, escreva $u \in V$ como $u = \lambda_1v_1 + \ldots + \lambda_nv_n$. Para todo $k \in \mathbb{K}$ temos

$$kv + u = k(k_1v_1 + \ldots + k_nv_n) + \lambda_1v_1 + \ldots + \lambda_nv_n = (kk_1 + \lambda_1)v_1 + \ldots + (kk_n + \lambda_n)v_n.$$

Portanto, pela definição de T temos que

$$T(kv + u) = (kk_1 + \lambda_1)w_1 + \ldots + (kk_n + \lambda_n)w_n.$$

Por outro lado,

$$kT(v)+T(u) = k(k_1w_1+\ldots+k_nw_n)+\lambda_1w_1+\ldots+\lambda_nw_n = (kk_1+\lambda_1)w_1+\ldots+(kk_n+\lambda_n)w_n.$$

Logo, $T(kv + u) = kT(v) + T(u)$, para todo $u, v \in V$ e $k \in \mathbb{K}$, implicando que T é linear.

Resta provar que T é única. Seja $R : V \to W$ uma transformação linear tal que $R(v_j) = w_j$ para todo $j = 1, \ldots, n$. Então

$$R(v) = R(\sum_{i=1}^{n} k_iv_i) = \sum_{i=1}^{n} k_iR(v_i) = \sum_{i=1}^{n} k_iw_i = T(v),$$

para todo $v \in V$. Logo $R = T$. ■

O conjunto de todas as transformações lineares de V em W herda uma estrutura natural de espaço vetorial sobre $\mathbb{K}$. Para entender isso, considere $T : V \to W$, $R : V \to W$ transformações lineares entre $\mathbb{K}$-espaços vetoriais e $k \in \mathbb{K}$. Definimos as operações de soma e produto por escalar

$$\begin{array}{rcl} T+R : V & \to & W \\ v & \mapsto & T(v)+R(v) \end{array} \quad \text{e} \quad \begin{array}{rcl} kT : V & \to & W \\ v & \mapsto & kT(v). \end{array}$$

As aplicações $T+R$ e kT são transformações lineares (exercício), e com essas operaçõs o conjunto de todas as transformações lineares de V em W é um $\mathbb{K}$-espaço vetorial, denotado por $\mathcal{L}(V, W)$.

3.2 Núcleo, Imagem e Isomorfismos Lineares

Definição 3.6. *Definimos o* ***núcleo*** *de uma transformação linear $T : V \to W$, denotado por $Nuc(T)$, como o subconjunto de V dado por*

$$Nuc(T) = \{v \in V;\ T(v) = \bar{0}_W\}.$$

A ***imagem*** *de $T : V \to W$ é o subconjunto de W, denotado $Im(T)$, definido por*

$$Im(T) = \{w \in W;\ w = T(v) \text{ para algum } v \in V\}.$$

Proposição 3.7. *Para toda transformação linear $T : V \to W$, $Nuc(T) \subseteq V$ é um subespaço vetorial de V e $Im(T) \subseteq W$ é um subespaço vetorial de W.*

Demonstração: Observe que $Nuc(T)$ e $Im(T)$ são não vazios, pois $\bar{0}_V \in Nuc(T)$ e $\bar{0}_W \in Im(T)$, pois como T é uma transformação linear segue que $T(\bar{0}_V) = \bar{0}_W$.

Sejam $u, v \in Nuc(T)$ e $k \in \mathbb{K}$ quaisquer. Assim $T(u) = \bar{0}_W = T(v)$, o que implica em

$$T(ku+v) = kT(u) + T(v) = k\bar{0}_W + \bar{0}_W = \bar{0}_W.$$

Logo, $ku+v \in Nuc(T)$. Agora, se $z, w \in Im(T)$ e $k \in \mathbb{K}$ então existem $u, v \in V$ tais que $z = T(u)$ e $w = T(v)$, donde

$$T(ku+v) = kT(u) + T(v) = kz + w.$$

Portanto, $kz + w \in Im(T)$ e o resultado segue. ∎

Teorema 3.8. ***(Teorema do Núcleo e da Imagem)*** *Seja $T : V \to W$ uma transformação linear. Então*

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} Nuc(T) + \dim_{\mathbb{K}} Im(T).$$

Demonstração: Seja $\mathcal{B} = \{v_1, \ldots, v_k\}$ uma base para $Nuc(T)$. Como $\mathcal{B}$ é LI, podemos completar tal conjunto de modo a obter uma base $\mathcal{C} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ de V. Provemos que $\mathcal{D} = \{T(v_{k+1}), \ldots, T(v_n)\}$ é uma base de $Im(T)$, ou seja, que $\mathcal{D}$ gera $Im(T)$ e é LI. De fato, se $w \in Im(T)$ então existe $u \in V$ tal que $w = T(u)$. Como $u \in V$, ele se escreve de maneira única como $u = \lambda_1 v_1 + \ldots + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \ldots + \lambda_n v_n$. Deste modo, como $v_1, \ldots, v_k \in Nuc(T)$ temos

$$\begin{aligned} w &= T(u) = T(\lambda_1 v_1 + \ldots + \lambda_k v_k + \lambda_{k+1} v_{k+1} + \ldots + \lambda_n v_n) \\ &= \lambda_1 T(v_1) + \ldots + \lambda_k T(v_k) + \lambda_{k+1} T(v_{k+1}) + \ldots + \lambda_n T(v_n) \\ &= \lambda_{k+1} T(v_{k+1}) + \ldots + \lambda_n T(v_n). \end{aligned}$$

Logo, $Im(T) = [T(v_{k+1}), \ldots, T(v_n)]$. Consideremos agora, uma combinação linear nula qualquer

$$a_{k+1} T(v_{k+1}) + \cdots + a_n T(v_n) = \bar{0}_W,$$

com $a_i \in \mathbb{K}$, para $i = k+1, \ldots, n$. Como T é linear segue que $T(a_{k+1} v_{k+1} + \cdots + a_n v_n) = \bar{0}_W$, ou seja, $a_{k+1} v_{k+1} + \cdots + a_n v_n \in Nuc(T)$. Assim podemos escrever

$$a_{k+1} v_{k+1} + \cdots + a_n v_n = a_1 v_1 + \ldots + a_k v_k$$

para alguns $a_1, \ldots, a_k \in \mathbb{K}$. Logo, $-a_1 v_1 - \cdots - a_k v_k + a_{k+1} v_{k+1} + \cdots + a_n v_n = \bar{0}_V$. Como $\mathcal{C}$ é base de V, e portanto é LI, concluímos que $a_i = 0$ para todo $i = 1, \ldots, n$. Assim, $a_{k+1} = \cdots = a_n = 0$, implicando que $\mathcal{D}$ é LI. Portanto, $\dim_{\mathbb{K}} Nuc(T) = k$ e $\dim_{\mathbb{K}} Im(T) = n - k$, ou seja, $\dim_{\mathbb{K}} Nuc(T) + \dim_{\mathbb{K}} Im(T) = k + n - k = n = \dim_{\mathbb{K}} V$. ■

Definição 3.9. *Seja $T : V \to W$ uma transformação linear. Definimos o* ***posto*** *de T como a dimensão da imagem de T e a* ***nulidade*** *de T como a dimensão do núcleo de T. Em símbolos, posto(T) $= \dim_{\mathbb{K}} Im(T)$ e nulidade(T) $= \dim_{\mathbb{K}} Nuc(T)$.*

Definição 3.10. *Seja $T : V \to W$ uma transformação linear. Dizemos que T é* ***injetora****, se para quaisquer $u, v \in V$ com $T(u) = T(v)$ tivermos $u = v$. Dizemos que T é* ***sobrejetora*** *se $Im(T) = W$. Se T for injetora e sobrejetora simultaneamente, dizemos que T é* ***bijetora****.*

Exemplo 3.11.

Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $T(x, y) = (x, y, x + y)$. Mostremos que T é injetora, mas não é sobrejetora. De fato, tome $(x, y), (z, w) \in \mathbb{R}^2$ tais que $T(x, y) = T(z, w)$. Então

$$(x, y, x + y) = (z, w, z + w),$$

o que implica em $x = z$ e $y = w$. Portanto $(x, y) = (z, w)$, ou seja, T é injetora.

Agora veja que

$$Im(T) = \{(x, y, x + y);\ x, y \in \mathbb{R}\} = \{x(1, 0, 1) + y(0, 1, 1);\ x, y \in \mathbb{R}\}.$$

Assim $\{(1, 0, 1), (0, 1, 1)\}$ é uma base para $Im(T)$ e, portanto, $\dim_{\mathbb{R}} Im(T) = 2 < 3 = \dim_{\mathbb{R}} \mathbb{R}^3$. Neste caso, T não pode ser sobrejetora pois $Im(T) \neq \mathbb{R}^3$.

PROPOSIÇÃO 3.12. *Seja $T : V \to W$ uma transformação linear. Então T é injetora se, e somente se, $Nuc(T) = \{\bar{0}_V\}$.*

DEMONSTRAÇÃO: Suponha que T seja injetora e tome $v \in Nuc(T)$. Então $T(v) = \bar{0}_W = T(\bar{0}_V)$, a segunda igualdade seguindo do fato de T ser linear. Como T é injetora temos $v = \bar{0}_V$, ou seja, $Nuc(T) = \{\bar{0}_V\}$. Reciprocamente, suponha que $Nuc(T) = \{\bar{0}_V\}$ e considere $u, v \in V$ tais que $T(u) = T(v)$. Como T é linear, temos $T(u - v) = T(u) - T(v) = \bar{0}_W$, o que significa que $u - v \in Nuc(T) = \{\bar{0}_V\}$. Portanto, $u = v$, implicando que T é injetora. ■

EXEMPLO 3.13. Sejam $n \geq 2$ e $T : \mathbb{R}^2 \to \mathbb{R}^n$ definida por

$$T(x, y) = (x, 0, \ldots, 0).$$

Observe que T não é injetora pois $(1, 5) \neq (1, 7)$ e $T(1, 5) = (1, 0, \ldots, 0) = T(1, 7)$. Também podemos verificar isso usando a proposição anterior, ou seja, provando que $Nuc(T) \neq \{(0, 0)\}$. Por definição, $(x, y) \in Nuc(T)$ se, e somente se, $T(x, y) = (x, 0, \ldots, 0) = (0, 0, \ldots, 0)$, ou seja, $x = 0$. Deste modo

$$Nuc(T) = \{(0, y);\ y \in \mathbb{R}\} \neq \{(0, 0)\}.$$

Agora observe que T também não é sobrejetora. De fato, pelo Teorema do Núcleo e da Imagem, temos que $\dim_{\mathbb{R}} Im(T) = \dim_{\mathbb{R}} \mathbb{R}^2 - \dim_{\mathbb{R}} Nuc(T) = 2 - 1 = 1$. Como $\dim_{\mathbb{R}} \mathbb{R}^n = n \geq 2$, segue que $Im(T) \neq \mathbb{R}^n$. implicando que T não é sobrejetora.

DEFINIÇÃO 3.14. *Dizemos que $T : V \to W$ é um **isomorfismo** se T for uma tranformação linear bijetora. Neste caso, dizemos que V e W são espaços vetoriais **isomorfos**.*

EXEMPLO 3.15.

1. Os espaços vetoriais $\mathbb{C}$ e $\mathbb{R}^2$ são isomorfos. Para ver isso, considere a tranformação linear $T : \mathbb{C} \to \mathbb{R}^2$ definida por $T(z) = (x, y)$, para todo $z = x + yi \in \mathbb{C}$. É fácil verificar que T é bijetora. De fato,

$$z \in Nuc(T) \Leftrightarrow T(z) = (0, 0) \Leftrightarrow (x, y) = (0, 0) \Leftrightarrow z = 0 + 0i = \bar{0}.$$

Portanto, $Nuc(T) = \{\bar{0}\}$, donde segue que T é injetora. Além disso, pelo Teorema do Núcleo e da Imagem,

$$2 = \dim_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{R}} Nuc(T) + \dim_{\mathbb{R}} Im(T) = 0 + \dim_{\mathbb{R}} Im(T) = \dim_{\mathbb{R}} Im(T).$$

Assim, como $Im(T)$ é um subespaço vetorial de $\mathbb{R}^2$, cuja dimensão também é 2, temos $Im(T) = \mathbb{R}^2$, o que prova que T é sobrejetora.

2. Os espaços $\mathcal{M}_{2\times 2}(\mathbb{R})$ e $\mathbb{R}^4$ são isomorfos via a transformação linear

$$\begin{array}{rcl} T : \mathcal{M}_{2\times 2}(\mathbb{R}) & \to & \mathbb{R}^4 \\ \begin{pmatrix} x & y \\ z & w \end{pmatrix} & \mapsto & (x, y, z, w). \end{array}$$

É deixado ao leitor verificar que T é bijetora.

O próximo passo é definir a inversa de uma transformação linear. Para isso, precisamos primeiramente definir a composição entre duas transformações lineares $T : V \to W$ e $R : W \to Z$ como

$$\begin{array}{rcl} R \circ T : V & \to & Z \\ v & \mapsto & R(T(v)). \end{array}$$

A aplicação $R \circ T$ é uma transformação linear e temos as seguintes propriedades:

Lema 3.16. *Sejam V, W, Z $\mathbb{K}$-espaços vetoriais, $T_1, T_2 : V \to W$ e $L_1, L_2 : W \to Z$ transformações lineares e $k \in \mathbb{K}$. Então*

(a) $L_1 \circ (T_1 + T_2) = L_1 \circ T_1 + L_1 \circ T_2$;

(b) $(L_1 + L_2) \circ T_1 = L_1 \circ T_1 + L_2 \circ T_1$;

(c) $k(L_1 \circ T_1) = (kL_1) \circ T_1 = L_1 \circ (kT_1)$;

(d) Se $T : V \to V$ for um operador linear, então $I \circ T = T \circ I = T$ onde I é o operador identidade.

Para o caso em que $T : V \to V$ é um operador linear denotaremos $T^n = T \circ \ldots \circ T$ (n vezes). Além disso, denotaremos $T^0 = I$ se T não for o operador nulo.

Definição 3.17. *Uma aplicação $T : V \to W$ entre espaços vetoriais é dita* ***inversível*** *se existir uma aplicação $T^{-1} : W \to V$ tal que $T^{-1} \circ T = I_V$ e $T \circ T^{-1} = I_W$, onde I_V e I_W são os operadores identidades em V e em W, respectivamente. Neste caso, T^{-1} é chamada* ***inversa*** *de T.*

Proposição 3.18. *Sejam $T : V \to W$ e $R : W \to Z$ aplicações inversíveis entre espaços vetoriais. Então* *(a)* T^{-1} *é inversível e* $(T^{-1})^{-1} = T$*;*

(b) $(kT)^{-1} = \frac{1}{k}T^{-1}$*, para todo* $k \in \mathbb{K}^*$*;*

(c) $R \circ T$ *é inversível e* $(R \circ T)^{-1} = T^{-1} \circ R^{-1}$*.*

Demonstração: O item (a) segue do fato de $T^{-1} \circ T = I_V$ e $T \circ T^{-1} = I_W$. O item (b) é válido pois

$$(kT) \circ (\frac{1}{k}T^{-1}) = (k\frac{1}{k})(T \circ T^{-1}) = 1I_W = I_W$$

e

$$(\frac{1}{k}T^{-1}) \circ (kT) = (\frac{1}{k}k)(T^{-1} \circ T) = 1I_V = I_V.$$

O item (c) segue das igualdades

$$(R \circ T) \circ (T^{-1} \circ R^{-1}) = R \circ (T \circ T^{-1}) \circ R^{-1} = (R \circ I_W) \circ R^{-1} = R \circ R^{-1} = I_Z$$

e

$$(T^{-1} \circ R^{-1}) \circ (R \circ T) = T^{-1} \circ (R^{-1} \circ R) \circ T = T^{-1} \circ (I_W \circ T) = T^{-1} \circ T = I_V.$$

∎

Proposição 3.19. *Uma aplicação* $T : V \to W$ *é bijetora se, e somente se,* T *é inversível.*

Demonstração: Suponha que T é bijetora. Então para todo $w \in W$ existe um único $v \in V$ tal que $w = T(v)$. Defina $T^{-1} : W \to V$ por $T^{-1}(w) = v$. Observe que

$$T \circ T^{-1}(w) = T(T^{-1}(w)) = T(v) = w \quad \text{e} \quad T^{-1} \circ T(v) = T^{-1}(T(v)) = T^{-1}(w) = v,$$

ou seja, T^{-1} é a inversa de T.

Reciprocamente, suponha que T seja inversível. Se $T(u) = T(v)$ então

$$u = I(u) = T^{-1} \circ T(u) = T^{-1} \circ T(v) = I(v) = v,$$

o que implica que T é injetora. Além disso, dado $w \in W$ tome $v = T^{-1}(w) \in V$. Então

$$T(v) = T(T^{-1}(w)) = T \circ T^{-1}(w) = I(w) = w,$$

o que implica que T é sobrejetora. Portanto, T é bijetora. ∎

Proposição 3.20. *Seja* $T : V \to W$ *uma transformação linear inversível. Então sua inversa* $T^{-1} : W \to V$ *também é uma transformação linear.*

Demonstração: Sejam $u, w \in W$ e $k \in \mathbb{K}$ quaisquer. Como T é inversível, então T é bijetora, ou seja, existem únicos $a, b \in V$ tais que $u = T(a)$ e $w = T(b)$. Como T é linear,

$$ku + w = kT(a) + T(b) = T(ka) + T(b) = T(ka + b).$$

Logo,

$$\begin{aligned} T^{-1}(ku + w) &= T^{-1}(T(ka + b)) = T^{-1} \circ T(ka + b)) = I_V(ka + b) \\ &= ka + b = kT^{-1}(T(a)) + T^{-1}(T(b)) = kT^{-1}(u) + T^{-1}(v). \end{aligned}$$

Como u, w e k são arbitrários, T^{-1} é uma transformação linear. ■

Exemplo 3.21. 1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear dada por $T(x, y) = (2x, -y)$. Determine T^{-1}.

Considere $\{(1, 0), (0, 1)\}$ a base canônica de $\mathbb{R}^2$. Como $T(1, 0) = (2, 0)$ e $T(0, 1) = (0, -1)$ temos que $T^{-1}(2, 0) = (1, 0)$ e $T^{-1}(0, -1) = (0, 1)$. Note que $\{(2, 0), (0, -1)\}$ é uma base de $\mathbb{R}^2$. Logo, todo $(x, y) \in \mathbb{R}^2$ pode ser escrito de modo único como

$$(x, y) = \frac{x}{2}(2, 0) - y(0, -1).$$

Como T é uma transformação linear, T^{-1} também é e, portanto,

$$\begin{aligned} T^{-1}(x, y) &= T^{-1}\left(\frac{x}{2}(2, 0) - y(0, -1)\right) = \frac{x}{2}T^{-1}(2, 0) - yT^{-1}(0, -1) \\ &= \frac{x}{2}(1, 0) - y(0, 1) = \left(\frac{x}{2}, -y\right). \end{aligned}$$

Logo, $T^{-1}(x, y) = \left(\frac{x}{2}, -y\right)$.

2. Seja $T : \mathcal{P}_3(\mathbb{R}) \to \mathcal{M}_{2\times 2}(\mathbb{R})$ a transformação linear definida por

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{pmatrix} a_0 & 2a_1 \\ -a_3 & a_2 \end{pmatrix}.$$

Seja $\mathcal{B} = \{1, x, x^2, x^3\}$ a base canônica de $\mathcal{P}_3(\mathbb{R})$. Observe que

$$T(1) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad T(x) = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \quad T(x^2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{e} \quad T(x^3) = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}.$$

Então

$$T^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1, \quad T^{-1}\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = x, \quad T^{-1}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = x^2 \quad \text{e} \quad T^{-1}\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} = x^3.$$

Como

$$\mathcal{C} = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

é uma base de $\mathcal{M}_{2\times 2}(\mathbb{R})$, toda matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{R})$ pode ser escrita de modo único como

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{b}{2} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} - c \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Então

$$\begin{aligned} T^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= T^{-1} \left(a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{b}{2} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} - c \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right) \\ &= aT^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{b}{2} T^{-1} \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} - cT^{-1} \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} + dT^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= a1 + \frac{b}{2}x - cx^3 + dx^2 \\ &= a + \frac{b}{2}x + dx^2 - cx^3. \end{aligned}$$

Logo, $T^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + \frac{b}{2}x + dx^2 - cx^3.$

Teorema 3.22. *Seja $T : V \to W$ uma transformação linear. Então T é injetora se, e somente se, T leva todo subconjunto LI de V em um subconjunto LI de W.*

Demonstração: Suponha que T é injetora e seja $S = \{v_1, \ldots, v_k\}$ um subconjunto LI de V. Provemos que $T(v_1), \ldots, T(v_k)$ são vetores LI de W. De fato, sejam $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ tais que

$$\lambda_1 T(v_1) + \ldots + \lambda_k T(v_k) = \bar{0}_W.$$

Então $T(\lambda_1 v_1 + \ldots + \lambda_k v_k) = \bar{0}_W$, ou seja, $\lambda_1 v_1 + \ldots + \lambda_k v_k \in Nuc(T)$. Como T é injetora, temos $Nuc(T) = \{\bar{0}_V\}$, donde concluímos que $\lambda_1 v_1 + \ldots + \lambda_k v_k = \bar{0}_V$. Como os vetores $v_1, \ldots, v_k$ são LI, segue que $\lambda_i = 0$ para todo $i = 1, \ldots, k$. Portanto, $\{T(v_1), \ldots, T(v_k)\}$ é LI.

Reciprocamente, seja $v \in V$ um vetor não nulo. Então $S = \{v\}$ é um subconjunto LI de V. Por hipótese, $\{T(v)\}$ é LI e portanto $T(v) \neq \bar{0}_W$. Logo, se $T(v) = \bar{0}_W$ então $v = \bar{0}_V$, ou seja, $Nuc(T) = \{\bar{0}_V\}$. Portanto, T é injetora. ∎

Teorema 3.23. *Se V e W são $\mathbb{K}$-espaços vetoriais isomorfos então* $\dim_K V = \dim_K W$.

Demonstração: Se V e W são espaços vetoriais isomorfos então existe um isomorfismo $T : V \to W$. Pela Proposição 3.20, $T^{-1} : W \to V$ também é uma transformação linear. Além disso, pela Proposição 3.18 item (a), a aplicação T^{-1} também é inversível e, portanto, T^{-1} é um isomorfismo. Suponha que $\dim_{\mathbb{K}} V = n$ e $\dim_{\mathbb{K}} W = m$. Se $\mathcal{B} = \{v_1, \ldots, v_n\}$

for uma base de V, então pelo teorema anterior $\{T(v_1), \ldots, T(v_n)\}$ é LI em W. Como $\dim_{\mathbb{K}} W = m$, temos que $n \leq m$ (pois todo subconjunto de W com mais do que m vetores é LD). Do mesmo modo, se $\mathcal{C} = \{w_1, \ldots, w_m\}$ for uma base de W, então pelo teorema anterior $\{T^{-1}(w_1), \ldots, T^{-1}(w_m)\}$ é LI em V. Como $\dim_{\mathbb{K}} V = n$, temos que $m \leq n$ (pois todo subconjunto de V com mais do que n vetores é LD). Assim, $\dim_{\mathbb{K}} V = n = m = \dim_{\mathbb{K}} W$. ■

Teorema 3.24. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$ tais que* $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$, *ambas finitas. Se $T : V \to W$ for uma transformação linear, são equivalentes:*

(i) T é inversível;

(ii) T é injetora;

(iii) T é sobrejetora;

(iv) T leva base de V em base de W.

Demonstração: $(i) \Rightarrow (ii)$ Óbvio.

$(ii) \Rightarrow (iii)$ Como T é injetora temos que $Nuc(T) = \{\bar{0}_V\}$, donde segue do Teorema do Núcleo e da Imagem que $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} Im(T)$. Por hipótese, $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W$ e portanto $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} Im(T)$. Como $Im(T)$ é um subespaço vetorial de W e ambos possuem a mesma dimensão, concluímos que $Im(T) = W$. Portanto, T é sobrejetora.

$(iii) \Rightarrow (ii)$ Como T é sobrejetora segue do Teorema do Núcleo e da Imagem que $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} Nuc(T) + \dim_{\mathbb{K}} W$. Logo, $\dim_{\mathbb{K}} Nuc(T) = 0$, pois por hipótese $\dim_{\mathbb{K}} V = dim_{\mathbb{K}} Im(T)$. Isso implica que $Nuc(T) = \{\bar{0}_V\}$, ou seja, T é injetora.

$(ii) \Rightarrow (iv)$ Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base de V. Pelo Teorema anterior $\mathcal{C} = \{T(v_1), \ldots, T(v_n)\}$ é um conjunto LI de W, pois $\mathcal{B}$ é LI. Como $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} V = n$, temos que $\mathcal{C}$ é uma base de W. Logo, T leva a base $\mathcal{B}$ de V na base $\mathcal{C}$ de W.

$(iv) \Rightarrow (i)$ Seja $\mathcal{B} = \{v_1, \ldots, v_n\}$ uma base de V. Por hipótese, $\mathcal{C} = \{T(v_1), \ldots, T(v_n)\}$ é uma base de W. Como os vetores $T(v_i)$ geram W, com $1 \leq i \leq n$, temos que todo vetor de W pertence a $Im(T)$ e, portanto, $Im(T) = W$. Seja agora $v \in Nuc(T)$ escrito como $v = k_1 v_1 + \ldots + k_n v_n$. Como T é linear, temos $\bar{0}_W = T(v) = k_1 T(v_1) + \ldots + k_n T(v_n)$, o que implica que $k_1 = \ldots = k_n = 0$ pois $\mathcal{C}$ é LI. Portanto, $v = \bar{0}_V$ implicando que T é injetora. Logo, T é bijetora e pela Proposição 3.19 temos que T é inversível. ■

Corolário 3.25. *Se V e W são $\mathbb{K}$-espaços vetoriais de mesma dimensão então eles são isomorfos.*

Demonstração: Suponha que V e W sejam $\mathbb{K}$-espaços vetoriais de mesma dimensão com bases ordenadas $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_n\}$, respectivamente. Pelo Teorema

3.5, existe uma única transformação linear $T : V \to W$ tal que $T(v_i) = w_i$, para cada $1 \leq i \leq n$. Pelo Teorema 3.24, como T leva base de V em base de W segue que T é inversível e, portanto, T é um isomorfismo. Logo, V e W são espaços vetoriais isomorfos. ■

3.3 Matriz de uma Transformação Linear

Nesta seção, veremos que é sempre possível associar a uma transformação linear $T : V \to W$ uma matriz de ordem $m \times n$, onde $m = \dim_{\mathbb{K}} W$ e $n = \dim_{\mathbb{K}} V$. Chamaremos tal matriz de **matriz associada a** T, cujas colunas são os vetores $T(v_1), \ldots, T(v_n)$, para $\{v_1, \ldots, v_n\}$ uma base de V.

Consideremos então V e W dois $\mathbb{K}$-espaços vetoriais de dimensão n e m, respectivamente. Sejam $T : V \to W$ uma transformação linear, $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ bases de V e W, respectivamente. Dado $v \in V$, escreva-o na base $\mathcal{B}$ unicamente como $v = k_1 v_1 + \ldots + k_n v_n$ com $k_i \in \mathbb{K}$ e $i = 1, \ldots, n$. Logo, as coordenadas de v na base $\mathcal{B}$ é

$$[v]_{\mathcal{B}} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}.$$

Além disso,

$$T(v) = T(k_1 v_1 + \ldots + k_n v_n) = k_1 T(v_1) + \ldots + k_n T(v_n). \tag{3.1}$$

Como $T(v) \in W$, podemos escrevê-lo como $T(v) = l_1 w_1 + \ldots + l_m w_m$ com $l_j \in \mathbb{K}$ e $j = 1, \ldots, m$, ou seja, as coordenadas de $T(v)$ na base $\mathcal{C}$ é

$$[T(v)]_{\mathcal{C}} = \begin{pmatrix} l_1 \\ \vdots \\ l_m \end{pmatrix}.$$

De modo análogo, como $T(v_i) \in W$ para cada $i = 1, \ldots, n$, então

$$\left\{ \begin{array}{c} T(v_1) = a_{11} w_1 + a_{21} w_2 + \ldots + a_{m1} w_m \\ T(v_2) = a_{12} w_1 + a_{22} w_2 + \ldots + a_{m2} w_m \\ \vdots \\ T(v_n) = a_{1n} w_1 + a_{2n} w_2 + \ldots + a_{mn} w_m. \end{array} \right. \tag{3.2}$$

Substituindo (3.2) em (3.1) temos

$$\begin{aligned} T(v) &= k_1(a_{11} w_1 + a_{21} w_2 + \ldots + a_{m1} w_m) + \ldots + k_n(a_{1n} w_1 + a_{2n} w_2 + \ldots + a_{mn} w_m) \\ &= (a_{11} k_1 + \ldots + a_{1n} k_n) w_1 + \ldots + (a_{m1} k_1 + \ldots + a_{mn} k_n) w_m. \end{aligned}$$

Pela unicidade da representação de $T(v)$ na base $\mathcal{C}$, temos que

$$\begin{cases} l_1 = a_{11}k_1 + \ldots + a_{1n}k_n \\ \vdots \\ l_m = a_{m1}k_1 + \ldots + a_{mn}k_n \end{cases},$$

o que na forma matricial nos dá

$$\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix},$$

ou seja,

$$[T(v)]_{\mathcal{C}} = A[v]_{\mathcal{B}},$$

onde $A = (a_{ij})_{m\times n}$. Tal matriz, que a partir de agora será denotada por $[T]_{\mathcal{C}}^{\mathcal{B}}$, é chamada **matriz de T em relação às bases $\mathcal{B}$ e $\mathcal{C}$**. Temos então o seguinte resultado:

TEOREMA 3.26. *Sejam V e W dois $\mathbb{K}$-espaços vetoriais tais que $\dim_{\mathbb{K}} V = n$ e $\dim_{\mathbb{K}} W = m$. Sejam $\mathcal{B}$ uma base ordenada de V e $\mathcal{C}$ uma base ordenada de W. Para cada transformação linear $T : V \to W$ existe uma matriz $[T]_{\mathcal{C}}^{\mathcal{B}}$ de ordem $m \times n$ sobre o corpo $\mathbb{K}$ tal que*

$$[T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{B}}[v]_{\mathcal{B}},$$

para todo $v \in V$. Além disso, a aplicação $\Psi : \mathcal{L}(V,W) \longrightarrow \mathcal{M}_{m\times n}(\mathbb{K})$ que associa cada transformação linear T à sua matriz $[T]_{\mathcal{C}}^{\mathcal{B}}$ é bijetora.

COROLÁRIO 3.27. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$ tais que $\dim_{\mathbb{K}} V = n$ e $\dim_{\mathbb{K}} W = m$. Então $\dim_{\mathbb{K}} \mathcal{L}(V,W) = mn$.*

DEMONSTRAÇÃO: Esse resultado segue diretamente do teorema anterior, uma vez que como $\mathcal{L}(V,W)$ e $\mathcal{M}_{m\times n}(\mathbb{K})$ são espaços vetoriais isomorfos então $\dim_{\mathbb{K}} \mathcal{L}(V,W) = \dim_K \mathcal{M}_{m\times n}(\mathbb{K}) = mn$. ■

EXEMPLO 3.28. 1. Sejam $I : V \to V$ o operador identidade, $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_n\}$ bases de V. Então cada $v_i \in V$ pode ser escrito como combinação linear dos vetores da base $\mathcal{C}$, ou seja, existem $a_{ij} \in \mathbb{R}$ tais que

$$\begin{cases} I(v_1) = v_1 = a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ I(v_2) = v_2 = a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n \\ \vdots \\ I(v_n) = v_n = a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n \end{cases}.$$

Desse modo, a matriz de I em relação às bases $\mathcal{B}$ e $\mathcal{C}$ é dada por

$$[I]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

que é exatamente a matriz mudança de base da base $\mathcal{B}$ para a base $\mathcal{C}$.

2. Seja $T : \mathbb{R}^3 \to \mathbb{R}^2$ a transformação linear dada por $T(x, y, z) = (2x + y - z, 3x - 2y + 4z)$ e sejam $\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ e $\mathcal{C} = \{(1, 3), (1, 4)\}$ bases de $\mathbb{R}^3$ e $\mathbb{R}^2$, respectivamente. Então

$$T(1, 1, 1) = (2, 5) = 3(1, 3) - 1(1, 4),$$
$$T(1, 1, 0) = (3, 1) = 11(1, 3) - 8(1, 4),$$
$$T(1, 0, 0) = (2, 3) = 5(1, 3) - 3(1, 4),$$

implicando que a matriz de T em relação às bases $\mathcal{B}$ e $\mathcal{C}$ é a matriz

$$[T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{pmatrix}.$$

3. Dadas as bases $\mathcal{B} = \{(1, 1), (0, 1)\}$ de $\mathbb{R}^2$ e $\mathcal{C} = \{(0, 3, 0), (-1, 0, 0), (0, 1, 1)\}$ de $\mathbb{R}^3$, determinemos a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^3$ cuja matriz é

$$[T]_{\mathcal{C}}^{\mathcal{B}} = \begin{pmatrix} 0 & 2 \\ -1 & 0 \\ -1 & 3 \end{pmatrix}.$$

Para isso, lembremos que para todo $v \in \mathbb{R}^2$ temos $[T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{B}}[v]_{\mathcal{B}}$. Como $v = (x, y) = x(1, 1) + (y - x)(0, 1)$ segue que

$$[v]_{\mathcal{B}} = \begin{pmatrix} x \\ y - x \end{pmatrix}.$$

Assim,

$$[T(v)]_{\mathcal{C}} = \begin{pmatrix} 0 & 2 \\ -1 & 0 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y - x \end{pmatrix} = \begin{pmatrix} 2(y - x) \\ -x \\ -x + 3(y - x) \end{pmatrix} = \begin{pmatrix} -2x + 2y \\ -x \\ -4x + 3y \end{pmatrix}.$$

Portanto,

$$\begin{aligned} T(x, y) &= (-2x + 2y)(0, 3, 0) - x(-1, 0, 0) + (-4x + 3y)(0, 1, 1) \\ &= (x, -10x + 9y, -4x + 3y), \end{aligned}$$

para todo $(x, y) \in \mathbb{R}^2$.

A partir de agora, denotaremos a matriz de um operador $T : V \to V$ em relação a uma mesma base $\mathcal{B}$ simplesmente por $[T]_{\mathcal{B}}$ ao invés de $[T]_{\mathcal{B}}^{\mathcal{B}}$.

Teorema 3.29. *Seja V um $\mathbb{K}$-espaço vetorial cuja dimensão é n e sejam $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_n\}$ bases ordenadas de V. Suponha que $T : V \to V$ seja um operador linear. Então existe uma única matriz inversível P de ordem n tal que*

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P.$$

Demonstração: Provamos no final do Capítulo 2 que as coordenadas de todo vetor $v \in V$ com relação à base $\mathcal{B}$ é dada pela igualdade

$$[v]_{\mathcal{B}} = [I]_{\mathcal{B}}^{\mathcal{C}}[v]_{\mathcal{C}}, \tag{3.3}$$

onde $[I]_{\mathcal{B}}^{\mathcal{C}}$ é a matriz de mudança da base $\mathcal{C}$ para a base $\mathcal{B}$. Como $T(v) \in V$, de (3.3) temos também que $[T(v)]_{\mathcal{B}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T(v)]_{\mathcal{C}}$. Além disso, pelo Teorema 3.26 temos que $[T(v)]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}$ e $[T(v)]_{\mathcal{C}} = [T]_{\mathcal{C}}[v]_{\mathcal{C}}$. Assim,

$$[T(v)]_{\mathcal{B}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T(v)]_{\mathcal{C}} \Leftrightarrow [T]_{\mathcal{B}}[v]_{\mathcal{B}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{C}}[v]_{\mathcal{C}} \Leftrightarrow [T]_{\mathcal{B}}[I]_{\mathcal{B}}^{\mathcal{C}}[v]_{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{C}}[v]_{\mathcal{C}}.$$

Como v é arbitrário, $[T]_{\mathcal{B}}[I]_{\mathcal{B}}^{\mathcal{C}} = [I]_{\mathcal{B}}^{\mathcal{C}}[T]_{\mathcal{C}}$, ou seja,

$$[T]_{\mathcal{C}} = ([I]_{\mathcal{B}}^{\mathcal{C}})^{-1}[T]_{\mathcal{B}}[I]_{\mathcal{B}}^{\mathcal{C}}.$$

Como $[I]_{\mathcal{B}}^{\mathcal{C}}$ é uma matriz inversível, o resultado segue tomando-se $P = [I]_{\mathcal{B}}^{\mathcal{C}}$. ■

O teorema anterior nos diz que as matrizes de um operador linear em relação às quaisquer duas bases de V são semelhantes. Mais especificamente, duas matrizes $A, B \in \mathcal{M}_{n\times n}(\mathbb{K})$ são **semelhantes** se existir uma matriz inversível $P \in \mathcal{M}_{n\times n}(\mathbb{K})$ tal que $B = P^{-1}AP$.

Teorema 3.30. *Sejam U, V e W $\mathbb{K}$-espaços vetoriais com bases $\mathcal{B}, \mathcal{C}$ e $\mathcal{D}$ respectivamente e considere $T : U \to V$ e $L : V \to W$ transformações lineares. Então*

1. $[L \circ T]_{\mathcal{D}}^{\mathcal{B}} = [L]_{\mathcal{D}}^{\mathcal{C}}[T]_{\mathcal{C}}^{\mathcal{B}}$;

2. *T é inversível se, e somente se,* $\det([T]_{\mathcal{C}}^{\mathcal{B}}) \neq 0$. *Neste caso,* $[T^{-1}]_{\mathcal{B}}^{\mathcal{C}} = ([T]_{\mathcal{C}}^{\mathcal{B}})^{-1}$.

De um modo particular temos:

Corolário 3.31. *Um operador linear $T : V \to V$ é inversível se, e somente se, o determinante da matriz de T em relação à qualquer base de V é diferente de zero.*

Demonstração: Seja $\mathcal{B}$ uma base de V e denote por $[T]_{\mathcal{B}}$ a matriz de T em relação à base $\mathcal{B}$. Temos que T é inversível se, e somente se, existe $T^{-1} : V \to V$ tal que $T \circ T^{-1} = I$ onde $I : V \to V$ é o operador identidade. Pelo item *1.* do teorema anterior,

$$I_n = [I]_{\mathcal{B}} = [T \circ T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}[T^{-1}]_{\mathcal{B}},$$

onde I_n é a matriz identidade de ordem $n = \dim_{\mathbb{K}} V$. Logo, $\det([T]_{\mathcal{B}}) \neq 0$. Reciprocamente, suponha que $\det([T]_{\mathcal{B}}) \neq 0$ e seja $v \in V$ tal que $T(v) = \bar{0}_V$. Pelo Teorema 3.26 temos que

$$[T]_{\mathcal{B}}[v]_{\mathcal{B}} = 0_{n\times 1},$$

onde $0_{n\times 1}$ é a matriz coluna nula, ou seja, a igualdade acima determina um sistema linear homogêneo. Como $\det([T]_{\mathcal{B}}) \neq 0$ podemos aplicar a Regra de Cramer para concluir que tal sistema admite somente a solução trivial, ou seja, $[v]_{\mathcal{B}} = 0_{n\times 1}$. Assim $v = \bar{0}_V$. Portanto $Nuc(T) = \{\bar{0}_V\}$, o que implica que T é injetora. Pelo Teorema 3.24, T é bijetora e, portanto, inversível.

Seja agora $\mathcal{C}$ uma outra base de V e denote por $[T]_{\mathcal{C}}$ a matriz de T em relação à base $\mathcal{C}$. Sabemos que $[T]_{\mathcal{B}}$ e $[T]_{\mathcal{C}}$ são semelhantes, ou seja, existe uma matriz inversível P tal que $[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$. Então

$$\det([T]_{\mathcal{C}}) = \det(P^{-1}[T]_{\mathcal{B}}P) = \det(P^{-1})\det([T]_{\mathcal{B}})\det(P) = \det([T]_{\mathcal{B}}).$$

Portanto, o determinante da matriz de T independe da base escolhida. Assim, T é inversível se, e somente se, o determinante de T é não nulo em relação à qualquer base de V. ■

3.4 Autovalores e Autovetores

Em toda esta seção $T : V \to V$ denotará um operador linear sobre um $\mathbb{K}$-espaço vetorial n-dimensional V. Como o último corolário nos garante que o determinante da matriz de T independe da base escolhida, vamos denotar a matriz de T em relação a uma base $\mathcal{B}$ fixada simplesmente por $[T]$. Faremos o mesmo com outros operadores.

Definição 3.32. *Dizemos que $\lambda \in \mathbb{K}$ é um* ***autovalor*** *de T (ou valor característico de T) se existir um vetor não nulo $v \in V$ tal que $T(v) = \lambda v$. Se λ for um autovalor de T então todo $v \in V$ tal que $T(v) = \lambda v$ é denominado um* ***autovetor*** *de T (ou vetor característico de T) associado ao autovalor λ.*

Observação 3.33.
1. Se v for um autovetor de T associado ao λ então, para todo $k \in \mathbb{K}$ não nulo, kv é um autovetor de T associado ao λ. De fato, $T(kv) = kT(v) = k\lambda v = \lambda(kv)$.

2. Autovetores associados a autovalores distintos são LI. Mostraremos essa afirmação para o caso de dois autovalores distintos. A demonstração para o caso de n autovalores distintos, com $n > 2$, é análoga (exercício!). Suponha v_1 e v_2 autovetores de T associados aos autovalores λ_1 e λ_2, respectivamente, com $\lambda_1 \neq \lambda_2$. Então $T(v_1) = \lambda_1 v_1$ e $T(v_2) = \lambda_2 v_2$. Considere a combinação linear trivial

$$k_1 v_1 + k_2 v_2 = \bar{0}_V. \tag{3.4}$$

Devemos mostrar que $k_1 = k_2 = 0$. Aplicando $T - \lambda_2 I$ a ambos os lados de (3.4) temos

$$\begin{aligned}(T - \lambda_2 I)(k_1 v_1 + k_2 v_2) = \bar{0}_V &\Leftrightarrow k_1 T(v_1) + k_2 T(v_2) - \lambda_2 k_1 v_1 - \lambda_2 k_2 v_2 = \bar{0}_V\\ &\Leftrightarrow k_1 \lambda_1 v_1 + k_2 \lambda_2 v_2 - k_1 \lambda_2 v_1 - k_2 \lambda_2 v_2 = \bar{0}_V\\ &\Leftrightarrow k_1(\lambda_1 - \lambda_2) v_1 = \bar{0}_V.\end{aligned}$$

Como $v_1 \neq \bar{0}_V$ e $\lambda_1 \neq \lambda_2$ segue que $k_1 = 0$. Substituindo em (3.4) temos $k_2 v_2 = \bar{0}_V$. Como $v_2 \neq \bar{0}_V$ segue que $k_2 = 0$.

Para cada autovalor λ de T, denotaremos por V_λ o conjunto de todos os autovetores de T associados ao λ juntamente com o vetor nulo, ou seja,

$$V_\lambda = \{v \in V;\ T(v) = \lambda v\}.$$

V_λ é denominado **o autoespaço correspondente ao autovalor** λ e é um $\mathbb{K}$-subespaço vetorial de V. De fato, para quaisquer $u, v \in V_\lambda$ e $k \in \mathbb{K}$ temos

$$T(ku + v) = kT(u) + T(v) = k\lambda u + \lambda v = \lambda(ku + v),$$

o que implica que $ku + v \in V_\lambda$.

TEOREMA 3.34. *Sejam $T : V \to V$ um operador linear e $I : V \to V$ o operador identidade. As seguintes condições são equivalentes:*

1. *λ é um autovalor de T;*

2. *O operador $T - \lambda I : V \to V$ é não inversível;*

3. $\det([T - \lambda I]) = 0$.

DEMONSTRAÇÃO: *1.* $\Leftrightarrow$ *2.* Temos que λ é um autovalor de T se, e somente se, existe $v \neq \bar{0}_V$ tal que $T(v) = \lambda v$. Observe que

$$T(v) = \lambda v \ \Leftrightarrow\ T(v) = \lambda I(v) \ \Leftrightarrow\ (T - \lambda I)(v) = \bar{0}_V,$$

ou seja, existe $v \neq \bar{0}_V$ tal que $v \in Nuc(T - \lambda I)$. Isso é equivalente a dizer que $T - \lambda I$ não é injetora, o que ocorre se, e somente se, $T - \lambda I$ não é inversível (Teorema 3.24).

2. $\Leftrightarrow$ *3.* Segue diretamente do Corolário 3.31, ou seja, o operador $T-\lambda I$ é não inversível se, e somente se, $\det([T-\lambda I]) = 0$. ■

Observe que $\det([T-\lambda I]) = 0$ se, e somente se, $\det([\lambda I - T]) = (-1)^n \det([T-\lambda I]) = 0$. Além disso,

$$[\lambda I - T] = [\lambda I] - [T] = \lambda [I] - [T] = \lambda I_n - [T],$$

onde I_n é a matriz identidade de ordem n. Portanto, λ é um autovalor de T se, e somente se, $\det(\lambda I_n - [T]) = 0$. Temos assim a seguinte definição:

Definição 3.35. *O* ***polinômio característico*** *de um operador linear $T : V \to V$ é definido como o polinômio mônico*[1] *de grau n dado por*

$$p(x) = \det(xI_n - [T]).$$

Observe que as raízes de $p(x)$ são precisamente os autovalores de T. Além disso, na definição acima a matriz $[T]$ está escrita em relação a uma base de V fixada. Veremos adiante que tal definição independe da base na qual associamos a matriz de T.

Exemplo 3.36. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ definido por $T(x, y) = (3x, 8x - y)$. Considerando $[T]$ a matriz de T na base canônica de $\mathbb{R}^2$ temos

$$xI_2 - [T] = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 8 & -1 \end{pmatrix} = \begin{pmatrix} x-3 & 0 \\ -8 & x+1 \end{pmatrix}.$$

O polinômio característico é dado por $p(x) = \det(xI_2 - [T]) = x^2 - 2x - 3$ cujas raízes são $x' = -1$ e $x'' = 3$, que são os autovalores de T. Os autovetores $v = (x, y)$ associados ao autovalor $\lambda_1 = -1$ satisfazem:

$$(-I_2 - [T])(v) = \begin{pmatrix} -4 & 0 \\ -8 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow -4x = -8x = 0 \Rightarrow x = 0.$$

Logo, todos os autovetores associados a $\lambda_1 = -1$ são da forma $(0, y)$, com $y \in \mathbb{R}$ não nulo. Assim, o autoespaço associado ao autovalor $\lambda_1 = -1$ é dado por $V_{-1} = \{(0, y); y \in \mathbb{R}\}$. Os autovetores $v = (x, y)$ associados ao autovalor $\lambda_2 = 3$ satisfazem:

$$(3I_2 - [T])(v) = \begin{pmatrix} 0 & 0 \\ -8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow -8x + 4y = 0 \Rightarrow y = 2x.$$

Logo, todos os autovetores associados a $\lambda_2 = 3$ são da forma $(x, 2x)$, com $x \in \mathbb{R}$ não nulo. Assim, o autoespaço associado ao autovalor $\lambda_2 = 3$ é dado por $V_3 = \{(x, 2x); x \in \mathbb{R}\}$.

De modo análogo ao que fizemos para operadores, podemos definir autovalor e polinômio característico para uma matriz quadrada A de ordem n arbitrária:

[1] Polinômio mônico é um polinômio em que o coeficiente que acompanha o termo de maior grau é 1.

Definição 3.37. *Sejam $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ e I_n a matriz identidade de ordem n. Dizemos que $\lambda \in \mathbb{K}$ é um autovalor de A se a matriz $\lambda I_n - A$ for não inversível, ou seja,* $\det(\lambda I_n - A) = 0$. *Definimos o polinômio característico de A como o polinômio mônico de grau n dado por* $p(x) = \det(xI_n - A)$.

Lema 3.38. *Matrizes semelhantes possuem o mesmo polinômio característico.*

Demonstração: Sejam A, $B \in \mathcal{M}_{n\times n}(\mathbb{K})$ semelhantes, ou seja, existe uma matriz inversível P de ordem n tal que $B = P^{-1}AP$. Então

$$\begin{aligned}
\det(xI_n - B) &= \det(xI_n - P^{-1}AP) = \det(P^{-1}PxI_n - P^{-1}AP) \\
&= \det(P^{-1}xI_nP - P^{-1}AP) = \det(P^{-1}(xI_n - A)P) \\
&= \det(P^{-1})\det(xI_n - A)\det(P) \\
&= \det(xI_n - A).
\end{aligned}$$

Portanto, A e B possuem o mesmo polinômio característico. ∎

Como o polinômio característico de um operador linear T é definido a partir de sua matriz em relação a alguma base de V e como as matrizes de T em relação às quaisquer duas bases de V são semelhantes, o lema anterior nos garante que o polinômio característico de T não depende da base na qual associamos sua matriz. Portanto, faz sentido considerar o polinômio característico de T como o polinômio característico de qualquer matriz associada a T.

Exemplo 3.39. Nem todo operador linear admite autovalores e, consequentemente, autovetores. Como um exemplo, considere o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ cuja matriz em relação à base canônica é

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

Então, o polinômio característico de T é dado por

$$p(x) = \det(xI_2 - A) = \det\begin{pmatrix} x & 1 \\ -1 & x \end{pmatrix} = x^2 + 1,$$

que não possui raízes reais. Portanto, T não possui autovalores. No entanto, se $V = \mathbb{C}^2$ é um espaço vetorial sobre $\mathbb{C}$ e $L : \mathbb{C}^2 \to \mathbb{C}^2$ é um operador linear que possui a mesma matriz associada A em relação à base canônica de $\mathbb{C}$, então $p(x)$ possui como raízes os escalares i e $-i$. Portanto, os autovalores de L são $\lambda_1 = i$ e $\lambda_2 = -i$. Neste caso, cálculos diretos nos mostram que os autoespaços associados a cada autovalor são:

$$V_i = \{(z, -iz);\ z \in \mathbb{C}\} \quad \text{e} \quad V_{-i} = \{(z, iz);\ z \in \mathbb{C}\}.$$

Definição 3.40. *Sejam V um $\mathbb{K}$-espaço vetorial, $T : V \to V$ um operador linear e $\lambda \in \mathbb{K}$ um autovalor de T. O número de vezes que $(x - \lambda)$ aparece como fator do polinômio característico de T é denominado* ***multiplicidade algébrica***[2] *de λ, cuja notação é $m_a(\lambda)$. A dimensão do autoespaço V_λ sobre $\mathbb{K}$ é denominada* ***multiplicidade geométrica*** *de λ, cuja notação é $m_g(\lambda)$.*

Exemplo 3.41. Seja $T : \mathbb{R}^4 \to \mathbb{R}^4$ um operador linear cuja matriz em relação a alguma base de $\mathbb{R}^4$ é dada por

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

Assim, o polinômio característico de T é dado por

$$p(x) = \det(xI_4 - A) = \det \begin{pmatrix} x-2 & -1 & 0 & 0 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x-2 & 0 \\ 0 & 0 & 0 & x-2 \end{pmatrix} = (x-2)^4.$$

Portanto, $\lambda = 2$ é o único autovalor de T de multiplicidade algébrica $m_a(2) = 4$. Agora, os autovetores $v = (x, y, z, w)$ associados ao autovalor $\lambda = 2$ satisfazem:

$$(2I_4 - A)(v) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow\; y = 0.$$

Logo, todos os autovetores associados a $\lambda = 2$ são da forma $(x, 0, z, w)$, com $x, z, w \in \mathbb{R}$ não todos nulos. Assim, o autoespaço associado ao autovalor $\lambda = 2$ é

$$V_2 = \{(x, 0, z, w);\; x, z, w \in \mathbb{R}\},$$

cuja dimensão é 3 (basta notar que $\{(1,0,0,0), (0,0,1,0), (0,0,0,1)\}$ é uma base de V_2). Portanto, a multiplicidade geométrica de $\lambda = 2$ é $m_g(2) = \dim_{\mathbb{R}}(V_2) = 3$. Observe que neste exemplo $m_a(2) \neq m_g(2)$.

3.5 Diagonalização de Operadores Lineares

Vimos na Seção 3.3 que todo operador linear definido em um espaço V de dimensão finita pode ser representado por uma matriz. Sendo as matrizes diagonais as mais simples, do

[2] Em outras palavras, a multiplicidade algébrica de λ é a multiplicidade de λ como raiz do polinômio característico.

ponto de vista das operações matriciais, nos perguntamos se é sempre possível representar um operador T por uma matriz diagonal. Mais precisamente, queremos saber se para todo operador T existe uma base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ é uma matriz diagonal. A resposta é que nem sempre existe tal base e, nesta seção, vamos verificar quais os casos em que ela existe.

Definição 3.42. *Seja $T : V \to V$ um operador linear sobre um $\mathbb{K}$-espaço vetorial V de dimensão finita. Dizemos que T é* ***diagonalizável*** *se existir uma base de V formada somente por autovetores de T.*

Uma justificativa para a nomenclatura introduzida na definição acima é a seguinte: suponha $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ uma base formada por autovetores de T, com cada autovetor v_i associado a um autovalor λ_i. Por definição de autovetor, temos

$$\begin{aligned} T(v_1) &= \lambda_1 v_1 = \lambda_1 v_1 + 0v_2 + \ldots + 0v_n \\ T(v_2) &= \lambda_2 v_2 = 0v_1 + \lambda_2 v_2 + \ldots + 0v_n \\ &\vdots \\ T(v_n) &= \lambda_n v_n = 0v_1 + 0v_2 + \ldots + \lambda_n v_n. \end{aligned} \tag{3.5}$$

Portanto, a matriz de T em relação à base $\mathcal{B}$ é dada por

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \tag{3.6}$$

que é uma matriz diagonal. Reciprocamente, se $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ for uma base segundo à qual a matriz de T é diagonal, digamos como em (3.6), então por definição de matriz associada temos que valem as igualdades em (3.5), ou seja, cada v_i é um autovetor de T. Portanto, T admite uma representação diagonal se, e somente se, V possui uma base formada por autovetores de T.

Proposição 3.43. *Seja V um $\mathbb{K}$-espaço vetorial de dimensão finita n. Se $T : V \to V$ é um operador linear que possui n autovalores distintos, então T é diagonalizável.*

Demonstração: Sejam $\lambda_1, \ldots, \lambda_n$ autovalores distintos de T. Para cada $i \in \{1, \ldots, n\}$, considere um autovetor v_i associado a λ_i. Pelo item 2. da Observação 3.33, quaisquer n autovetores associados a n autovalores distintos são LI, ou seja, $\{v_1, \ldots, v_n\}$ é LI. Como $\dim_{\mathbb{K}} V = n$, tal conjunto é uma base de V formada por autovetores de T. Por definição, T é diagonalizável. ∎

Exemplo 3.44. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ um operador cuja matriz em relação a alguma base de $\mathbb{R}^3$ é

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

O polinômio característico de T é dado por

$$\begin{aligned} p(x) = \det \begin{pmatrix} x-2 & 1 & 0 \\ 1 & x-2 & 0 \\ 0 & 0 & x-1 \end{pmatrix} &= (x-2)^2(x-1) - (x-1) \\ &= (x-1)((x-2)^2 - 1) \\ &= (x-1)^2(x-3). \end{aligned}$$

Portanto, T possui 2 autovalores distintos, a saber $\lambda_1 = 1$ e $\lambda_2 = 3$. Os autovetores $v = (x, y, z)$ associados a $\lambda_1 = 1$ satisfazem:

$$(I_3 - A)(v) = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow -x + y = 0 \Rightarrow y = x.$$

Portanto

$$V_1 = \{(x, x, z);\ x, z \in \mathbb{R}\},$$

cuja base é o conjunto $\{(1,1,0), (0,0,1)\}$. Os autovetores $v = (x, y, z)$ associados a $\lambda_2 = 3$ satisfazem:

$$(3I_3 - A)(v) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x + y = 0,\ z = 0 \Rightarrow y = -x,\ z = 0.$$

Portanto

$$V_3 = \{(x, -x, 0);\ x \in \mathbb{R}\},$$

cuja base é o conjunto $\{(1, -1, 0)\}$. Assim, o conjunto $\{(1,1,0), (0,0,1), (1,-1,0)\}$ é uma base de $\mathbb{R}^3$ formada por autovetores de T, ou seja, T é diagonalizável.

Exercícios 3.45. Verifique quais dos operadores abaixo são diagonalizáveis.

1. Operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz é dada por

$$[T] = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

2. Operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz é dada por

$$[T] = \begin{pmatrix} 3 & -3 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{pmatrix}.$$

3. Operador linear $T : \mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz é dada por

$$[T] = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & -2 \\ 3 & 2 & 1 \end{pmatrix}.$$

4. Operador linear $T : \mathbb{C}^4 \to \mathbb{C}^4$ cuja matriz é dada por

$$[T] = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Referências Bibliográficas

[1] J.L. Boldrini, S.I.R. Costa, V.L. Figueiredo, H.G. Wetzler, *Álgebra Linear*, 3^a edição, Harbra Ltda, São Paulo, 1986.

[2] C.A. Callioli, H.H. Domingues, R.C.F. Costa, *Álgebra Linear e aplicações*, 6^a edição, Atual, São Paulo, 1993.

[3] F. U. Coelho, M. L. Lourenço, *Um curso de álgebra linear*, 2^a edição, Edusp, São Paulo, 2005.

[4] K. Hoffman, R. Kunze, *Álgebra Linear*, 2^a edição, LTC, Rio de Janeiro, 1979.

[5] S. Lipschutz, M.L. Lipson, *Álgebra Linear*, 4^a edição, Bookman, Porto Alegre, 2011.