Álgebra Linear

Maria Elenice Rodrigues Hernandes

Universidade Estadual de Maringá

SUMÁRIO

INTRODUÇÃO

Este curso introdutório de Álgebra Linear tem como foco o estudo de uma estrutura algébrica chamada espaço vetorial. Essa é uma estrutura muito importante em toda a matemática, em que os estudantes já tiveram um primeiro contato na disciplina de geometria analítica, no estudo dos espaços $\mathbb{R}^2$ e $\mathbb{R}^3$ que são exemplos de espaços vetoriais.

No capítulo 1, iniciamos com o estudo de matrizes com entradas reais ou complexas, posteriormente apresentamos métodos para obter a solução, caso exista, de um sistema de equações lineares. Nesse sentido, o processo de escalonamento de matrizes nos fornece um método efetivo para tal propósito.

No capítulo 2, voltamos nossa atenção para o estudo de espaços vetoriais. Um espaço vetorial, ou mais precisamente um espaço vetorial sobre um conjunto $\mathbb{K}$, nada mais é que um conjunto munido de uma operação de soma entre quaisquer elementos do conjunto e de um produto de um elemento desse conjunto por um escalar de $\mathbb{K}$. Neste curso $\mathbb{K}$ denota o conjunto dos números reais ou o conjunto dos números complexos. Vamos manipular elementos de um espaço vetorial, definir o conceito de base de modo que qualquer elemento de um espaço vetorial possa ser escrito de maneira única em termos de uma base dada. Com isso seremos capazes de formalizar a noção de dimensão de um espaço vetorial.

No capítulo 3, nosso objetivo é o estudo das transformações lineares, que são aplicações entre espaços vetoriais. Nesse contexto, veremos as noções de autovalores, autovetores e o processo de diagonalização de operadores lineares ou correspondentemente de matrizes.

O capítulo final apresenta uma breve revisão de conceitos básicos sobre o conjunto dos números complexos e resultados que nos auxiliam para determinar as raízes de um polinômio de grau maior que 2.

CAPÍTULO 1

MATRIZES E SISTEMAS DE EQUAÇÕES LINEARES

Neste capítulo abordamos um dos mais frutíferos conceitos da Álgebra Linear: matrizes e sistemas lineares. Dentre deste enfoque destacamos o conceito de escalonamento de matrizes, que nos permite determinar se uma matriz é invertível, e quando isto 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. Outros métodos para determinar as possíveis soluções de um sistema linear serão apresentados como, por exemplo, a Regra de Cramer.

1.1 Matrizes

O foco desta seção é o estudo de matrizes. Vamos definir as operações de soma, produto de matrizes e produto de uma matriz por um escalar e apresentar suas propriedades.

Ao longo deste texto, $\mathbb{K}$ denota o conjunto dos números reais $\mathbb{R}$ ou o conjunto dos números complexos $\mathbb{C}$.

Definição 1.1. *Sejam n, m inteiros positivos quaisquer. Uma **matriz** A de ordem n por m ($n \times m$) é uma tabela de elementos $a_{ij} \in \mathbb{K}$, com $1 \leq i \leq n$ e $1 \leq j \leq m$ dispostos em n linhas e* *m colunas, o qual vamos representar por*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Também denotamos uma matriz A por $(a_{ij})_{n\times m}$ ou simplesmente (a_{ij}). Os elementos a_{ij} são denominados as entradas da matriz A.

Vamos denotar por $\mathcal{M}_{n\times m}(\mathbb{K})$ o conjunto de todas as matrizes de ordem $n \times m$ com entradas em $\mathbb{K}$.

Exemplo 1.2. *A matriz*

$$A = \begin{pmatrix} 1 & 0 \\ -5 & \frac{2}{3} \\ 5i-2 & 10 \end{pmatrix}$$

é um elemento de $\mathcal{M}_{3\times 2}(\mathbb{C})$. Apesar de não usual, podemos considerar matrizes de ordem 1×1 como, por exemplo, a matriz (4). Finalmente,

$$\begin{pmatrix} 13 & 3 & -4 & 27 \\ 1 & 1 & 0 & \frac{4}{5} \end{pmatrix} \in \mathcal{M}_{2\times 4}(\mathbb{R}).$$

Definição 1.3. *Duas matrizes $A = (a_{ij})$ e $B = (b_{ij})$ de ordem $n \times m$ são ditas iguais, o qual denotamos $A = B$, se*

$$a_{ij} = b_{ij}$$

para todo $i = 1, \dots, n$ e $j = 1, \dots, m$.

Algumas matrizes recebem nomes especiais por apresentarem boas propriedades. Mencionemos algumas delas.

1. **Matriz Nula**

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

   Por exemplo, as matrizes nulas em $\mathcal{M}_{3\times 1}(\mathbb{K})$ e $\mathcal{M}_{2\times 2}(\mathbb{K})$ são respectivamente

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

2. **Matriz Quadrada**

   Uma matriz $A = (a_{ij})_{n\times m}$ é dita quadrada se $n = m$, ou seja, se o número de linhas e o número de colunas coincide. Neste caso, denominamos por diagonal principal (ou simplesmente diagonal) de A, os elementos da matriz

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

   em que $i = j$, ou seja, os elementos a_{ii}. Neste caso, dizemos simplesmente que A é uma matriz de ordem n.

3. **Matriz Identidade**

   Uma matriz $A = (a_{ij})$ é dita matriz identidade, se ela é uma matriz quadrada em que

$$a_{ij} = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases}$$

   Exemplo 1.4. *A matriz identidade de ordem 4×4 com entradas em $\mathbb{K}$ é dada por*

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

4. **Matriz Diagonal**

   Uma matriz $A = (a_{ij})$ é dita matriz diagonal se ela é uma matriz quadrada tal que $a_{ij} = 0$ para $i \neq j$, ou seja, os elementos que não estão na diagonal são nulos.

   Exemplo 1.5. *As matrizes abaixo são exemplos de matrizes diagonais.*

$$\begin{pmatrix} 7i & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \quad e \quad \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -\frac{16}{250} \end{pmatrix}.$$

As matrizes nula e identidade em $\mathcal{M}_{n\times n}(\mathbb{K})$ são casos particulares de matrizes diagonais.

5. **Matriz Triangular**

   Uma matriz $A = (a_{ij})$ é dita triangular, se ela é uma matriz quadrada tal que todos os elementos abaixo ou acima da diagonal são nulos.

   - *Triangular Superior*: é uma matriz triangular em que $a_{ij} = 0$ para $i > j$, ou seja, os elementos abaixo da diagonal são nulos. Mais precisamente, se A é uma matriz $n \times n$, obtemos

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

   - *Triangular Inferior*: é uma matriz triangular em que $a_{ij} = 0$ para $i < j$, ou seja, os elementos acima da diagonal são nulos, ou ainda,

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

   Exemplo 1.6. *Abaixo apresentamos um exemplo de matriz triangular superior e outra triangular inferior, respectivamente.*

$$\begin{pmatrix} 1 & 3 \\ 0 & 7-2i \end{pmatrix} \quad e \quad \begin{pmatrix} 15 & 0 & 0 \\ 9 & 2 & 0 \\ \frac{1}{4} & 3 & -1 \end{pmatrix}.$$

6. **Matriz Simétrica**

   Uma matriz $A = (a_{ij})$ é dita simétrica, se ela é uma matriz quadrada satisfazendo

$$a_{ij} = a_{ji},$$

   ou seja, há uma simetria entre os elementos abaixo da diagonal com os elementos acima da diagonal.

   Exemplo 1.7. *Este é um exemplo de uma matriz simétrica de ordem* 3*:*

$$\begin{pmatrix} -5i & 9 & 4 \\ 9 & 2 & 0 \\ 4 & 0 & 18 \end{pmatrix}.$$

7. **Matriz Antissimétrica**

   Uma matriz $A = (a_{ij})$ é dita antissimétrica, se ela é uma matriz quadrada de modo que $a_{ij} = -a_{ji}$.

   Observe que neste caso, os elementos da diagonal são nulos, já que $a_{ii} = -a_{ii}$ se, e somente se, $a_{ii} = 0$.

   Exemplo 1.8. *Esta é uma matriz antissimétrica:*

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

   No entanto, a matriz

$$\begin{pmatrix} 15 & -1 & 20 \\ 1 & \frac{1}{2} & 3i \\ -20 & -3i & -7 \end{pmatrix}$$

   não *é antissimétrica.*

1.2 Operações com Matrizes

Nesta seção vamos definir algumas operações com matrizes, como a soma de matrizes, o produto de um escalar por uma matriz e o produto de matrizes.

Primeiramente, dizemos que duas matrizes $A = (a_{ij})$ e $B = (b_{ij})$ de mesma ordem $n \times m$ são iguais, o qual denotamos $A = B$, se $a_{ij} = b_{ij}$ para todo $1 \leq i \leq n$ e $1 \leq j \leq m$.

- Soma de Matrizes

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

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

  Exemplo 1.9. *Veja que*

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

A operação de soma definida anteriormente satisfaz boas propriedades.

Proposição 1.10. *Sejam $A, B, C \in \mathcal{M}_{n\times m}(\mathbb{K})$. Valem as seguintes propriedades:*

(i) $A + (B + C) = (A + B) + C$; *(Associatividade)*

(ii) $A + B = B + A$; *(Comutatividade)*

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

(iv) $A + (-A) = 0$, *em que se* $A = (a_{ij})$, *então* $-A$ *denota a matriz* $(-a_{ij})$, *chamada matriz oposta de* A. *(Elemento oposto)*

Demonstração: (i) Considere $A = (a_{ij}), B = (b_{ij})$ e $C = (c_{ij})$ em $\mathcal{M}_{n\times m}(\mathbb{K})$. Denotando $B + C = (d_{ij})$, sabemos que $d_{ij} = b_{ij} + c_{ij}$. Assim

$$A + (B + C) = (a_{ij} + d_{ij}) = (a_{ij} + (b_{ij} + c_{ij})) = ((a_{ij} + b_{ij}) + c_{ij}) = (A + B) + C,$$

em que a penúltima igualdade segue da propriedade associativa de $\mathbb{K}$.

As demais propriedades decorrem analogamente e serão deixadas como exercício. ■

Uma vez definido a operação de adição em $\mathcal{M}_{n\times m}(\mathbb{K})$, naturalmente podemos definir a operação de *subtração* da maneira usual: dadas as matrizes $A, B \in \mathcal{M}_{n\times m}(\mathbb{K})$, a subtração de A por B, denotada $A - B$, é a matriz definida por

$$A - B = A + (-B).$$

- Multiplicação de um Escalar por uma Matriz

Sejam $A = (a_{ij})_{n\times m}$ e $k \in \mathbb{K}$ um escalar. Definimos o produto do escalar $k \in \mathbb{K}$ pela matriz A, denotado por kA, como sendo a matriz

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

Exemplo 1.11. *Neste exemplo temos o escalar* i *multiplicado por uma matriz* 2×5*:*

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

O produto de um escalar por uma matriz também satisfaz algumas propriedades:

Proposição 1.12. *Dados* $A, B \in \mathcal{M}_{n\times m}(\mathbb{K})$ *e* $k, k_1, k_2 \in \mathbb{K}$, *temos que*

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

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

(iii) $k_1(k_2A) = (k_1k_2)A$; *(Associatividade)*

(iv) $0A = 0_{n\times m}$, *onde* $0_{n\times m}$ *denota a matriz nula de ordem* $n \times m$;

(vi) $1A = A$.

Demonstração: (i) Sejam $A = (a_{ij})$ e $B = (b_{ij})$ e denotemos por $A + B = (c_{ij})$, em que $c_{ij} = a_{ij} + b_{ij}$. Dado $k \in \mathbb{K}$ temos que

$$k(A+B) = (kc_{ij}) = (k(a_{ij} + b_{ij})) = (ka_{ij} + kb_{ij}) = kA + kB,$$

em que a penúltima igualdade segue da propriedade distributiva em $\mathbb{K}$.

As demais propriedades serão deixadas como exercício. ∎

- Transposta de uma Matriz

Dada uma matriz $A = (a_{ij})_{n\times m}$ denomina-se a *transposta* de A, o qual denotamos por A^t, a matriz cujas linhas são as colunas de A, isto é, se $A^t = (b_{ij})_{m\times n}$ então $b_{ij} = a_{ji}$.

Exemplo 1.13. *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. *Dadas as matrizes* A *e* B *de ordem* $n \times m$ *e* $k \in \mathbb{K}$, *temos que*

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

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

(iii) $(A^t)^t = A$.

Demonstração: Sejam as matrizes $A = (a_{ij})$ e $B = (b_{ij})$.

(i) Denotemos por $A + B = (c_{ij})$ em que $c_{ij} = a_{ij} + b_{ij}$. Assim,

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

(ii) Dado $k \in \mathbb{K}$ temos que $kA = (ka_{ij})$. Assim $(kA)^t = (ka_{ji})$. Por outro lado,

$$kA^t = (ka_{ji}).$$

Portanto, $(kA)^t = kA^t$.

(iii) Sabemos que $A^t = (a_{ji})$ e assim $(A^t)^t = (a_{ij}) = A$. ■

- Produto de Matrizes

A operação produto de matrizes é algo mais complicado de definir, pois gostaríamos que o produto definido satisfizesse algumas propriedades naturais, como a associatividade, distributividade com respeito à soma, entre outras.

Sejam $A = (a_{ij})_{n\times m}$ e $B = (b_{jk})_{m\times p}$ matrizes em que o 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 $(c_{ik})_{n\times p}$ tal que

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

Exemplo 1.15. *Vamos calcular o produto das matrizes*

$$A = \begin{pmatrix} 1 & -1 & i \\ 2 & 1 & 0 \end{pmatrix} \quad e \quad B = \begin{pmatrix} i & 1 & -1 & \sqrt{3} \\ 0 & x & i & 1 \\ 0 & 0 & 1 & 8 \end{pmatrix}.$$

Primeiramente, observe que o número de colunas de A coincide com o número de linhas de B. Denotando $AB = (c_{ik})$ como anteriormente, temos que

$$\begin{aligned} c_{11} &= \sum_{j=1}^{3} a_{1j}b_{j1} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} \\ &= 1 \cdot i + (-1) \cdot 0 + i \cdot 0 = i. \end{aligned}$$

Seguindo de modo análogo, obtemos

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

Observação 1.16. *O produto de matrizes* ***não é comutativo****, ou seja, em geral*

$$AB \neq BA.$$

Na verdade, um dos membros deste produto pode nem estar definido.

Definição 1.17. *Dizemos que duas matrizes A e B comutam se $AB = BA$.*

Exemplo 1.18. *Sejam $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Essas matrizes comutam, pois*

$$AB = \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = BA.$$

Proposição 1.19. *Sejam as matrizes A, B e C com as ordens adequadas para que cada produto abaixo esteja definido:*

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

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

(*iii*) *$(A+B)C = AC + BC$; (Distributividade)*

(*iv*) *$(AB)C = A(BC)$; (Associatividade)*

(*v*) *$(AB)^t = B^tA^t$;*

(*vi*) *$0A = 0$ e $A0 = 0$, em que 0 denota a matriz nula.*

Demonstração: (*i*) Sejam $A = (a_{ij})_{n\times m}$ e $I = (b_{jk})_{m\times m}$. Denotando $AI = (c_{ik})_{n\times m}$, por definição, $c_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk}$. Entretanto a matriz identidade satisfaz que

$$b_{jk} = \begin{cases} 1, & \text{se } j = k, \\ 0, & \text{se } j \neq k. \end{cases}$$

Logo,

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

Portanto, $AI = (c_{ik}) = (a_{ik}) = (a_{ij}) = A$.

(ii) Sejam $A = (a_{ij})_{n\times m}$, $B = (b_{jk})_{m\times p}$ e $C = (c_{jk})_{m\times p}$. Denotemos $A(B+C) = (d_{ik})_{n\times p}$. Assim, segue da propriedade distributiva em $\mathbb{K}$ que

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

Portanto, $A(B + C) = AB + AC$.

(v) Se $AB = (c_{ik})_{n\times k}$ com $c_{ik} = \sum_{j=1}^{m} a_{ij}b_{jk}$, então $(AB)^t = (c_{ki})_{k\times n}$. Logo,

$$c_{ki} = \sum_{j=1}^{m} a_{ji}b_{kj} = \sum_{j=1}^{m} b_{kj}a_{ji},$$

em que na última igualdade utilizamos a propriedade comutativa em $\mathbb{K}$.

Por outro lado, $B^t = (b_{kj})$ e $A^t = (a_{ji})$. Assim $B^tA^t = (d_{ki})$ em que

$$d_{ki} = \sum_{j=1}^{m} b_{kj}a_{ji} = c_{ki}.$$

Donde segue que $(AB)^t = B^tA^t$.

As demais propriedades serão deixadas como exercício. ■

Definição 1.20. *Uma matriz quadrada $A = (a_{ij})_{n\times n}$ é dita **invertível** se existir uma matriz quadrada B de mesma ordem tal que $AB = BA = I$, em que I denota a matriz identidade de ordem $n \times n$.*

A matriz B é dita a inversa da matriz A, o qual vamos denotar $B = A^{-1}$. Observe que a definição acima garante que existe uma inversa à direita e à esquerda. Na verdade, a proposição abaixo nos garante que se uma matriz possui uma inversa à direita B, então B também é uma inversa à esquerda, ou seja, se A é uma matriz invertível, então existe uma única matriz $B = A^{-1}$ tal que $AB = BA = I$.

Proposição 1.21. *Dado uma matriz quadrada A de ordem $n \times n$, se existirem matrizes B e C de ordem $n \times n$ tais que $AB = I$ e $CA = I$, então $B = C$.*

Demonstração: Pelos itens (i) e (iv) da Proposição 1.19 temos que

$$C = CI = C(AB) = (CA)B = IB = B.$$

Portanto, a inversa de uma matriz A, se existir, é única. ∎

Em geral, o cálculo da inversa de uma matriz é um processo trabalhoso. Mais adiante vamos apresentar alguns métodos para calcular a inversa de uma matriz.

Exemplo 1.22. *Verifiquemos que a inversa da matriz* $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$ *é a matriz* $\begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix}$.

De fato,

$$\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix} = \begin{pmatrix} \frac{3}{5} + \frac{2}{5} & \frac{1}{5} - \frac{1}{5} \\ \frac{6}{5} - \frac{6}{5} & \frac{2}{5} + \frac{3}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Proposição 1.23. *Sejam A e B matrizes invertíveis de ordem $n \times n$. Então,*

(i) A^{-1} *é invertível e* $(A^{-1})^{-1} = A$;

(ii) AB *é invertível e* $(AB)^{-1} = B^{-1}A^{-1}$;

(iii) A^t *é invertível e* $(A^t)^{-1} = (A^{-1})^t$.

Demonstração: (i) Seja $B = A^{-1}$. Devemos provar que B é invertível. Mais ainda, queremos provar que $B^{-1} = A$, ou seja, que $BA = I$. De fato,

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

(ii) Provemos que $AB(B^{-1}A^{-1}) = I$. Note que,

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

(iii) Pela propriedade (v) do produto de matrizes e pela proposição acima temos que:

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

∎

Exemplo 1.24. *Considere a matriz A dada no Exemplo 1.22. Veja que utilizando as propriedades dadas na proposição anterior, podemos facilmente calcular $(A^t)^{-1}$ em que*

$$A^t = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$$

Como no Exemplo 1.22 temos a inversa da matriz A, segue do item (iii) da proposição anterior que

$$(A^t)^{-1} = (A^{-1})^t = \begin{pmatrix} \frac{3}{5} & \frac{1}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{pmatrix}^t = \begin{pmatrix} \frac{3}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

Dado A uma matriz quadrada de ordem $n \times n$, denotamos por A^k o produto da matriz A por ela mesmo k vezes, isto é,

$$A^k = \underbrace{A \cdot A \cdot \ldots \cdot A}_{k \text{ vezes}}.$$

Analogamente, denotamos por

$$A^{-k} = \underbrace{A^{-1} \cdot A^{-1} \cdot \ldots \cdot A^{-1}}_{k \text{ vezes}}.$$

Definição 1.25. *Dada uma matriz quadrada A de ordem n associamos um número chamado traço de A, denotada por $tr(A)$, o qual é definida como*

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

Exemplo 1.26. *Vamos calcular o traço da matriz*

$$B = \begin{pmatrix} \frac{7}{4} & -25 & 56i+7 \\ 0 & -3 & 468 \\ 11i & \sqrt{2} & 1-2i \end{pmatrix}.$$

Neste caso, $tr(B) = \frac{7}{4} + (-3) + (1-2i) = \frac{7}{4} - 2 - 2i = -\frac{1}{4} + 2i.$

1.3 Escalonamento de Matrizes

O processo de escalonamento de matrizes é um dos mais úteis no estudo de matrizes e na solução de sistemas lineares. Comentamos anteriormente que determinar a inversa de uma matriz pode ser uma tarefa árdua. Vamos verificar que o processo de escalonamento nos dá um método sistemático (mas não menos trabalhoso) para determinar a inversa de uma matriz. Este processo também nos auxilia no cálculo das possíveis soluções de um sistema linear. A cada passo do processo de escalonamento, obtemos um novo sistema mais simples que o anterior e o mais surpreendente: com a propriedade de possuir o mesmo conjunto solução que o sistema inicial.

Vamos iniciar definindo determinadas operações sobre as linhas L_i de uma matriz A, chamadas *operações elementares*:

1. Permutar a linha L_i pela linha L_j $(L_i \leftrightarrow L_j)$.
2. Multiplicar a linha L_i por um escalar $k \in \mathbb{K}^*$ $(L_i \to kL_i)$.
3. Substituir a linha L_i pela adição desta mesma linha com k vezes uma linha L_j, onde $k \in \mathbb{K}^*$ e $i \neq j$ $(L_i \to L_i + kL_j)$.

Exemplo 1.27. *Vamos aplicar as seguintes operações elementares nas linhas da seguinte matriz:*

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

Definição 1.28. *Sejam A e B matrizes de ordem $n \times m$. Dizemos que B é* ***equivalente por linhas*** *à matriz A, que denotamos $A \sim B$, se B for obtida de A através de um número finito de operações elementares sobre as linhas de A.*

No exemplo anterior, todas as matrizes são equivalentes por linhas.

Vamos verificar abaixo que as operações elementares nas linhas de uma matriz são reversíveis, ou seja, se A é equivalente por linhas à uma matriz B, então B é equivalente por linhas à matriz A. De fato, se e representa uma das operações elementares nas linhas de uma matriz A de ordem $n \times m$, denotemos por $e(A)$ a matriz obtida de A, aplicando-lhe a operação elementar e. Temos o seguinte resultado:

Proposição 1.29. *Toda operação elementar e nas linhas de uma matriz em $\mathcal{M}_{n\times m}(\mathbb{K})$ é reversível, ou seja, existe uma operação elementar e' tal que $e'(e(A)) = A$ e $e(e'(A)) = A$, para todo $A \in \mathcal{M}_{n\times m}(\mathbb{K})$.*

Demonstração: Para todo $A \in \mathcal{M}_{n\times m}(\mathbb{K})$, se e é uma operação elementar do tipo $L_i \leftrightarrow L_j$, então tome $e' = e$, pois neste caso temos que $e(A)$ troca a linha L_i pela L_j e $e'(e(A))$ troca a linha L_j pela linha L_i, obtendo novamente a matriz A. Se e for a operação elementar $L_i \to kL_i$ com $k \in \mathbb{K}^*$, tome e' a operação elementar $L_i \to \frac{1}{k}L_i$. Assim, a i-ésima linha de $e(A)$ será kL_i, o qual operando e' obtemos i-ésima linha $\frac{1}{k}kL_i = L_i$, ou seja, obtemos novamente a matriz A.

E finalmente, se e for a operação elementar $L_i \to L_i + kL_j$ com $k \in \mathbb{K}^*$, tome e' a operação elementar $L_i \to L_i - kL_j$. Deste modo, $e(A)$ possui a i-ésima linha $L_i + kL_j$ e $e'(e(A))$ tem a i-ésima linha da forma $L_i + kL_j - (kL_j) = L_i$.

Analogamente se operarmos $e(e'(A))$. ∎

É fácil ver que e' é a única operação elementar com a propriedade que

$$e'(e(A)) = A$$

para toda matriz $A \in \mathcal{M}_{n\times m}(\mathbb{K})$.

Também não é difícil verificar que a relação $\sim$ de equivalência por linhas é uma relação de equivalência em $\mathcal{M}_{n\times m}(\mathbb{K})$, ou seja, dadas quaisquer matrizes $A, B, C \in \mathcal{M}_{n\times m}(\mathbb{K})$ temos que

(i) $A \sim A$ (reflexiva);

(ii) se $A \sim B$ então $B \sim A$ (simétrica);

(iii) se $A \sim B$ e $B \sim C$ então $A \sim C$ (transitiva).

Definição 1.30. *Dizemos que uma matriz A está na* ***forma escada*** *se:*

(a) o primeiro elemento não nulo de uma linha não nula é 1;

(b) cada coluna que contém o primeiro elemento não nulo de alguma linha, possui todos os seus outros elementos iguais a zero;

(c) toda linha nula ocorre abaixo de todas as linhas não nulas;

(d) se as linhas $L_1, \ldots L_r$ são as linhas não nulas e se o primeiro elemento não nulo da linha L_i ocorre na coluna k_i, então $k_1 < k_2 < \cdots < k_r$, ou seja, o número de zeros precedendo o primeiro elemento não nulo de uma linha aumenta a cada linha, até que sobrem apenas as linhas nulas, se houverem.

Dizemos que uma matriz está na **forma escalonada** se satisfaz as condições (c) e (d) acima e trocamos (b) por:

(b') cada coluna que contém o primeiro elemento não nulo de alguma linha possui todos os elementos abaixo dele nulos.

A forma escada de uma matriz é um caso particular de uma matriz na forma escalonada.

Exemplo 1.31. *1. A matriz abaixo não está na forma escada, pois falha o item (b), mas está na forma escalonada:*

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

2. A matriz abaixo não está na forma escalonada, pois falha os itens (a) e (d):

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

3. Sejam as matrizes

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

A matriz A está na forma escada e a matriz B na forma escalonada.

Exemplo 1.32. *Vamos obter a forma escada da matriz abaixo por meio de um número finito de operações elementares em suas linhas.*

$$\begin{pmatrix} 0 & -1 & 0 & -2 \\ 4 & 0 & 0 & 1 \\ -6 & 2 & 0 & 1 \end{pmatrix} \underset{L_2 \leftrightarrow L_1}{\rightarrow} \begin{pmatrix} 4 & 0 & 0 & 1 \\ 0 & -1 & 0 & -2 \\ -6 & 2 & 0 & 1 \end{pmatrix} \underset{L_1 \rightarrow \frac{1}{4}L_1}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & -1 & 0 & -2 \\ -6 & 2 & 0 & 1 \end{pmatrix}$$

$$\underset{L_3 \rightarrow L_3 + 6L_1}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & -1 & 0 & -2 \\ 0 & 2 & 0 & \frac{5}{2} \end{pmatrix} \underset{L_2 \rightarrow (-1)L_2}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 0 & \frac{5}{2} \end{pmatrix} \underset{L_3 \rightarrow L_3 - 2L_2}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}$$

$$\underset{L_3 \rightarrow -\frac{2}{3}L_3}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{L_1 \rightarrow L_1 - \frac{1}{4}L_3}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{L_2 \rightarrow L_2 - 2L_3}{\rightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Uma prova da unicidade do próximo teorema será apresentada posteriormente.

Teorema 1.33. *Toda matriz A é equivalente por linhas a uma única matriz na forma escada.*

Demonstração: Suponha $A = (a_{ij})_{n \times m}$ uma matriz não nula. Então, existe um menor inteiro positivo k_1 tal que $a_{ik_1} \neq 0$ para alguma linha i, com $1 \leq i \leq n$. Escolha uma tal linha i e troque (se necessário) de modo que esta seja a primeira linha. Considere as seguintes operações elementares

$$L_1 \to \frac{1}{a_{1k_1}} L_1, \qquad \text{já que } a_{1k_1} \neq 0$$

$$L_j \to L_j - a_{jk_1} L_1 \quad \forall\ 2 \leq j \leq n.$$

Com isso, obtivemos uma matriz B cujo primeiro elemento não nulo da primeira linha é 1 e todos os demais elementos desta coluna são iguais a zero. Se as linhas L_i da matriz $B = (b_{ij})$ com $2 \leq i \leq n$ não são todas nulas, então existe um menor inteiro positivo k_2 tal que $b_{ik_2} \neq 0$ para alguma linha L_i com $2 \leq i \leq n$. Escolha uma tal linha e troque (se necessário) para que esta seja a segunda linha. Considere as seguintes operações:

$$L_2 \to \frac{1}{b_{2k_2}} L_2, \qquad \text{já que } b_{2k_2} \neq 0$$

$$L_j \to L_j - b_{jk_2} L_2 \quad \forall\ 1 \leq j \leq n,\ j \neq 2.$$

Observe que $k_1 < k_2$. Com isso, temos agora uma nova matriz cujas colunas anteriores a k_2 permaneceram como na matriz B, o primeiro elemento não nulo da segunda linha é 1 e todos os demais elementos desta coluna (das linhas L_1 e L_i com $3 \leq i \leq n$) são nulos. Continuando este processo e permutando, se necessário, obteremos uma matriz M com $k_1 < k_2 < \cdots < k_r$ e que satisfaz o item (c). Com isso M está na forma escada. ∎

Seguindo os passos da demonstração do teorema acima, obtemos de modo análogo que qualquer matriz A é equivalente por linhas a uma matriz na forma escalonada, todavia ela não é única.

Definição 1.34. *O posto de uma matriz A de ordem $n \times m$, denotado por p_A, é o número de linhas não nulas de uma matriz B, que é uma redução de A à forma escalonada (ou escada). A nulidade de A é o número $m - p_A$. Vamos denotar a nulidade de A por N_A.*

Observe que aplicar uma operação elementar nas linhas de uma matriz A de ordem $n \times m$, é o mesmo que aplicar esta operação elementar na matriz identidade $n \times n$ e em seguida, multiplicar esta nova matriz por A.

Exemplo 1.35. *1. Realizando a operação $L_2 \to L_2 - L_1$ seguida de $L_3 \to L_3 - 3L_1$ nas linhas da matriz A obtemos a matriz B, ou seja,*

$$A = \begin{pmatrix} -1 & 4 & 2 \\ 0 & 1 & -3 \\ -1 & 1 & 1 \end{pmatrix} \underset{L_2 \to L_2 - L_1}{\longrightarrow} \begin{pmatrix} -1 & 4 & 2 \\ 1 & -3 & -5 \\ -1 & 1 & 1 \end{pmatrix} \underset{L_3 \to L_3 - 3L_1}{\longrightarrow} \begin{pmatrix} -1 & 4 & 2 \\ 1 & -3 & -5 \\ 2 & -11 & -5 \end{pmatrix} = B.$$

Realizando essas mesmas operações na matriz identidade obtemos a matriz E, ou seja,

$$I=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\underset{L_2\to L_2-L_1}{\to}\begin{pmatrix}1&0&0\\-1&1&0\\0&0&1\end{pmatrix}\underset{L_3\to L_3-3L_1}{\to}\begin{pmatrix}1&0&0\\-1&1&0\\-3&0&1\end{pmatrix}=E.$$

Observe que fazendo o produto de E por A obtemos B:

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

2. *Novamente, aplicando a operação $L_2\to L_2+L_1$ nas linhas da matriz A obtemos*

$$A=\begin{pmatrix}-1&1&3\\4&2&0\end{pmatrix}\underset{L_2\to L_2+L_1}{\to}\begin{pmatrix}-1&1&3\\3&3&3\end{pmatrix}=B.$$

Aplicando a mesma operação na matriz identidade 2×2 obtemos a matriz E, o qual multiplicando por A obtemos

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

Definição 1.36. *Uma **matriz elementar** E de ordem n é uma matriz quadrada de mesma ordem, obtida da matriz identidade I de ordem n, a partir da aplicação de uma operação elementar nas linhas de I, ou seja, $E=e(I)$, onde e é uma operação elementar.*

O próximo resultado nos diz que aplicar uma operação elementar nas linhas de uma matriz A é o mesmo que aplicar tal operação elementar nas linhas da matriz identidade I e posteriormente multiplicar a matriz obtida pela matriz A.

Proposição 1.37. *Seja e uma operação elementar nas linhas de matrizes em $\mathcal{M}_{n\times m}(\mathbb{K})$. Considere a matriz elementar $E=e(I)$, em que I é a matriz identidade de ordem n. Então, para toda matriz $A\in\mathcal{M}_{n\times m}(\mathbb{K})$ temos que*

$$e(A)=EA.$$

Demonstração: Como $IA=A$ basta efetuar uma das operações elementares em I, obtendo a matriz E e efetuar o produto $E\cdot A$ e observar que coincide com $e(A)$. ■

Corolário 1.38. *Sejam $A, B \in \mathcal{M}_{n\times m}(\mathbb{K})$. Então, A é equivalente por linhas a B se, e somente se, existem matrizes elementares $E_1, \ldots, E_s$ de ordem n tais que*

$$E_s \cdots E_2 E_1 A = B.$$

Demonstração: Por definição, A é equivalente por linhas a B se, e somente se, existem operações elementares $e_1, \ldots, e_s$ tais que

$$e_s(\cdots(e_2(e_1(A)))\cdots) = B.$$

Denotanto por $E_i = e_i(I)$ para cada $1 \leq i \leq s$, onde I é a matriz identidade de ordem n, segue do resultado anterior que $e_i(A) = e_i(I)A = E_i A$ assim

$$\begin{aligned} B = e_s(\cdots(e_2(e_1(A)))\cdots) &= e_s(\cdots(e_2(E_1A))\cdots) = e_s(\cdots(E_2E_1A))\cdots) = \ldots \\ &= E_s \cdots E_2 E_1 A. \end{aligned}$$

■

Corolário 1.39. *Uma matriz elementar E é invertível e sua inversa é a matriz elementar E', que corresponde à operação inversa da operação efetuada nas linhas de E, ou seja, $E' = E^{-1}$.*

Demonstração: Seja $E = e(I)$ uma matriz elementar e e sua correspondente operação elementar. Sabemos que as operações elementares são reversíveis, assim se e' é a operação elementar inversa de e, denotemos por $E' = e'(I)$.

Pela Proposição 1.37 temos

$$\begin{aligned} & I = e'(e(I)) = e'(E) = e'(I)E = E'E \\ \text{e} \quad & I = e(e'(I)) = e(E') = e(I)E' = EE'. \end{aligned}$$

Logo, E é inversível e $E^{-1} = E'$. ■

$$\begin{array}{cc|ccc} I & I & \text{operação elementar} & & \text{oper. elementar inversa} \\ e\downarrow & \downarrow e' & L_i \leftrightarrow L_j & \longrightarrow & L_j \leftrightarrow L_i \\ E & E' & L_i \to kL_i & \longrightarrow & L_i \to \frac{1}{k}L_i \\ & & L_i \to L_i + kL_j & \longrightarrow & L_i \to L_i - kL_j \end{array}$$

Exemplo 1.40. *Considere a matriz elementar E de ordem 4 obtida da matriz identidade aplicando a seguinte operação elementar:*

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{L_4 \to L_4 - 5L_2}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -5 & 0 & 1 \end{pmatrix} = E.$$

Agora, vamos considerar a operação elementar inversa da considerada acima, obtendo a matriz E', ou seja,

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underset{L_4 \to L_4 + 5L_2}{\longrightarrow} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{pmatrix} = E'.$$

Pelo resultado anterior, E' é a inversa da matriz E, como podemos comprovar abaixo:

$$E \cdot E' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -5 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1.4 Sistemas de Equações Lineares

Um sistema de equações lineares com m equações e n variáveis é um conjunto de equações do tipo

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

com $a_{ij}, b_i \in \mathbb{K}$ com $\mathbb{K}$ um corpo qualquer, $1 \leq i \leq m$ e $1 \leq j \leq n$. Denominamos por x_j as variáveis ou incógnitas do sistema, a_{ij} os coeficientes das variáveis e b_i de termos independentes.

Definição 1.41. *Uma **solução** do sistema (1.1) é uma n-upla de números $(x_1, x_2, \ldots, x_n)$ com $x_i \in \mathbb{K}$ satisfazendo simultâneamente as m equações de (1.1).*

Exemplo 1.42. *Uma classe de sistemas de equações lineares, bem conhecidos da geometria analítica, são os sistemas em $\mathbb{R}^2$ e $\mathbb{R}^3$ (espaços que serão definidos em detalhes posteriormente). Nestes casos, as soluções algébricas dos sistemas, podem ser visualizadas geometricamente.*

Em $\mathbb{R}^2$, cada equação de um sistema de equações lineares dado por

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2, \end{cases} \tag{1.2}$$

com $a_{ij}, b_j \in \mathbb{R}$ e $i, j = 1, 2$, descreve uma reta em $\mathbb{R}^2$, cujas soluções para tal sistema equivale a descobrir as posições relativas entre as duas retas. Mais precisamente, se as retas são concorrentes, algebricamente significa que o sistema acima possui uma única solução. Se as retas são paralelas, concluímos que o sistema não possui solução e finalmente se as retas são coincidentes, então o sistema possui infinitas soluções.

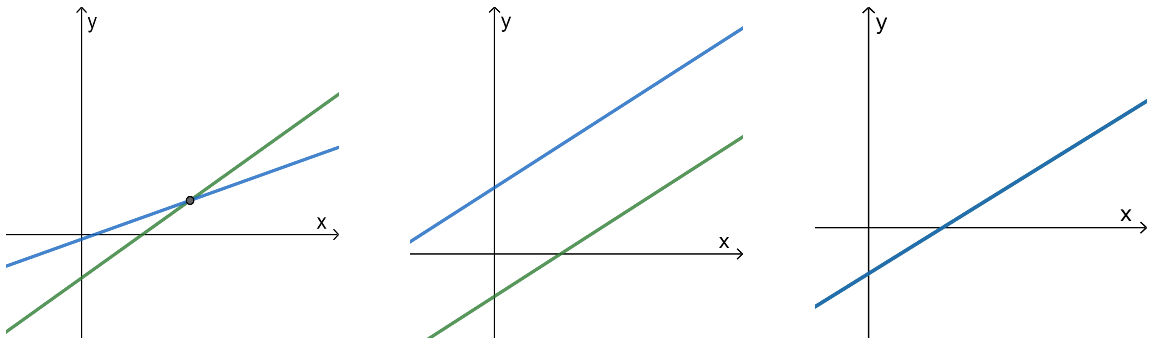

Figura 1.1: Representação geométrica das possíveis soluções do sistema (1.2)

De maneira análoga, em $\mathbb{R}^3$ cada equação de um sistema de equações lineares dado por

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3, \end{cases} \tag{1.3}$$

com $a_{ij}, b_j \in \mathbb{R}$ e $i, j = 1, 2, 3$, descreve um plano.

Em geometria analítica vimos que as posições relativas entre dois planos quaisquer são: os planos podem ser transversais, ou seja, se intersectam em uma reta, ou ainda podem ser paralelos ou coincidentes. Assim, dado um sistema com três equações e três variáveis como acima, algumas possíveis posições relativas entre estes três planos são: Os três planos podem se intersectar em um ponto e, portanto, teríamos uma única solução para o sistema. Os três planos podem se intersectar em uma reta, ou serem coincidentes, que algebricamente significa que o sistema possui infinitas soluções.

Podemos ainda ter três planos paralelos, ou dois planos que se intersectam em uma reta, mas o terceiro plano não intersecta esta reta, ou seja, em ambos os casos o sistema não teria solução.

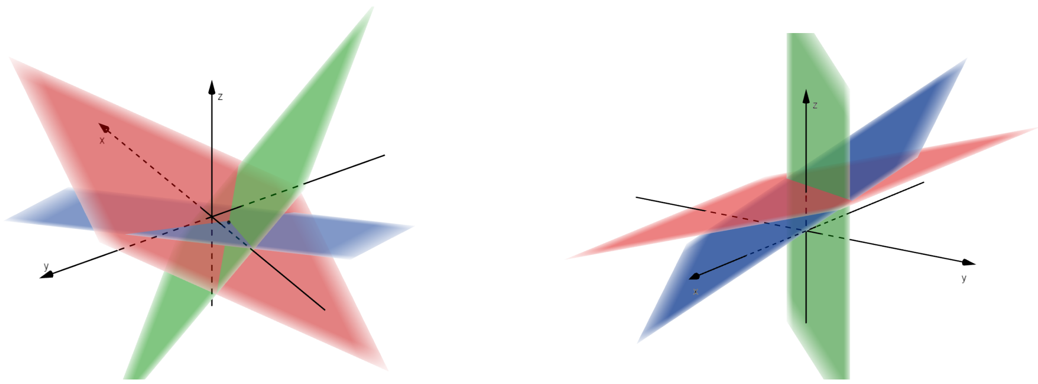

Figura 1.2: Três planos que se intersectam em um ponto e uma reta

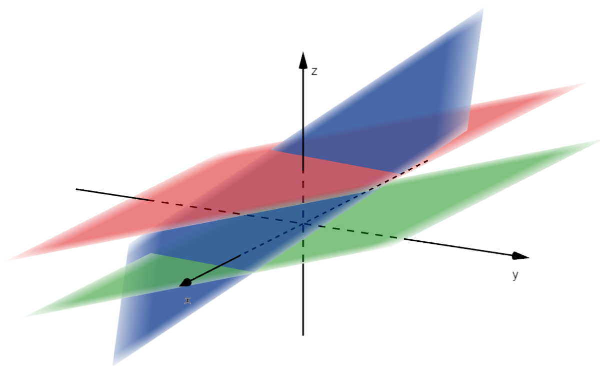

Figura 1.3: Dois planos paralelos e um plano transversal a esses planos

Mas fica a pergunta se não teríamos outras possibilidades para as soluções de tal sistema.

Claramente podemos analisar, em cada um dos casos descritos acima, as soluções dos sistemas se tivéssemos mais equações que as consideradas.

Nosso propósito nesta seção é apresentar um método que nos permita determinar se um sistema de equações lineares possui ou não solução e, caso o tenha, determinar quais são essas soluções.

Podemos associar a um sistema da forma (1.1) uma equação matricial dada por:

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

ou seja, se denotamos por $C = (a_{ij})_{m\times n}$ a matriz dos coeficientes, por $X = (x_j)_{n\times 1}$ a matriz das variáveis e por $B = (b_i)_{m\times 1}$ a matriz dos termos independentes, com $i = 1, \ldots, m$ e $j = 1, \ldots, n$, então o sistema (1.1) pode ser reescrito na forma matricial pela relação

$$C \cdot X = B. \tag{1.4}$$

Desse modo, uma matriz X_0 é solução do sistema (1.4) se, e somente se, $C \cdot X_0 = B$.

Vamos definir ainda uma outra matriz que será importante para nossos propósitos. Denominamos a matriz

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

por **matriz ampliada** do sistema (1.1).

O teorema a seguir nos apresenta um método para determinar as soluções de um sistema linear, por meio de outro sistema linear obtido a partir do primeiro, cuja busca de soluções (caso exista) se torna mais simples. Mais ainda, o sistema inicial e o sistema mais simples possuem o mesmo conjunto solução. Veremos que encontrar as soluções de um sistema linear recai em um problema matricial, no qual o processo de escalonamento de matrizes nos oferece um método sistemático para a obtenção de tais soluções.

Teorema 1.43. *Dois sistemas de equações lineares com m equações e n variáveis, cujas matrizes ampliadas são equivalentes por linhas, possuem o mesmo conjunto solução.*

Demonstração: Considere os sistemas de equações lineares I e II com m equações e n variáveis dados por

$$I : C \cdot X = B \quad \text{e} \quad II : C' \cdot X = B'$$

em que C, C' são as matrizes dos coeficientes e B, B' são as matrizes dos termos independentes de I e II, respectivamente. Sejam A e A' suas respectivas matrizes ampliadas que, por hipótese, são equivalentes por linhas. Pelo Corolário 1.38, $A' = M \cdot A$, onde M é um produto de matrizes elementares. Como matrizes elementares são invertíveis (Corolário 1.39) e pelo fato de que se duas matrizes são invertíveis, seu produto também o é, concluímos que M é invertível. Assim $A = M^{-1}A'$.

Vamos denotar as matrizes ampliadas da forma $A = (C\ B)$ e $A' = (C'\ B')$. Assim

$$A = (C\ B) \ \Leftrightarrow\ MA = (MC\ MB) \ \Leftrightarrow\ A' = (MC\ MB) \ \Leftrightarrow\ (C'\ B') = (MC\ MB).$$

Logo, $C' = M \cdot C$ e $B' = M \cdot B$. Portanto,

$$C \cdot X = B \;\Leftrightarrow\; MCX = MB \;\Leftrightarrow\; C'X = B'.$$

Donde concluímos que, uma matriz X é solução do sistema I se, e somente se, for solução do sistema II. ■

No próximo teorema veremos que um sistema de equações lineares pode admitir uma única solução, infinitas soluções ou nenhuma solução. Esse método de resolução de sistemas lineares é chamado de método do escalonamento, ou método de eliminação de Gauss (ou Gaussiana).

Teorema 1.44. *Considere um sistema de equações lineares com m equações e n variáveis. Sejam A a matriz ampliada do sistema e C sua matriz dos coeficientes.*

(i) o sistema admite solução se, e somente se, $p_A = p_C$, ou seja, o posto das matrizes A e C coincidem;

(ii) se $p_A = p_C = n$, então o sistema possui uma única solução;

(iii) se $p_A = p_C = r < n$, então o sistema possui infinitas soluções. Neste caso, podemos escolher $n - r$ variáveis livres e as demais r variáveis serão expressas em função dessas.

Demonstração: Vamos provar a ida de (i) por contra-positiva.

Suponha que $p_A \neq p_C$. Neste caso, $p_A > p_C$ (pois menor não pode ser), o que implica que a matriz A reduzida à forma escada A' deve ter pelo menos uma linha do tipo

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

com $c_k \neq 0$. Logo, o sistema associado à matriz A', que pelo Teorema 1.43 possui o mesmo conjunto solução que o sistema inicial, tem uma equação do tipo,

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

e, portanto, o sistema não admite solução.

Agora suponha que $p_A = p_C$ e consideremos os seguintes casos:

$1^{o.}$ *caso:* $p_A = p_C = n$.

Neste caso, significa que reduzindo a matriz ampliada do sistema à forma escada A' teremos:

$$A' = \begin{pmatrix} 1 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & \cdots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_n \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}_{m\times(n+1)}.$$

Portanto, a solução do sistema associado à matriz A' será $x_1 = c_1,\ x_2 = c_2, \ldots,\ x_n = c_n$, ou seja, o sistema inicial também terá solução única.

$2^{o.}$ *caso:* $p_A = p_C = r < n$.

Neste caso, temos várias possibilidades para a matriz A' que é a redução de A à forma escada. A menos de uma permutação das variáveis podemos supor primeiramente que a matriz A' é dada por:

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

cujo sistema associado possui n variáveis com $r < n$ equações e satisfaz:

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

ou seja, o sistema terá infinitas soluções com $x_{r+1}, \ldots, x_n$ variáveis livres, mais especificamente $n - r$.

Todavia a matriz ampliada poderia 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 possui n variáveis com $r < n$ equações e satisfaz:

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

com $x_1, x_{r+2}, \ldots, x_n$ variáveis livres, ou seja, $n - r$. E assim, sucessivamente, a menos de uma permutação das variáveis, concluímos que o sistema possui infinitas soluções.

Os itens (ii) e (iii) foram provados nos casos anteriores e os mesmos provam a recíproca do item (i). ∎

Definição 1.45. *Dizemos que um sistema de equações lineares é possível e determinado (SPD), se o sistema possui uma única solução. Que é possível e indeterminado (SPI), se admite infinitas soluções e dizemos que é impossível (SI), se não admite solução.*

Exemplo 1.46. *Verifique se os sistemas abaixo possuem uma única, infinitas ou nenhuma solução.*

$$I: \begin{cases} x+y+z &= 3 \\ 2y+z &= 2 \\ y+2z &= 2, \end{cases} \quad II: \begin{cases} x+y+z &= 3 \\ y+2z &= 2 \\ x-z &= 1 \end{cases} \quad e \quad III: \begin{cases} x+y+z &= -10 \\ 2x+y+z &= -20 \\ y+z &= -40. \end{cases}$$

Considere a matriz ampliada associada ao sistema I e determinemos sua matriz na forma escalonada:

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 2 \\ 0 & 1 & 2 & 2 \end{pmatrix} \underset{L_3 \to L_3 - \frac{1}{2}L_2}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & \frac{3}{2} & 1 \end{pmatrix}.$$

Assim, o posto da matriz ampliada coincide com o posto da matriz dos coeficientes, que é exata-

mente o número de variáveis do sistema, isto é, $p_A = p_C = 3$. *Portanto, pelo teorema anterior o sistema possui uma única solução. Determinemos tal solução. O sistema associado a esta última matriz é*

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

ou seja, $z = \frac{2}{3}$. *Da segunda equação temos* $y = 1 - \frac{z}{2} = 1 - \frac{1}{2}\cdot\frac{2}{3} = \frac{2}{3}$. *E da primeira equação*

$$x = 3 - y - z = 3 - \frac{2}{3} - \frac{2}{3} = 3 - \frac{4}{3} = \frac{5}{3}.$$

Segue do Teorema 1.43 que os sistemas inicial e este último possuem mesmo conjunto solução, dado por

$$S = \left\{\left(\frac{5}{3}, \frac{2}{3}, \frac{2}{3}\right)\right\}.$$

Considere agora a matriz ampliada associada ao sistema de equações lineares II, *e determinemos sua matriz na forma escalonada:*

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

Neste caso, $p_A = p_C = 2 < 3$, *ou seja, o sistema possui infinitas soluções. Vamos determinar tais soluções. Associemos um novo sistema linear a esta última matriz,*

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

Como isso obtemos $y = 2 - 2z$ *o qual substituindo na primeira equação temos*

$$x = 3 - y - z = 3 - 2 + 2z - z = z + 1.$$

Portanto, o conjunto solução do sistema é $S = \{(z+1, 2-2z, z);\ z \in \mathbb{R}\}$. *Geometricamente significa que a interseção dos três planos descritos pelas equações do sistema inicial (ou dos dois planos descritos no segundo sistema) se intersectam em uma reta, parametrizada por* $(z+1, 2-2z, z)$, *ou ainda, chamando* $z = t$ *temos a equação paramétrica da reta*

$$\begin{cases} x = 1 + t \\ y = 2 - 2t \\ z = t. \end{cases}$$

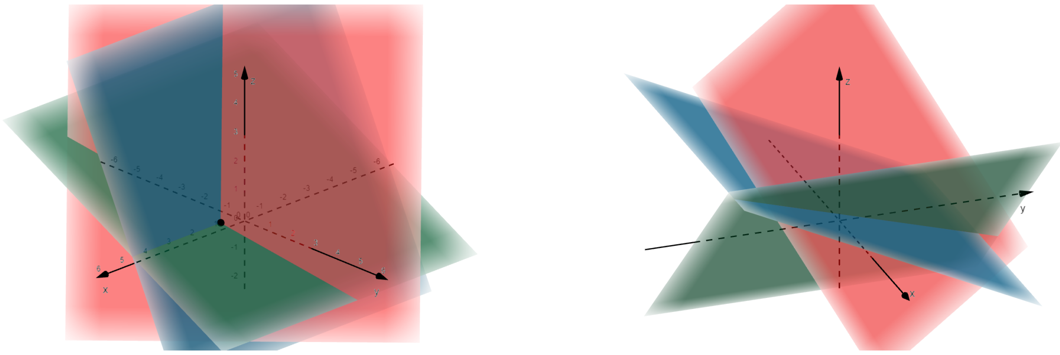

Figura 1.4: Representação geométrica das equações do sistema I e II

Finalmente, considere a matriz ampliada associada ao sistema III, e determinemos sua matriz na forma escalonada:

$$\begin{pmatrix} 1 & 1 & 1 & -10 \\ 2 & 1 & 1 & -20 \\ 0 & 1 & 1 & -40 \end{pmatrix} \underset{L_2 \to L_2 - 2L_1}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & -10 \\ 0 & -1 & -1 & 0 \\ 0 & 1 & 1 & -40 \end{pmatrix} \underset{L_3 \to L_3 + L_1}{\longrightarrow} \begin{pmatrix} 1 & 1 & 1 & -10 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & -40 \end{pmatrix}.$$

Portanto, $p_A = 3$ e $p_C = 2$. Como $p_A \neq p_C$ concluímos que o sistema não possui solução. Apenas observe que a equação associada à última linha do sistema acima, seria

$$0x + 0y + 0z = -40,$$

ou seja, não existe solução.

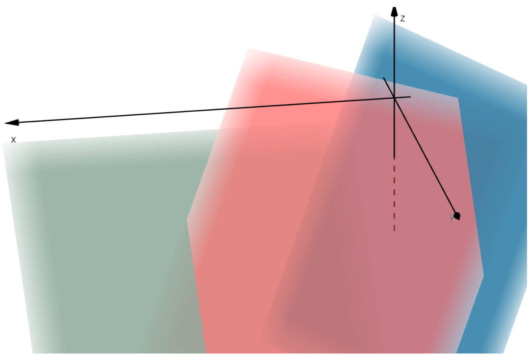

Figura 1.5: Representação geométrica das equações do sistema III

Uma classe interessante de sistemas de equações lineares são os chamados sistemas homogêneos, pois eles sempre admitem solução.

Definição 1.47. *(**Sistema Homogêneo**) Um sistema de equações lineares em que todos os termos independentes são iguais a zero, é dito um sistema homogêneo. Digamos, um sistema com m equações e n variáveis dado por forma:*

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

com $a_{ij}, b_i \in \mathbb{K}$, $1 \le i \le m$ e $1 \le j \le n$ é um sistema homogêneo.

Observe que um sistema homogêneo sempre tem solução, já que $x_1 = \cdots = x_n = 0$ é solução, chamada solução trivial. Logo, se um sistema homogêneo tiver uma única solução, então essa é a solução trivial.

1.5 Determinante

O determinante de uma matriz quadrada é uma função que associa a cada matriz $\mathcal{M}_{n\times n}(\mathbb{K})$ um número em $\mathbb{K}$. Conceito extremamente útil na matemática e que aparece nas mais diversas áreas. Particularmente em Álgebra Linear, esse conceito vai nos dar condições para a existência da inversa de uma matriz, nos propicia determinar as soluções de sistemas de equações lineares cujo número de equações e de variáveis coincidem, vai ser importante no cálculo de autovalores de uma transformação linear, entre tantas outras aplicações.

Definir a função determinante, caracterizando-a como a única função satisfazendo certas condições, não é uma tarefa fácil. Seria necessário introduzir uma série de conceitos e resultados, que muitas vezes não dispomos deste tempo num curso de graduação. Alguns autores definem utilizando o conceito de permutação.

Nestas notas, vamos definir o determinante de uma matriz utilizando para tanto um método denominado **desenvolvimento de Laplace** que apresenta uma fórmula recursiva para o cálculo do determinante, que recai em determinantes de matrizes de ordem mais baixa com relação à ordem da matriz inicial.

Dado uma matriz A de ordem $n \times n$ com entradas em $\mathbb{K}$, o **determinante** de $A = (a_{ij})$, denotado por $det(A)$, é o número definido como segue:

Se $A = (b)$ é uma matriz 1×1 então $det(A) = b$. Se $n > 1$ fixemos uma linha i de A, então o determinante da matriz A é dada por

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot det(A_{ij}), \tag{1.5}$$

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

Observe que quanto mais elementos nulos tiver uma linha i, mais interessante é aplicar o método de Laplace fixando essa linha, uma vez que por (1.5) para cada coluna j, o elemento a_{ij} aparece multiplicando os fatores do somatório, ou seja, se $a_{ij}=0$ simplesmente não precisamos calcular essa parcela.

Algumas vezes utilizaremos a notação de duas barras verticais da forma $|A|$ para denotar o determinante de uma matriz A.

Exemplo 1.48. *1. Se* $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$*, então por (1.5) se fixamos, por exemplo, a primeira linha, temos que*

$$det(A)=(-1)^{1+1}\cdot a\cdot d+(-1)^{1+2}\cdot b\cdot c=ad-bc.$$

2. Fixada a primeira linha de uma matriz de ordem 3 *dada por*

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

e aplicando (1.5), obtemos:

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

O método prático de resolução de determinantes de matrizes 3×3 *é devido ao francês Pierre Fréderic Sarrus (1768-1861), cujo método ficou conhecido como* Regra de Sarrus.

3. Vamos agora aplicar o método de Laplace para uma matriz particular. Seja

$$A=\begin{pmatrix} 2 & 5 & -1 & 0 \\ 1 & -\frac{1}{2} & 3 & 7 \\ 2 & 0 & -3 & 0 \\ -3 & 1 & 0 & 12 \end{pmatrix}.$$

Neste caso, é natural que fixemos a $3^{a.}$ *linha, pois é a que possui o maior número de zeros.*

Assim por (1.5) temos que

$$det(A) = (-1)^{3+1} \cdot 2 \cdot \begin{vmatrix} 5 & -1 & 0 \\ -\frac{1}{2} & 3 & 7 \\ 1 & 0 & 12 \end{vmatrix} + (-1)^{3+3} \cdot (-3) \cdot \begin{vmatrix} 2 & 5 & 0 \\ 1 & -\frac{1}{2} & 7 \\ -3 & 1 & 12 \end{vmatrix}.$$

Neste passo poderíamos calcular os determinantes das duas matrizes de ordem 3 pela regra de Sarrus, todavia para continuarmos ilustrando o método de Laplace vamos fixar para a primeira matriz a $3^{a.}$ *linha e para a segunda matriz a* $1^{a.}$ *linha. Logo*

$$\begin{aligned} det(A) &= 2 \cdot \left[(-1)^{3+1} \cdot 1 \cdot \begin{vmatrix} -1 & 0 \\ 3 & 7 \end{vmatrix} + (-1)^{3+3} \cdot 12 \cdot \begin{vmatrix} 5 & -1 \\ -\frac{1}{2} & 3 \end{vmatrix} \right] \\ &\quad + (-3) \cdot \left[(-1)^{1+1} \cdot 2 \cdot \begin{vmatrix} -\frac{1}{2} & 7 \\ 1 & 12 \end{vmatrix} + (-1)^{1+2} \cdot 5 \cdot \begin{vmatrix} 1 & 7 \\ -3 & 12 \end{vmatrix} \right] \\ &= 2 \cdot \left[1 \cdot (-7) + 12 \cdot (15 - \tfrac{1}{2})\right] + (-3) \cdot [2 \cdot (-6 - 7) - 5(12 + 21)] \\ &= 2 \cdot 167 + (-3) \cdot 191 \\ &= -239. \end{aligned}$$

Proposição 1.49. *Seja* $A = (a_{ij})_{n \times n}$ *uma matriz quadrada. Então:*

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

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

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

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

(*e*) *Se uma matriz B for obtida de A somando a uma linha de A um múltiplo de outra linha* ($L_i \to L_i + kL_r$) *com* $k \in \mathbb{K}^*$*, então* $det(B) = det(A)$*;*

(*f*) $det(A^t) = det(A)$.

(*g*) $det(kA) = k^n det(A)$*;*

Demonstração:

(*a*) Segue de (1.5), que fixando justamente a linha nula, teremos que $det(A) = 0$.

(b) Para a prova desta propriedade, precisamos dos conceitos fundamentais para a definição de determinante.

(c) Suponha que $A = (a_{ij})_{n\times n}$ possui duas linhas L_i e L_r iguais. Seja B a matriz obtida de A trocando L_i com L_r. Por (1.5), como as linhas L_i e L_r são iguais, se fixamos a linha i temos que $det(B) = det(A)$. Entretanto pelo item (b), o determinante troca de sinal quando trocamos duas linhas, ou seja, $det(B) = -det(A)$. Portanto, a única possibilidade é $det(A) = 0$.

(d) Seja B a matriz obtida de A após a seguinte operação elementar: $L_i \to kL_i$. Fixando esta linha em (1.5) temos $det(B) = \sum_{j=1}^{n}(-1)^{i+j}ka_{ij}det(B_{ij})$. Observe que a matriz B_{ij} não possui a linha i qualquer que seja $j = 1, \ldots, n$. E como apenas a linha i está multiplicada por k, concluímos que $B_{ij} = A_{ij}$. Portanto,

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

(e) Seja B a matriz obtida de A pela operação elementar $L_i \to L_i + kL_r$ e aplique (1.5) nesta linha, obtendo $det(B) = \sum_{j=1}^{n}(-1)^{i+j}(a_{ij}+ka_{rj})det(B_{ij})$. Pelo mesmo argumento anterior veja que $B_{ij} = A_{ij}$ e ainda

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

onde

$$B = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{i1}+ka_{r1} & \ldots & a_{in}+ka_{rn} \\ \vdots & \ldots & \vdots \\ a_{r1} & \ldots & a_{rn} \\ \vdots & \ldots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \quad \text{e} \quad C = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{r1} & \ldots & a_{rn} \\ \vdots & \ldots & \vdots \\ a_{r1} & \ldots & a_{rn} \\ \vdots & \ldots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}.$$

Como C possui duas linhas iguais temos que $det(C) = 0$. Logo, $det(B) = det(A)$.

(f) Para a prova desta propriedade, precisamos dos conceitos fundamentais para a definição de determinante.

(g) Segue da Proposição 1.49 item (d) que se multiplicarmos uma linha da matriz por um escalar k o determinante sai multiplicado por este escalar. Como A possui ordem n temos que $det(kA) = k^n det(A)$. ■

Decorre da propriedade (f) do resultado anterior, que as propriedades do determinante que valem para as linhas de uma matriz, também valem para as colunas. No desenvolvimento de Laplace podemos fixar uma coluna e calcular o determinante de modo similar quando fixamos uma linha. Por exemplo, se uma matriz A tem duas colunas iguais, então seu determinante é igual a zero, pois sua transposta A^t têm duas linhas iguais, o que implica pelo item (c) do resultado anterior que $det(A) = det(A^t) = 0$. Assim por diante.

Proposição 1.50. *Seja $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ dada por $A = (a_{ij})_{n\times n}$ uma matriz triangular, então*

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

com $i = 1, \ldots, n$.

Demonstração: Seja $A = (a_{ij})_{n\times n}$ uma matriz triangular. Vamos supor que A seja uma matriz triangular superior, pois o resultado vai seguir de modo análogo no caso de triangular inferior. A prova será feita por indução sobre n.

Para $n = 2$ temos

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix},$$

com $a, b, c \in \mathbb{K}$. Neste caso, $det(A) = ac - 0b = ac = \prod_{i=1}^{2} a_{ii}$, em que $a_{11} = a$ e $a_{22} = c$. Suponha por indução que o resultado seja válido para matrizes triangulares de ordem $n-1$. Aplicando (1.5) ao fixar a última linha de A temos que $det(A) = (-1)^{n+n} a_{nn} det(A_{nn})$, em que A_{nn} é a submatriz de ordem $n-1$ de A, obtida retirando-se a linha n e coluna n. Como A é triangular superior segue que $a_{ij} = 0$ para todo $i > j$, com $i, j = 1, \ldots, n$. Em particular, o mesmo vale para $i, j = 1, \ldots, n-1$, ou seja, A_{nn} é uma matriz triangular superior. Por hipótese de indução,

$$det(A) = a_{nn} det(A_{nn}) = a_{nn} a_{11} \ldots a_{n-1n-1} = \prod_{i=1}^{n} a_{ii}.$$

■

Definição 1.51. *Uma matriz A de ordem $n \times m$ é dita uma* **matriz em blocos***, se A está subdividida em matrizes menores, chamadas blocos.*

Por exemplo,

$$\begin{pmatrix} -1 & 4 & | & 5 & 18 & | & 3 \\ 0 & 7 & | & 1 & 1 & | & 9 \\ -- & -- & -- & -- & -- & -- & -- \\ 25 & -1 & | & -i & 0 & | & 4 \end{pmatrix}.$$

Esta é uma classe interessante de matrizes, pois as propriedades de adição e multiplicação de matrizes em blocos podem ser obtidos efetuando o cálculo com os blocos, como se fossem produtos feitos isoladamente.

Proposição 1.52. *Sejam $A_1, A_2, \ldots, A_r$ matrizes quadradas e A uma matriz quadrada em blocos da forma*

$$\begin{pmatrix} A_1 & * & * & * & * & * \\ & A_2 & * & * & * & * \\ & & & \ddots & * & * \\ & O & & & A_{r-1} & * \\ & & & & & A_r \end{pmatrix} \quad ou \quad \begin{pmatrix} A_1 & & & & & \\ * & A_2 & & & O & \\ * & * & * & \ddots & & \\ * & * & * & * & A_{r-1} & \\ * & * & * & * & * & A_r \end{pmatrix}.$$

Então, $det(A) = det(A_1) \cdot det(A_2) \cdot \ldots \cdot det(A_r)$.

Exemplo 1.53. *Seja uma matriz X em blocos da forma*

$$X = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

em que A é uma matriz $r \times r$, C é uma matriz $s \times s$, B é uma matriz $r \times s$ e 0 é a matriz nula $s \times r$. É possível provar que

$$det(X) = det(A) \cdot det(C).$$

Exemplo 1.54. *Considere a matriz de ordem 7 dada por*

$$A = \left(\begin{array}{cc|c|cccc} -1 & 4 & 77 & 5 & 18 & -150 & 3 \\ \underline{3} & \underline{2} & 4^9 & 1 & 1 & 37 & 9 \\ 0 & 0 & \underline{5} & 0 & -9 & -\sqrt{13} & \frac{14}{67} \\ \cline{4-7} 0 & 0 & 0 & \sqrt{51} & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & i & 0 & 0 \\ 0 & 0 & 0 & \frac{54}{129} & \sqrt{2} & 2 & 0 \\ 0 & 0 & 0 & -135 & 44 & \frac{7}{9} & -1 \end{array}\right).$$

Imagina o terror que nos passa pela cabeça só de imaginar calcular o determinante dessa matriz de ordem 7!! Na verdade, o que precisamos é enxergar que essa matriz é especial, pois é uma matriz em blocos. No caso, decidimos subdividir a matriz A em 3 blocos A_1, A_2 e A_3 ao longo da diagonal, a saber, o bloco A_1 tem ordem 2, A_2 tem ordem 1 e o bloco A_3 tem ordem 4. A matriz de ordem 4 poderia nos dar um trabalho maior, mas veja que ela é uma matriz triangular *inferior. Assim*

$$\begin{aligned} det(A) &= det(A_1) \cdot det(A_2) \cdot det(A_3) \\ &= [(-1) \cdot 2 - 3 \cdot 4] \cdot 5 \cdot [\sqrt{51} \cdot i \cdot 2 \cdot (-1)] \\ &= 7 - 14 \cdot 5 \cdot (-2i\sqrt{51}) \\ &= 140i\sqrt{51}. \end{aligned}$$

Teorema 1.55. *Sejam A e B são matrizes quadradas de ordem $n \times n$ e $k \in \mathbb{K}$. Então:*

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

Demonstração: A demonstração deste resultado será feito posteriormente.

1.6 Cálculo da Inversa de uma Matriz

Vamos apresentar um método para o cálculo da inversa de uma matriz.

Teorema 1.56. *Seja A uma matriz quadrada. As afirmações abaixo são equivalentes:*

(i) A é invertível;

(ii) A é equivalente por linhas à matriz identidade I;

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

Demonstração: $(i) \Rightarrow (ii)$ Suponhamos que A seja invertível. Como A é uma matriz quadrada observe que o sistema de equações lineares homogêneo $AX = 0$ (que possui n equações e n variáveis) admite uma única solução, pois $AX = 0$ se, e somente se,

$$A^{-1}AX = A^{-1}0 \Leftrightarrow X = 0.$$

Por outro lado, seja R a matriz na forma escada, equivalente por linhas à matriz A. Como vimos anteriormente no Teorema 1.43, sistemas associados à matrizes ampliadas equivalentes por linhas, possuem mesmo conjunto solução. Assim, os sistemas $AX = 0$ e $RX = 0$ possuem mesmo conjunto solução, ou seja, admitem uma única solução $X = 0$. Entretanto, como R está na forma escada e a única solução de $RX = 0$ é $X = 0$ temos que $R = I$. Portanto, $A \sim I$.
$(ii) \Rightarrow (iii)$ Se $A \sim I$ então existem matrizes elementares $E_1, \ldots, E_s$ tais que $E_s \cdots E_1 A = I$. Como cada E_i é invertível, o produto $E_s \cdots E_1$ também é invertível, assim $A = I(E_s \cdots E_1)^{-1}$. Logo,

$$A = E_1^{-1} \cdot \ldots \cdot E_s^{-1},$$

em que cada E_i^{-1} é uma matriz elementar com $i = 1, \ldots, s$.

$(iii) \Rightarrow (i)$ Seja $A = E_1 \cdot \ldots \cdot E_s$ onde cada E_i é uma matriz elementar, $i = 1, \ldots, s$. Como cada E_i é invertível segue que seu produto também o é. Portanto, A é invertível. ■

Corolário 1.57. *Se uma matriz quadrada A é linha equivalente à matriz identidade, então A é invertível e a sua inversa é obtida a partir da matriz identidade, aplicando-se a mesma sequência de operações nas linhas, ou ainda,*

$$(A : I) \longrightarrow (I : A^{-1}).$$

Demonstração: Pelo Teorema anterior, se $A \sim I$, então A é invertível. E mais, existem matrizes elementares $E_1, \ldots, E_s$ tais que

$$I = E_s \cdot \ldots \cdot E_1 \cdot A = (E_s \cdot \ldots \cdot E_1 \cdot I)A \;\Rightarrow\; A^{-1} = E_s \cdot \ldots \cdot E_1 \cdot I.$$

■

Exemplo 1.58. *Verifique se as matrizes abaixo admitem inversa:*

$$(a)\; A = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad (b)\; B = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 2 & 0 \end{pmatrix} \qquad (c)\; C = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & 3 \end{pmatrix}$$

Vamos utilizar o corolário anterior para verificar se A é invertível e, neste caso, determinar a sua inversa. Se a forma escada de A for a matriz identidade, então as mesmas operações realizadas nas linhas da matriz identidade vão nos dar a inversa da matriz A. Considere a matriz $(A : I)$:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \underset{\substack{L_2 \to L_2 - 3L_1 \\ L_3 \to L_3 - L_1}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & -5 & -6 & -3 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right) \underset{L_2 \to -\frac{1}{5}L_2}{\longrightarrow}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 1 & \frac{6}{5} & \frac{3}{5} & -\frac{1}{5} & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right) \underset{\substack{L_1 \to L_1 - 2L_2 \\ L_3 \to L_3 + L_2}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 1 & \frac{6}{5} & \frac{3}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{5} & -\frac{2}{5} & -\frac{1}{5} & 1 \end{array}\right) \underset{L_3 \to 5L_3}{\longrightarrow}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 1 & \frac{6}{5} & \frac{3}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & 1 & -2 & -1 & 5 \end{array}\right) \underset{\substack{L_1 \to L_1 + \frac{2}{5}L_3 \\ L_2 \to L_2 - \frac{6}{5}L_3}}{\longrightarrow} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 3 & 1 & -6 \\ 0 & 0 & 1 & -2 & -1 & 5 \end{array}\right).$$

Logo, a inversa da matriz A é dada por

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

Vamos deixar os cálculos das matrizes B e C como exercício.

Neste momento estamos preparados para provar que o determinante do produto é o produto dos determinantes. Para tanto, vamos provar o seguinte lema:

Lema 1.59. *Se E for uma matriz elementar de ordem n e A uma matriz qualquer de mesma ordem, então $det(EA) = det(E) \cdot det(A)$.*

Demonstração: Seja e uma operação elementar nas linhas da matriz A. Na Proposição 1.37 provamos que fazer uma operação elementar nas linhas de A, é o mesmo que fazer na identidade obtendo a matriz elementar E e multiplicar E à esquerda de A, isto é, $e(A) = EA$, onde $E = e(I)$. Analisemos as seguintes matrizes elementares e veja que o cálculo do determinante segue da Proposição 1.49:

(i) Se E é obtida de I por uma permutação de linhas segue que $det(E) = det(e(I)) = -det(I) = -1$. Logo,

$$det(EA) = det(e(A)) = -det(A) = (-1)det(A) = det(E) \cdot det(A).$$

(ii) Se E é obtida de I pela multiplicação de uma das linhas por uma constante não nula k, então $det(E) = det(e(I)) = k \cdot det(I) = k$. Donde segue que

$$det(EA) = det(e(A)) = k \cdot det(A) = det(E) \cdot det(A).$$

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

$$det(EA) = det(e(A)) = det(A) = 1 \cdot det(A) = det(E) \cdot det(A).$$

■

Teorema 1.60. *Sejam A e B matrizes quadradas de ordem n. Então*

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

Demonstração: Vamos provar este resultado inicialmente apenas no caso em que A e B são invertíveis. Neste caso, segue do Teorema 1.56 que A é um produto de matrizes elementares. Logo, $det(AB) = det(E_1 \ldots E_k B)$. Aplicando o lema anterior sucessivas vezes obtemos

$$\begin{aligned} det(AB) &= det(E_1) \cdot det(E_2 \ldots E_k B) = \ldots = det(E_1) \cdot det(E_2) \ldots det(E_k) \cdot det(B) \\ &= det(E_1) \ldots det(E_{k-1} \cdot E_k) \cdot det(B) = \ldots = det(E_1 E_2 \ldots E_k) \cdot det(B) \\ &= det(A) det(B). \end{aligned}$$

O caso em que A ou B não é invertível será provado posteriormente. ■

Nosso próximo passo é apresentarmos um critério para verificar se uma matriz é invertível, sem ter que tentar calcular a sua inversa. Além disso, veremos um processo alternativo para calcularmos a inversa de uma matriz.

Definição 1.61. *Dada uma matriz quadrada A, a matriz adjunta de A, que denotamos por $Adj(A) = (a'_{ij})$, é definida por*

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

em que A_{ij} é a submatriz de ordem $n-1$ obtida de A, retirando-se a linha i e a coluna j.

A proposição a seguir vai ser apresentada sem demonstração.

Proposição 1.62. *Sejam A uma matriz quadrada e I a matriz identidade, ambas de ordem n. Então*

$$A \cdot (Adj(A))^t = det(A) \cdot I.$$

Teorema 1.63. *Uma matriz quadrada A admite inversa se, e somente se, $det(A) \neq 0$. Nesse caso,*

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

Além disso, $det(A^{-1}) = \frac{1}{det(A)}$.

Demonstração: Se A admite inversa então $AA^{-1} = I$. Como o Teorema 1.60 foi provado no caso em que A e B são invertíveis, segue do mesmo que

$$1 = det(I) = det(AA^{-1}) = det(A) det(A^{-1}).$$

Portanto, $det(A) \neq 0$ e mais ainda $det(A^{-1}) = \frac{1}{det(A)}$.

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

$$A \cdot (Adj(A))^t = det(A) \cdot I \Rightarrow A \cdot \frac{1}{det(A)} (Adj(A))^t = I.$$

Como a inversa de A é única temos que

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

$\blacksquare$

Observação 1.64. *No Teorema 1.60, se A ou B não são invertíveis, temos que $A \cdot B$ não é invertível. De fato, suponha por contradição que AB seja invertível. Logo, existe uma matriz C satisfazendo $(AB)C = I$. Por associatividade de matrizes, $A(BC) = I$ e, portanto, A é invertível. Por outro lado, $C(AB) = I$ e assim $(CA)B = I$, garantindo que B é invertível. O que é uma contradição.*

Portanto, do Teorema 1.63 segue que $det(AB) = 0$. Como $det(A) = 0$ ou $det(B) = 0$ obtemos que $det(AB) = det(A)det(B)$ também neste caso.

Exemplo 1.65. *Considere a matriz*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{2\times 2}(\mathbb{K}).$$

Sabemos que A é invertível se, e somente se, $det(A) = ad - bc \neq 0$. Vamos calcular a inversa de A usando sua matriz adjunta. Por definição, $Adj(A) = (a'_{ij})$ em que

$$a'_{11} = (-1)^{1+1}det(d) = d, \qquad a'_{12} = (-1)^{1+2}det(c) = -c,$$

$$a'_{21} = (-1)^{2+1}det(b) = -b, \qquad a'_{22} = (-1)^{2+2}det(a) = a.$$

Desse modo,

$$A^{-1} = \frac{1}{det(A)} \cdot Adj(A)^t = \frac{1}{det(A)} \cdot \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^t.$$

Portanto,

$$A^{-1} = \frac{1}{det(A)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

ou seja, a inversa de uma matriz quadrada de ordem 2, pode ser obtida de modo direto da matriz original A, bastando multiplicar o inverso do determinante de A pela matriz obtida de A, invertendo as posições dos elementos da sua diagonal principal e trocando os sinais dos elementos da diagonal secundária.

Calcule a inversa dessa mesma matriz utilizando o processo de escalonamento.

1.7 Regra de Cramer

O cálculo da inversa de uma matriz nos fornece um outro método de resolução de sistemas de equações lineares. Entretanto este método se aplica apenas a sistemas lineares cujo número de equações coincide com o número de variáveis. Considere um sistema de equações lineares com n equações e n variáveis.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{cases} \tag{1.6}$$

o qual podemos escrever matricialmente por $AX = B$, em que $A = (a_{ij})$ é a matriz dos coeficientes, $B = (b_j)$ a matriz dos termos independentes e $X = (x_j)$ a matriz das variáveis, para $i, j = 1, \ldots, n$. Seja A_i a matriz obtida de A substituindo a i-ésima coluna de A pela matriz coluna B.

Teorema 1.66. ***(Regra de Cramer)*** *O sistema (1.6) admite solução única se, e somente se, $det(A) \neq 0$. Nesse caso, a solução $S = \{(w_1, \ldots, w_n)\}$ é dada por*

$$w_1 = \frac{det(A_1)}{det(A)}, \quad w_2 = \frac{det(A_2)}{det(A)}, \ \ldots, \ w_n = \frac{det(A_n)}{det(A)}.$$

Demonstração: Suponha que o sistema $AX = B$ admite uma única solução e seja $(A : B)$ a matriz ampliada do sistema. Seja $(R : B')$ a matriz na forma escada equivalente por linhas à matriz ampliada $(A : B)$. Logo, pelo Teorema 1.43, o sistema $RX = B'$ e $AX = B$ possuem o mesmo conjunto solução, que é única. Entretanto, R está na forma escada, assim para termos uma única solução do sistema, temos que $R = I$. Logo, $A \sim R = I$. Portanto, pelo Teorema 1.56 a matriz A é invertível e, portanto, $det(A) \neq 0$.

Reciprocamente, se $det(A) \neq 0$ temos que A é invertível, logo

$$A^{-1}(AX) = A^{-1}B \Rightarrow (A^{-1}A)X = A^{-1}B \Rightarrow IX = A^{-1}B \Rightarrow X = A^{-1}B.$$

Portanto, essa é a solução do sistema, o produto da inversa da matriz A pela matriz B. Todavia, nosso propósito é obtermos a solução explícita do sistema.

Sabemos que $A^{-1} = \frac{1}{det(A)}(Adj(A))^t$. Assim se $A = (a_{ij})$ e $Adj(A) = (c_{ij})$, então

$$c_{ij} = (-1)^{i+j} det(A_{ij}).$$

Logo,

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

Logo, $x_i = \dfrac{c_{1i}b_1 + \ldots + c_{ni}b_n}{det(A)}$, para todo $i = 1, \ldots, n$. Por outro lado, vamos calcular o determinante da matriz A_j, definida acima, aplicando Laplace na coluna j, temos que:

$$A_j = \begin{pmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{pmatrix}$$

e $det(A_i) = \displaystyle\sum_{j=1}^{n} b_j(-1)^{j+i}det(A_{ji}) = \sum_{j=1}^{n} b_j c_{ji}$. Portanto, a solução $S = \{(w_1, \ldots, w_n)\}$ do sistema dado é tal que

$$w_i = \frac{det(A_i)}{det(A)}$$

para todo $i = 1, \ldots, n$. ■

Corolário 1.67. *Um sistema de equações lineares homogêneo $AX = 0$ com n equações e n variáveis, possui infinitas soluções se, e somente se, $det(A) = 0$.*

Demonstração: Este corolário é imediato do teorema anterior e do fato de que todo sistema linear homogêneo sempre admite solução. ■

CAPÍTULO 2

ESPAÇOS VETORIAIS

Neste capítulo vamos explorar os conceitos referentes a espaços vetoriais, e veremos o quanto o conceito de vetor é muito mais amplo do que foi abordado num curso de geometria analítica.

Seja V um conjunto não vazio, munido de duas operações binárias: uma adição e uma multiplicação por um escalar em um corpo $\mathbb{K}$, respectivamente, dadas por:

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

Em geral, vamos trabalhar com o corpo dos números reais $\mathbb{R}$ ou dos números complexos $\mathbb{C}$.

Definição 2.1. *Seja V um conjunto não vazio, munido de uma soma e de um produto por escalar como acima. Dizemos que V é um* ***espaço vetorial sobre um corpo*** $\mathbb{K}$ *se satisfaz as seguintes condições: para quaisquer $u, v, w \in V$ e $k, k_1, k_2 \in \mathbb{K}$ temos*

A1. $(u + v) + w = u + (v + w)$; *(Associatividade da soma)*

A2. $u + v = v + u$; *(Comutatividade da soma)*

A3. Existe $e \in V$ tal que $u + e = u$. Vamos denotar o elemento e por 0; (Elemento neutro)

A4. Para cada $u \in V$, existe $u' \in V$ tal que $u + u' = 0$. O elemento u' será denotado por $-u$;

M1. $k(u + v) = ku + kv$; *(Distributiva com respeito à soma de vetores)*

M2. $(k_1 + k_2)u = k_1u + k_2u$; *(Distributiva com respeito à soma de escalares)*

M3. $k_1(k_2u) = (k_1k_2)u$; *(Associativa do produto por escalar)*

M4. $1 \cdot u = u$.

Um espaço vetorial sobre $\mathbb{K}$ também vai ser denominado de $\mathbb{K}$-espaço vetorial.

Os elementos de um espaço vetorial são denominados de **vetores**, os elementos de $\mathbb{K}$ de escalares. E ainda, o elemento 0 de V será chamado de **vetor nulo** e o elemento $-u$ de **vetor oposto** de u, já que $-u$ é o oposto (com respeito à soma) do vetor u.

Exemplo 2.2. *Vamos apresentar alguns exemplos de espaços vetoriais, cuja verificação das propriedades que caracterizam um espaço vetorial será deixado como exercício:*

1. *Considere o conjunto de todos os pares ordenados de números reais que denotamos por $\mathbb{R}^2$, ou seja, $\mathbb{R}^2 = \{(x, y);\ x, y \in \mathbb{R}\}$ munido das operações usuais: dados $(x, y), (z, w) \in \mathbb{R}^2$*

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

O elemento $(0, 0)$ é o elemento neutro de $\mathbb{R}^2$ e para cada $u = (x, y) \in \mathbb{R}^2$, o elemento $(-x, -y)$ é o oposto de u, ou seja, $-u = (-x, -y)$. O conjunto $\mathbb{R}^2$ com as operações acima é um espaço vetorial sobre $\mathbb{R}$.

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

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

não é um espaço vetorial, pois não possui elemento neutro. De fato, seja $e = (e_1, e_2)$ tal que $(x, y) + (e_1, e_2) = (x, y)$. Por outro lado, $(x, y) + (e_1, e_2) = (x + e_1, 0)$. Logo, $(x + e_1, 0) = (x, y)$ o que implica que $e_1 = 0$ e $y = 0$, ou seja, não existe elemento neutro.

3. *Mais geralmente, denotemos por $\mathbb{R}^n$ o conjunto de todas as n-uplas ordenadas de números reais, ou seja,*

$$\mathbb{R}^n = \{(x_1, \ldots, x_n);\ x_1, \ldots, x_n \in \mathbb{R}\}.$$

Dados $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$ e $k \in \mathbb{R}$ consideramos as operações:

$$\begin{aligned} &(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n) \quad e \\ &k(x_1, \ldots, x_n) = (kx_1, \ldots, kx_n). \end{aligned}$$

De modo análogo, $(0, \ldots, 0)$ é o elemento neutro de $\mathbb{R}^n$ e para cada $u = (x_1, \ldots, x_n)$, seu oposto é o elemento $-u = (-x_1, \ldots, -x_n)$. É possível provar que $\mathbb{R}^n$ é um espaço vetorial real ou sobre $\mathbb{R}$. Em particular, $\mathbb{R}^n$ também é um espaço vetorial sobre $\mathbb{Q}$.

4. Analogamente ao caso anterior, definimos

$$\mathbb{C}^n = \{(z_1, \ldots, z_n);\ z_1, \ldots, z_n \in \mathbb{C}\}$$

o conjunto de todas as n-uplas ordenadas de números complexos. Seja $\mathbb{K}$ o corpo $\mathbb{C}, \mathbb{R}$ ou $\mathbb{Q}$. Sejam $(z_1, \ldots, z_n), (w_1, \ldots, w_n) \in \mathbb{C}^n$, $k \in \mathbb{K}$ e as operações

$$\begin{aligned} &(z_1, \ldots, z_n) + (w_1, \ldots, w_n) = (z_1 + w_1, \ldots, z_n + w_n) \quad e \\ &k(z_1, \ldots, z_n) = (kz_1, \ldots, kz_n). \end{aligned}$$

Assim $\mathbb{C}^n$ munido das operações acima é um espaço vetorial sobre $\mathbb{K}$.

5. O conjunto das matrizes de ordem $n \times m$ com entradas em $\mathbb{K}$, denotado por $\mathcal{M}_{n\times m}(\mathbb{K})$, munido das operações usuais de soma de matrizes e produto de uma matriz por um escalar $k \in \mathbb{K}$, ou seja, dados $A, B \in \mathcal{M}_{n\times m}(\mathbb{K})$ com $A = (a_{ij})$, $B = (b_{ij})$ e $k \in \mathbb{K}$ em que

$$A + B = (a_{ij} + b_{ij}) \quad e \quad kA = (ka_{ij})$$

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

6. Sejam X um conjunto não vazio e $\mathbb{K}$ o corpo $\mathbb{R}$ ou $\mathbb{C}$. Consideremos o conjunto

$$\mathcal{F}(X) = \{f : X \to \mathbb{K};\ \textit{funções}\}.$$

Note que $\mathcal{F}(X)$ é não vazio, já que X é não vazio. Vamos considerar as seguintes operações: dados $f, g \in \mathcal{F}(X)$ e $k \in \mathbb{K}$ defina

$$(f + g)(x) = f(x) + g(x) \quad e \quad (kf)(x) = kf(x).$$

Elemento Neutro: a função nula 0 definida por $0(x) = 0$ para todo $x \in X$ é o elemento neutro de $\mathcal{F}(X)$.

Elemento Oposto: dado $f \in \mathcal{F}(X)$, a função $-f \in \mathcal{F}(X)$ definida por $(-f)(x) = -f(x)$ será o elemento oposto de f.

Com estas operações $\mathcal{F}(X)$ é um espaço vetorial sobre $\mathbb{K}$.

7. Seja $\mathcal{P}$ o conjunto de todos os polinômios na variável x com coeficientes no corpo $\mathbb{K}$ e qualquer grau maior ou igual a 0. Um elemento deste conjunto é da forma

$$a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0,$$

com $a_i \in \mathbb{K}$. *Sejam* $n, m \in \mathbb{N}$ *digamos com* $n \geq m$, $k \in \mathbb{K}$ *e* $p, q \in \mathcal{P}$ *dados por* $p = a_n x^n + \ldots + a_1 x + a_0$ *e* $q = b_m x^m + \ldots + b_1 x + b_0$, *com* $a_i, b_j \in \mathbb{K}$, $0 \leq i \leq n$ *e* $0 \leq j \leq m$. *Definimos a soma e o produto por escalar nesse conjuntos, respectivamente, por:*

$$\begin{aligned} & p + q = a_n x^n + \ldots + a_{m+1} x^{m+1} + (a_m + b_m) x^m + \ldots + (a_1 + b_1) x + (a_0 + b_0) \\ & e \quad kp = k a_n x^n + \ldots + k a_1 x + k a_0. \end{aligned}$$

Elemento Neutro: o polinômio nulo é o elemento neutro de $\mathcal{P}$.

Elemento Oposto: dado $p = a_n x^n + \ldots + a_1 + a_0 \in \mathcal{P}$, *o polinômio*

$$-p = (-a_n) x^n + \ldots + (-a_1) x + (-a_0) \in \mathcal{P}$$

é o elemento oposto de p.

Com isso, $\mathcal{P}$ *é um espaço vetorial sobre* $\mathbb{K}$. *Uma notação usual para o conjunto de todos os polinômios em* $\mathcal{P}$ *é* $\mathbb{K}[x]$.

Apresentamos agora algumas propriedades básicas que serão muito úteis posteriormente:

Observação 2.3. *Sejam* V *um espaço vetorial sobre um corpo* $\mathbb{K}$, $v \in V$ *e* $k \in \mathbb{K}$. *Então:*

1. $kv = 0$ *se, e somente se,* $k = 0$ *ou* $v = 0$;
2. $(-1)v = -v$.

Apresentemos a solução do item 1.

Suponhamos primeiramente que $v = 0$. *Como* 0 *é o vetor nulo, segue da definição de espaço vetorial, item* $M1$, *que* $k0 = k(0 + 0) = k0 + k0$. *Segue ainda do item* $A4$ *que somando o oposto do vetor* $k0$ *de ambos os lados da última igualdade obtemos*

$$k0 + (-k0) = k0 + k0 + (-k0) \quad \Rightarrow \quad 0 = k0 + 0 \quad \Rightarrow \quad kv = k0 = 0.$$

Analogamente se $k = 0$.

Reciprocamente, suponha que $kv = 0$ *com* $k \neq 0$. *Então, como* $\mathbb{K}$ *é um corpo, o elemento* k *admite inverso. Deste modo, multiplicando o inverso de* k *de ambos os lados da igualdade obtemos por* $M3$, $M4$ *e a recíproca que acabamos de provar, que*

$$kv = 0 \;\Rightarrow\; (k^{-1}) kv = k^{-1} 0 \;\Rightarrow\; (k^{-1} k) v = 0 \;\Rightarrow\; 1v = 0 \;\Rightarrow\; v = 0.$$

Para o item 2. queremos verificar que dado $v \in V$, *o oposto de* v *é o elemento* $(-1)v$. *De* *fato, pela propriedade $M4$ e $M2$, temos*

$$(-1)v + v = (-1)v + 1 \cdot v = ((-1) + 1) \cdot v = 0 \cdot v = 0.$$

Como queríamos verificar.

2.1 Subespaços Vetoriais

Uma questão que se coloca é: quais são os subconjuntos de um espaço vetorial que também satisfazem as propriedades de se manter sendo um espaço vetorial? Este é o conceito que vamos explorar agora.

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

O resultado a seguir nos dá um método para verificarmos quando um subconjunto de um espaço vetorial é um subespaço do mesmo.

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

Demonstração: Seja S um subespaço de V. Como por hipótese S é não vazio, existe um elemento $u \in S$. Mas sendo S um subespaço de V temos que S é um espaço vetorial e assim dado qualquer $k \in \mathbb{K}$ temos que $ku \in S$. Logo, se $v \in S$, então pelo mesmo argumento de S ser um espaço vetorial temos que $ku + v \in S$.

Reciprocamente, suponhamos que para todo $u, v \in S$ e $k \in \mathbb{K}$ tenhamos $ku + v \in S$. Como S é não vazio, existe um elemento $v \in S$. Em particular, v é um elemento do espaço vetorial V e, portanto, o oposto de v também pertence a V. Logo, pelo exercício anterior,

$$0 = -v + v = (-1)v + v \in S,$$

ou seja, $0 \in S$. Se $u, v \in S \subset V$, então $u + v = 1 \cdot u + v \in S$. Se $u \in S$ é um vetor arbitrário e $k \in \mathbb{K}$ é um escalar qualquer, então $ku = ku + 0 \in S$. As demais propriedades de espaço vetorial são satisfeitas para os elementos de S, pois S é um subconjunto de V que é um espaço vetorial. Logo, S é um subespaço de V. ∎

Observe que todo subespaço S de um espaço vetorial V possui o vetor nulo. Claramente V e $\{0\}$ são subespaços do espaço vetorial V, ditos subespaços triviais.

Exemplo 2.6. 1. *Seja S o subconjunto de $\mathbb{R}^2$ dado por*

$$S = \{(x, x);\ x \in \mathbb{R}\}.$$

Vamos verificar que S é, na verdade, um subespaço de $\mathbb{R}^2$.

Observe que $(0, 0) \in S$ 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

$$ku + v = (kx + y, kx + y) \in S.$$

Logo, pela proposição anterior, S é um subespaço de $\mathbb{R}^2$. Geometricamente, podemos enxergar S como sendo o conjunto dos pontos da reta $y = x$, ou seja,

$$S = \{(x, x);\ x \in \mathbb{R}\} = \{(x, y) \in \mathbb{R}^2;\ y = x\}.$$

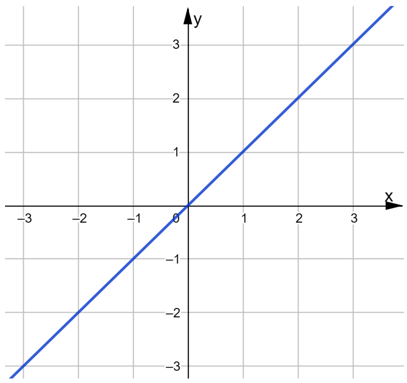

Figura 2.1: Subespaço de $\mathbb{R}^2$ - Reta $y = x$

*Observe que um subconjunto de $\mathbb{R}^2$, cujos pontos estejam em uma reta que não passa por $(0, 0)$, **não** pode ser um subespaço de $\mathbb{R}^2$, pois não tem o elemento neutro do espaço vetorial.*

2. *O conjunto*

$$S = \{(x, y, z) \in \mathbb{R}^3;\ x = y + 1\}$$

não é um subespaço vetorial de $\mathbb{R}^3$, pois $(0, 0, 0) \notin S$. Geometricamente, a equação descreve um plano em $\mathbb{R}^3$ que não passa pela origem.

3. *Verifique que o conjunto $S = \{(x, y, z) \in \mathbb{R}^3;\ ax + by + cz = 0,\ \text{com } (a, b, c) \neq 0\}$ é um subespaço de $\mathbb{R}^3$. Geometricamente S descreve um plano em $\mathbb{R}^3$ que passa pela origem.*

4. *O conjunto*

$$S = \{(x, y, z) \in \mathbb{R}^3;\ x = 0 \quad e \quad y \geq z\}$$

não é um subespaço vetorial de $\mathbb{R}^3$ *com as operações usuais.*

Observe que $(0,0,0) \in S$*. Agora, se* $u \in S$ *e* $k \in \mathbb{R}$ *segue que* $u = (0, y, z)$ *com* $y \geq z$*. Logo,*

$$ku = k(0, y, z) = (0, ky, kz).$$

Todavia, se $k > 0$ *como* $y \geq z$ *temos que* $ky \geq kz$*. Entretanto, se* $k < 0$ *teremos* $ky \leq kz$*, o que implica* $(0, ky, kz) \notin S$*. Logo,* S *não é subespaço de* $\mathbb{R}^3$*.*

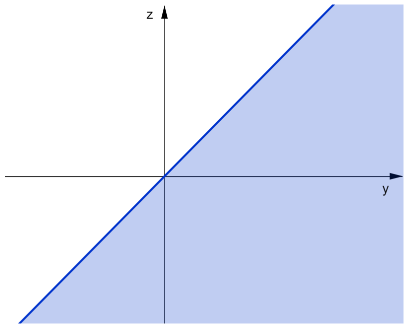

Figura 2.2: Projeção no plano yz de $S = \{(x, y, z) \in \mathbb{R}^3;\ x = 0 \text{ e } y \geq z\}$

5. *Por um exercício da lista vimos que uma matriz* A *é simétrica se, e somente se,* $A = A^t$*. Verifiquemos que 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* $\mathcal{M}_{n\times n}(\mathbb{K})$*.*

Observe que a matriz nula $0 \in S$*, pois* $0 = 0^t$*. Sejam* $A, B \in S$ *e* $k \in \mathbb{K}$*. Deste modo,* $A^t = A$ *e* $B^t = B$*. Logo, pela Proposição 1.14 itens (i) e (ii) temos*

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

ou seja, $kA + B \in S$*. Pela Proposição 2.5,* S *é um subespaço de* $\mathcal{M}_{n\times n}(\mathbb{K})$*.*

6. *O conjunto das funções contínuas definidas em* X *é um subespaço vetorial de* $\mathcal{F}(X)$*.*

De fato, a função nula é uma função contínua. E sabemos que soma de funções contínuas é uma função contínua, e produto de um escalar por uma função contínua é uma função contínua.

7. *Seja* $\mathcal{P}_n$ *o subconjunto de* $\mathcal{P}$ *que consiste dos polinômios de grau* $\leq n$*, para* n *fixo, ou seja,*

$$\mathcal{P}_n = \{a_n x^n + \ldots + a_2 x^2 + a_1 x + a_0;\ a_i \in \mathbb{K},\ i = 0, 1, \ldots, n\}.$$

Por exemplo, $\mathcal{P}_2 = \{a_2x^2 + a_1x + a_0;\ a_i \in \mathbb{K},\ i = 0, 1, 2\}$ *é o subconjunto de* $\mathcal{P}$ *de todos os polinômios de grau menor ou igual a* 2.

Vamos verificar que $\mathcal{P}_n$ *é um subespaço de* $\mathcal{P}$*. Com efeito, observe que o polinômio nulo pertence a* $\mathcal{P}_n$*, basta tomar* $a_0 = a_1 = \ldots = a_n = 0$*. Sejam*

$$p = a_nx^n + \ldots + a_2x^2 + a_1x + a_0 \quad e \quad q = b_nx^n + \ldots + b_2x^2 + b_1x + b_0$$

polinômios em $\mathcal{P}_n$ *e* $k \in \mathbb{K}$*. Então,*

$$kp + q = (ka_n + b_n)x^n + \ldots + (ka_2 + b_2)x^2 + (ka_1 + b_1)x + (ka_0 + b_0) \in \mathcal{P}_n,$$

pois a soma de polinômios de grau $\leq$ *é um polinômio de grau* $\leq n$*.*

Observação 2.7. *O conjunto dos polinômios de grau* n ***não*** *é um subespaço vetorial de* $\mathcal{P}$*, pois a soma de dois polinômios de grau* n *pode ter grau estritamente menor que* n*. Por exemplo, sejam* $p = 2x^3 - 4x^2 + x - 1$ *e* $q = -2x^3 + 4x + 5$ *polinômios de grau 3. Note que*

$$p + q = -4x^2 + 5x + 4$$

possui grau menor do que 3.

Podemos definir algumas operações entre subespaços. Uma questão natural é se a união ou a interseção de subespaços de um espaço vetorial V é ainda um subespaço de V? A proposição a seguir nos dá uma resposta no caso da interseção.

Proposição 2.8. *Sejam* V *um espaço vetorial sobre* $\mathbb{K}$ *e* S_i *com* $i \in J$ *uma coleção arbitrária de subespaços de* V*. Então* $\bigcap_{i \in J} S_i$ *é um subespaço de* V*.*

Demonstração: Como todo S_i com $i \in J$ é um subespaço de V temos que $0 \in S_i$ para todo $i \in J$. Logo, $0 \in \bigcap_{i \in J} S_i$. Agora, sejam $u, v \in \bigcap_{i \in J} S_i$ e $k \in \mathbb{K}$. Assim $u, v \in S_i$ para todo $i \in J$. Como S_i é subespaço de V segue que $ku + v \in S_i$ para todo $i \in J$. Logo $ku + v \in \bigcap_{i \in J} S_i$, concluindo o resultado. ■

Exercício 2.9. *Considere os seguintes subespaços de* $\mathbb{R}^3$*:*

$$S_1 = \{(x, y, z) \in \mathbb{R}^3;\ ax + by + cz = 0,\ (a, b, c) \neq 0\} \quad e$$

$$S_2 = \{(x, y, z) \in \mathbb{R}^3;\ dx + ey + fz = 0,\ (d, e, f) \neq 0\}$$

que descrevem planos em $\mathbb{R}^3$ *que passam pela origem. Suponha que* $a \neq 0$ *e que* $ea - bd \neq 0$*. Determine* $S_1 \cap S_2$ *e verifique que ele é um subespaço de* $\mathbb{R}^3$*.*

Veremos no exemplo abaixo que a união de subespaços não é necessariamente um subespaço.

Exemplo 2.10. *Considere* $U = \{(x, y) \in \mathbb{R}^2;\ x = 0\}$ *e* $W = \{(x, y) \in \mathbb{R}^2;\ y = x\}$ *subconjuntos de* $\mathbb{R}^2$*, que facilmente podemos verificar que são subespaços de* $\mathbb{R}^2$*. Entretanto, o conjunto* $U \cup W$ *não é um subespaço de* $\mathbb{R}^2$*, uma vez que os vetores* $u = (0, 1)$ *e* $v = (1, 1)$ *pertencem a* $U \cup W$*, mas* $u + v = (1, 2) \notin U \cup W$*.*

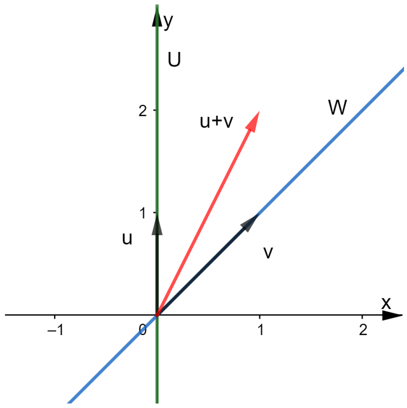

Figura 2.3: União dos subespaços U e W não é um subespaço

Sejam S_1 e S_2 subespaços de um espaço vetorial V. Definimos a soma de S_1 e S_2 por

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

Verifique que $S_1 + S_2$ é um subespaço de V.

Dado um espaço vetorial V sobre $\mathbb{K}$, muitas vezes é conveniente escrever seus elementos como soma de elementos de dois (ou mais) subespaços. A fim de obter uma unicidade de tal decomposição definimos:

Definição 2.11. *Sejam* W_1, W_2 *dois subespaços vetoriais de* V*. Diremos que a soma* $W_1 + W_2$ *é direta se* $W_1 \cap W_2 = \{0\}$*. Neste caso, escrevemos* $W_1 \oplus W_2$*.*

Proposição 2.12. *Sejam* W_1, W_2 *dois subespaços de um espaço vetorial* V *sobre* $\mathbb{K}$*. Então* $V = W_1 \oplus W_2$ *se, e somente se, todo vetor* $v \in V$ *se escreve de maneira única como* $v = w_1 + w_2$ *com* $w_i \in W_i, i = 1, 2$*.*

Demonstração: Por hipótese, cada elemento v de V se escreve da forma $v = w_1 + w_2$ com $w_i \in W_i, i = 1, 2$. Suponha que v também se escreva da forma $v = u_1 + u_2$ para algum

$u_i \in W_i, i = 1, 2$. Logo, $w_1 + w_2 = v = u_1 + u_2$, ou seja,

$$w_1 - u_1 = u_2 - w_2 \in W_1 \cap W_2,$$

já que $w_1 - u_1 \in W_1$ e $u_2 - w_2 \in W_2$. Como $W_1 \cap W_2 = \{0\}$ segue que $w_1 = u_1$ e $w_2 = u_2$.

Reciprocamente, suponhamos que todo vetor v de V se escreve de modo único da forma $v = w_1 + w_2$ com $w_1 \in W_1$ e $w_2 \in W_2$. Então $V = W_1 + W_2$. Se $W_1 \cap W_2 \neq \{0\}$, então existe um vetor não nulo u em $W_1 \cap W_2$. Como W_2 é um subespaço de V e $u \in W_2$, então $-u \in W_2$. Assim $0 = u + (-u)$ com $u \in W_1$ e $-u \in W_2$. Por outro lado, $0 = 0 + 0$ com $0 \in W_1$ e $0 \in W_2$. Uma vez que $u \neq 0$, temos duas maneiras distintas de escrever um elemento de V como soma de um vetor de W_1 com outro de W_2, o que não ocorre. Portanto, $W_1 \cap W_2 = \{0\}$. ■

2.2 Base e Dimensão de um Espaço Vetorial

Nesta seção vamos abordar um tema importante em Álgebra Linear, que é o conceito de base e de dimensão de um espaço vetorial. Nestas notas vamos abordar apenas espaços vetoriais de dimensão finita.

Para tanto vamos iniciar com o conceito de combinação linear.

Definição 2.13. *Sejam $v_1, \ldots, v_n$ vetores de um $\mathbb{K}$-espaço vetorial V. Um vetor $u \in V$ é dito uma **combinação linear** dos vetores $v_1, \ldots, v_n$ se existem escalares $k_1, \ldots, k_n \in \mathbb{K}$ tais que*

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

Exemplo 2.14. 1. *O vetor $u = (-1, -1) \in \mathbb{R}^2$ é combinação linear dos vetores $v_1 = (1, 2)$ e $v_2 = (4, 1)$?*

Para que essa afirmação seja verdadeira é necessário que existam $k_1, k_2 \in \mathbb{R}$ tais que

$$(-1, -1) = k_1(1, 2) + k_2(4, 1).$$

Logo, precisamos verificar se o sistema

$$\begin{cases} k_1 + 4k_2 = -1 \\ 2k_1 + k_2 = -1 \end{cases}$$

possui solução. Veja que a solução do sistema é $k_1 = -\frac{3}{7}$ *e* $k_2 = -\frac{1}{7}$. *Portanto,*

$$(-1,-1) = -\frac{3}{7}\cdot(1,2) + \left(-\frac{1}{7}\right)\cdot(4,1).$$

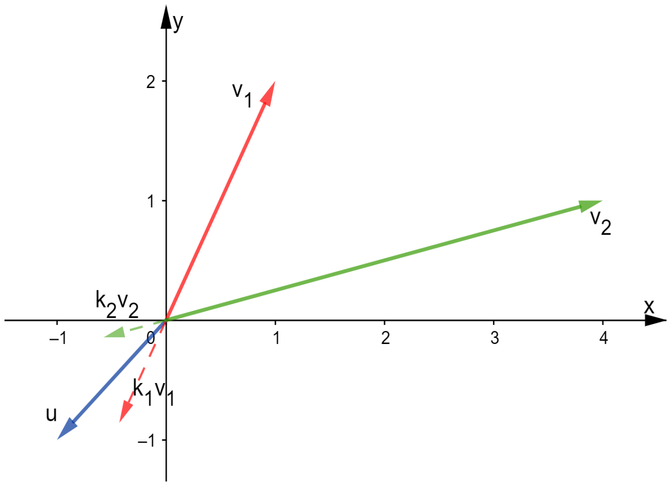

Figura 2.4: $u = (-1,-1) = k_1v_1 + k_2v_2 = -\frac{3}{7}(1,2) + \left(-\frac{1}{7}\right)(4,1)$

2. *Quais são os vetores* $(x,y,z) \in \mathbb{R}^3$ *que podem ser escritos como combinação linear de* $(1,0,0)$ *e* $(0,0,1)$. *Sejam* $k_1, k_2 \in \mathbb{R}$ *tais que*

$$(x,y,z) = k_1(1,0,0) + k_2(0,0,1) = (k_1, 0, k_2).$$

Logo, $k_1 = x$, $y = 0$ *e* $k_2 = z$. *Portanto* $\{(x,0,z);\ x,z \in \mathbb{R}\}$ *é o conjunto de todos os vetores escritos como combinação linear de* $(1,0,0)$ *e* $(0,0,1)$. *Geometricamente, trata-se do plano* xz, *o qual é um subespaço de* $\mathbb{R}^3$.

Definição 2.15. *Sejam* $v_1,\ldots,v_n$ *vetores de um* $\mathbb{K}$*-espaço vetorial* V. *Denotamos por* $[v_1,\ldots,v_n]$ *o conjunto de todas as combinações lineares de* $v_1,\ldots,v_n$, *ou seja,*

$$[v_1,\ldots,v_n] = \{k_1v_1 + \ldots + k_nv_n;\ k_1,\ldots,k_n \in \mathbb{K}\}.$$

Vamos verificar que o conjunto $[v_1,\ldots,v_n]$ *é um subespaço vetorial de* V, *chamado* ***subespaço vetorial gerado*** *por* $v_1,\ldots,v_n$.

Proposição 2.16. *Sejam* $v_1,\ldots,v_n$ *vetores de um* $\mathbb{K}$*-espaço vetorial* V. *O conjunto* $[v_1,\ldots,v_n]$ *é um subespaço de* V.

Demonstração: De fato, o elemento neutro 0 de V pertence a $[v_1, \ldots, v_n]$, pois

$$0 = 0v_1 + 0v_2 + \ldots + 0v_n.$$

Dados $u, v \in [v_1, \ldots, v_n]$ e $k \in \mathbb{K}$, existem $\alpha_i, \beta_i \in \mathbb{K}$ para $i = 1, \ldots, n$ tais que $u = \alpha_1 v_1 + \ldots + \alpha_n v_n$ e $v = \beta_1 v_1 + \ldots + \beta_n v_n$. Assim

$$\begin{aligned} ku + v &= k(\alpha_1 v_1 + \ldots + \alpha_n v_n) + \beta_1 v_1 + \ldots + \beta_n v_n \\ &= (k\alpha_1 + \beta_1)v_1 + \ldots + (k\alpha_n + \beta_n)v_n \in [v_1, \ldots, v_n], \end{aligned}$$

pois $k\alpha_i + \beta_i \in \mathbb{K}$ para todo $i = 1, \ldots, n$. Portanto, pela Proposição 2.5, $[v_1, \ldots, v_n]$ é um subespaço vetorial de V. ■

Exemplo 2.17. *1. Determinemos o subespaço gerado por* $(1,3,2)$ *e* $(-1,0,1)$ *em* $\mathbb{R}^3$. *De fato, queremos determinar*

$$\begin{aligned} (x,y,z) &= k_1(1,3,2) + k_2(-1,0,1) \\ &= (k_1 - k_2,\ 3k_1,\ 2k_1 + k_2), \end{aligned}$$

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

Observe que renomeando $k_1 - k_2 = x,\ 3k_1 = y$ *e* $2k_1 + k_2 = z$ *e resolvendo* k_1 *e* k_2 *em função de* x, y *e* z, *obtemos da segunda igualdade que* $k_1 = \frac{y}{3}$. *Substituindo esse valor na terceira equação, obtemos* $k_2 = z - 2\frac{y}{3}$. *Desse modo, pela primeira equação concluímos que* $x - y + z = 0$, *ou seja,*

$$[(1,3,2),(-1,0,1)] = \{(k_1 - k_2, 3k_1, 2k_1 + k_2);\ k_1, k_2 \in \mathbb{R}\} = \{(x,y,z) \in \mathbb{R}^3;\ x - y + z = 0\}.$$

Geometricamente temos que o subespaço gerado por $(1,3,2)$ *e* $(-1,0,1)$ *é o plano em* $\mathbb{R}^3$ *dado pela equação* $x - y + z = 0$.

2. O subespaço de $\mathcal{M}_{2\times 2}(\mathbb{R})$ *gerado pelas matrizes* $\begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix}$ *e* $\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}$ *é dado por:*

$$\begin{aligned} \left[\begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix} \right] &= \left\{ a \begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix};\ a, b \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} -a & 3a + 2b \\ -2b & a \end{pmatrix};\ a, b \in \mathbb{R} \right\}. \end{aligned}$$

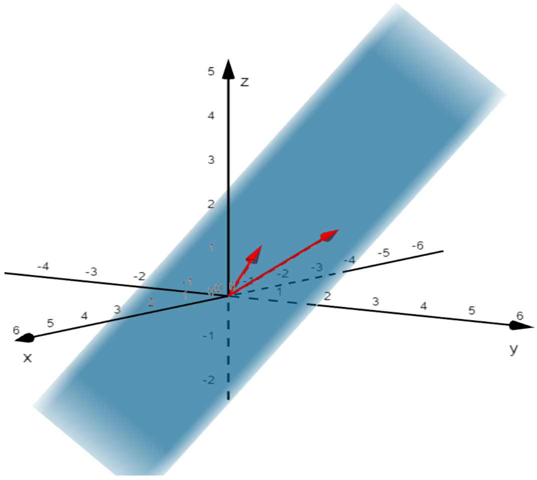

Figura 2.5: Plano de equação $x - y + z = 0$

A definição que vamos apresentar agora vai ser essencial para definirmos o conceito de base de um espaço vetorial.

Definição 2.18. *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 2.19.

1. *Um conjunto com dois vetores $\{u, v\}$ de um espaço vetorial V é LI se, e somente se, u e v não são múltiplos.*

2. *O conjunto $\left\{(1,2,0,4), (-1,5,\frac{1}{2},6), (\frac{7}{2},-7,-1,-6)\right\}$ é um subconjunto linearmente dependente de $\mathbb{R}^4$, pois*

$$\left(\frac{7}{2}, -7, -1, -6\right) = \frac{3}{2}(1,2,0,4) - 2\left(-1, 5, \frac{1}{2}, 6\right).$$

Proposição 2.20. *Seja V um espaço vetorial sobre $\mathbb{K}$.*

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

2. *Todo conjunto que contém um subconjunto 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 $0 = 0v_1 + \ldots + 0v_n$.

2. Claramente essa afirmação é verdadeira, pois se S é um conjunto que possui um subconjunto S' que é LD, então existe um vetor em S' (em particular em S) que é uma combinação linear de outros vetores de S' (em particular de S).

3. Seja S um conjunto LI de V. Como nenhum vetor de S é uma combinação linear dos demais, em particular não o é para qualquer elemento de um subconjunto de S.

4. A ida segue de (3). Reciprocamente, suponha que S seja LD. Logo existe algum vetor u de S que é uma combinação linear finita de certos vetores $v_1, \ldots, v_n$ de S. Logo, o subconjunto $\{u, v_1, \ldots, v_n\}$ 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 2.21. *Um conjunto $\{v_1, \ldots, v_n\}$ de um espaço vetorial V é LI se, e somente se, para toda combinação linear*

$$a_1 v_1 + \ldots + a_n v_n = 0$$

com $a_i \in \mathbb{K}$ implicar que $a_1 = \ldots = a_n = 0$.

Demonstração: Vamos provar os dois casos por contra-positiva. Com efeito, suponha que exista um escalar $a_i \in \mathbb{K}$ não nulo satisfazendo $a_1 v_1 + \ldots + a_n v_n = 0$. 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 $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, e portanto o conjunto $\{v_1, \ldots, v_n\}$ é LD.

Por outro lado, suponha que $\{v_1, \ldots, v_n\}$ é LD. Assim existe pelo menos um v_j que é combinação linear dos demais, ou seja, existem escalares $a_i \in \mathbb{K}$ com $i = 1, \ldots, n,\ i \neq j$ tais que

$$\begin{aligned} & v_j = a_1 v_1 + \ldots + a_{j-1} v_{j-1} + a_{j+1} v_{j+1} + \ldots + a_n v_n \\ \Rightarrow \quad & a_1 v_1 + \ldots + a_{j-1} v_{j-1} - v_j + a_{j+1} v_{j+1} + \ldots + a_n v_n = 0, \end{aligned}$$

ou seja, existe uma combinação linear nula com pelo menos um escalar $a_j = -1$ não nulo.

■

Exemplo 2.22. *1. O conjunto de matrizes em $\mathcal{M}_{2\times 3}(\mathbb{R})$ abaixo é linearmente independente:*

$$\gamma = \left\{ \begin{pmatrix} 1 & -1 & 0 \\ 4 & \sqrt{2} & 4 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 2 \\ -1 & -2 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 6 & 0 & 3 \end{pmatrix} \right\}.$$

De fato, considere $a, b, c \in \mathbb{R}$ tais que

$$a \cdot \begin{pmatrix} 1 & -1 & 0 \\ 4 & \sqrt{2} & 4 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 0 & 2 \\ -1 & -2 & 5 \end{pmatrix} + c \cdot \begin{pmatrix} 1 & 1 & 1 \\ 6 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} a+c & -a+c & 2b+c \\ 4a-b+6c & \sqrt{2}a-2b & 4a+5b+3c \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

A igualdade acima é válida se, e somente se,

$$\begin{cases} a+c=0 \\ -a+c=0 \\ 2b+c=0 \end{cases} \quad e \quad \begin{cases} 4a-b+6c=0 \\ \sqrt{2}a-2b=0 \\ 4a+5b+3c=0. \end{cases}$$

Pelas duas primeiras equações do primeiro sistema temos que $a = c = 0$ que implica que $2b = 0$, ou seja, $b = 0$. No segundo sistema temos que a matriz dos coeficientes do sistema

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

possui determinante $det(A) = 24 + 33\sqrt{2}$ que é diferente de zero. Assim pela regra de Crammer a única solução deste segundo sistema é a trivial.

Portanto, γ é um conjunto linearmente independente de $\mathcal{M}_{2\times 3}(\mathbb{R})$.

2. *O conjunto $\{1, x, \ldots, x^n\}$ em $\mathcal{P}_n$ é LI. De fato, considere uma combinação linear qualquer*

$$a_0 \cdot 1 + a_1 x + \ldots + a_n x^n = 0 = 0 + 0x + \ldots + 0x^n$$

e por igualdade de polinômios sua única solução é $a_0 = \ldots = a_n = 0$.

3. *É possível provar que o conjunto $\{1, x, \ldots, x^n, \ldots\}$ é um conjunto LI de $\mathcal{P}$.*

Proposição 2.23. *Sejam $v_1, \ldots, v_n$ vetores de $\mathbb{K}^n$ com $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. O conjunto $\{v_1, \ldots, v_n\}$ com $v_i = (a_{1i}, a_{2i}, \ldots, a_{ni})$ é LI se, e somente se, o determinante da matriz*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

cujas colunas são as coordenadas dos vetores v_i's é não nulo.

Demonstração: Sejam $v_1, \ldots, v_n$ vetores LI, onde $v_i = (a_{1i}, a_{2i}, \ldots, a_{ni})$. Considere uma combinação linear nula de $v_1, \ldots, v_n$, ou seja, sejam $x_1, \ldots, x_n \in \mathbb{K}$ tais que

$$x_1(a_{11}, a_{21}, \ldots, a_{n1}) + x_2(a_{12}, a_{22}, \ldots, a_{n2}) + \ldots + x_n(a_{1n}, a_{2n}, \ldots, a_{nn}) = (0, 0, \ldots, 0).$$

Efetuando os cálculos, obtemos o sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= 0, \end{cases}$$

o qual é um sistema homogêneo com n equações e n variáveis. Segue da Regra de Cramer que esse sistema admite solução única se, e somente se, o determinante da matriz dos coeficientes, que é exatamente a matriz A, é diferente de zero. ■

Definição 2.24. *Sejam V um espaço vetorial sobre um corpo $\mathbb{K}$ e $\beta \subset V$ um conjunto não vazio. Diz-se que β é uma **base** de V,*

1. *se β é linearmente independente;*

2. *se β gera V, ou seja, o subespaço gerado por β coincide com V, ou ainda, $[\beta] = V$.*

Observe que se $\beta = \{v_1, \ldots, v_n\}$ é uma base de um espaço vetorial V, então dado um elemento qualquer de $v \in V$ ele se escreve de maneira única como combinação linear de $v_1, \ldots, v_n$. De fato, suponha que existam $k_1, \ldots, k_n \in \mathbb{K}$ e $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tais que $v = k_1v_1 + \cdots + k_nv_n$ e $v = \lambda_1v_1 + \cdots + \lambda_nv_n$. Logo,

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

O que implica que $(k_1 - \lambda_1)v_1 + \ldots + (k_n - \lambda_n)v_n = 0$. Como β é LI, segue que $k_i - \lambda_i = 0$ para todo $i = 1, \ldots, n$. E portanto, $k_i = \lambda_i$.

Exemplo 2.25. *1. Provemos que o conjunto $\beta = \{(0, 4), (2, -1)\}$ é uma base de $\mathbb{R}^2$. Como*

$$\begin{vmatrix} 0 & 2 \\ 4 & -1 \end{vmatrix} = -8 \neq 0$$

segue que β é L.I. Falta provarmos que β gera $\mathbb{R}^2$, ou seja, qualquer vetor de $\mathbb{R}^2$ é combinação linear dos elementos de β. Seja $(a,b) \in \mathbb{R}^2$ e considere $x, y \in \mathbb{R}$ tais que

$$(a,b) = x(0,4) + y(2,-1).$$

Logo, $a = 2y$ e $b = 4x - y$, ou seja, $y = \frac{a}{2}$ e $x = \frac{b}{4} + \frac{y}{4} = \frac{2b+a}{8}$. Ou ainda

$$x(0,4) + y(2,-1) = \frac{2b+a}{8}(0,4) + \frac{a}{2}(2,-1) = (a,b).$$

Portanto, β gera $\mathbb{R}^2$, donde concluímos que β é uma base de $\mathbb{R}^2$.

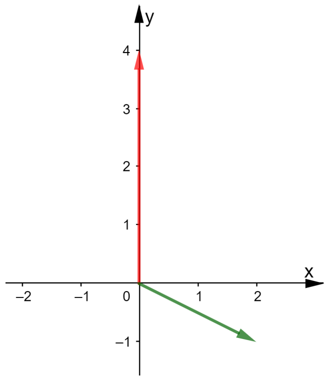

Figura 2.6: Base $\beta = \{(0,4),(2,-1)\}$ de $\mathbb{R}^2$

2. *O conjunto $\alpha = \{(1,2),(-1,0),(2,3)\}$ gera $\mathbb{R}^2$, entretanto não é LI e, portanto, não é uma base de $\mathbb{R}^2$.*

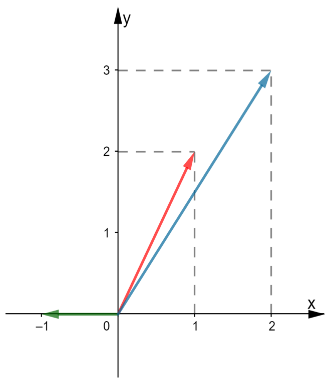

Figura 2.7: $\alpha = \{(1,2),(-1,0),(2,3)\}$ gera $\mathbb{R}^2$, mas não é base

De fato, dado qualquer vetor $(x,y) \in \mathbb{R}^2$ temos que (x,y) é combinação linear dos vetores de α, já que dados $a, b, c \in \mathbb{R}$ temos que

$$(x,y) = a(1,2) + b(-1,0) + c(2,3),$$

se, e somente se,

$$\begin{cases} a - b + 2c = x \\ 2a + 3c = y. \end{cases}$$

Deste modo, $b = a + 2c - x$ *e* $a = \frac{1}{2}(y - 3c)$*. Substituindo o valor de* a *em* b *obtemos que* $b = \frac{1}{2}(y + c) - x$*, ou seja, dado* (x, y) *qualquer, escolha qualquer valor para* c*, tome* $a = \frac{1}{2}(y - 3c)$ *e* $b = \frac{1}{2}(y + c) - x$ *e obtemos*

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

Portanto, α *gera* $\mathbb{R}^2$*. No entanto, se*

$$a(1, 2) + b(-1, 0) + c(2, 3) = (0, 0),$$

então como fizemos acima, tomando qualquer $c \in \mathbb{R}$*, obtemos* $a = \frac{1}{2}(-3c)$ *e* $b = \frac{c}{2}$ *obtemos uma combinação linear nula dos vetores de* α*, em que* a, b *e* c *não são necessariamente nulos, ou seja,* α *não é LI.*

As bases descritas nos itens 1,2 e 3 abaixo são chamadas de **bases canônicas**.

Exemplo 2.26. *1. O conjunto* $\beta = \{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$ *é uma base* $\mathbb{R}^n$ *como espaço vetorial sobre* $\mathbb{R}$ *e também é uma base de* $\mathbb{C}^n$ *como espaço vetorial sobre* $\mathbb{C}$*.*

De fato, denotemos por $\mathbb{K}$ *o corpo* $\mathbb{R}$ *ou* $\mathbb{C}$*. Claramente* β *é um conjunto LI de* $\mathbb{K}^n$*, pois pela Proposição 2.23, o determinante da matriz cujas colunas são os vetores do conjunto* β *satisfaz:*

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1 \neq 0.$$

E β *gera* $\mathbb{K}^n$ *uma vez que dado qualquer vetor* $(x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ *temos que*

$$(x_1, x_2, \ldots, x_n) = x_1(1, 0, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \ldots + x_n(0, \ldots, 0, 1),$$

com $x_i \in \mathbb{K}$ *para* $i = 1, \ldots, n$*, ou seja, é uma combinação linear dos vetores de* β*. Portanto,* β *é uma base de* $\mathbb{K}^n$*.*

2. O conjunto $\gamma = \{x^n, x^{n-1}, \ldots, x, 1\}$ *é uma base do espaço* $\mathcal{P}_n$*.*

Verificamos anteriormente que γ *é um conjunto LI de* $\mathcal{P}_n$*. Mais ainda,* γ *gera* $\mathcal{P}_n$*, pois* *dado um polinômio qualquer $p = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ de $\mathcal{P}_n$, podemos escrevê-lo como*

$$p = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \ldots + a_1 \cdot x + a_0 \cdot 1.$$

3. *Seja $\delta = \{A_{ij} = (a_{kl});$ com $a_{ij} = 1$ e $a_{kl} = 0$ caso contrário, $i = 1, \ldots, n, j = 1, \ldots, m\}$ é uma base de $\mathcal{M}_{n\times m}(\mathbb{R})$.*

   Por exemplo, o conjunto

$$\delta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

   é uma base para $\mathcal{M}_{3\times 2}(\mathbb{R})$.

Proposição 2.27. *Sejam $v_1, \ldots, v_n$ vetores não nulos que geram um $\mathbb{K}$-espaço vetorial V. Então, dentre estes vetores podemos extrair uma base de V.*

Demonstração: Seja $\beta = \{v_1, \ldots, v_n\}$ tal que $V = [v_1, \ldots, v_n]$. Se β é LI não temos nada a fazer. Senão, existe pelo menos um vetor de β que é combinação linear dos demais. Sem perda de generalidade, suponha que v_n seja tal elemento. Logo,

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

Mais ainda V é gerado por $v_1, \ldots, v_{n-1}$. Seja $\beta_1 = \{v_1, \ldots, v_{n-1}\}$. Se β_1 for LI temos o resultado. Senão, existe um elemento de β_1 que é combinação linear dos demais. Retirando este elemento obtemos um conjunto que ainda gera V. Continuando este processo chegaremos, após um número finito de passos, a um subconjunto de β formado por vetores LI e que geram V, ou seja, a uma base. ■

Teorema 2.28. *Se um espaço vetorial V é gerado por n vetores, então qualquer conjunto com mais de n vetores em V é LD.*

Demonstração: Sejam $v_1, \ldots, v_n$ vetores tais que $V = [v_1, \ldots, v_n]$. Pela proposição anterior, podemos extrair deste conjunto uma base para V. Seja $\beta = \{v_1, \ldots, v_r\}$ uma tal base com $r \leq n$. Sejam os vetores $w_1, \ldots, w_m$ em V com $m > r$. Como $v_1, \ldots, v_r$ geram V temos que

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

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

$$x_1w_1 + \ldots + x_mw_m = \left(\sum_{j=1}^{m} x_j\alpha_{j1}\right) v_1 + \ldots + \left(\sum_{j=1}^{m} x_j\alpha_{jr}\right) v_r = 0.$$

Como $\{v_1, \ldots, v_r\}$ é LI obtemos o sistema

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

Temos assim um sistema linear homogêneo com r equações e m variáveis e como $r < m$, o sistema admite solução não trivial. Logo, $w_1, \ldots, w_m$ são LD. ■

Corolário 2.29. *Se um espaço vetorial V admite uma base $\beta = \{v_1, \ldots, v_n\}$ com n vetores, então qualquer outra base de V possui também n elementos.*

Demonstração: De fato, seja $\beta' = \{w_1, \ldots, w_m\}$ uma outra base de V. Como β gera V e β' é LI, segue da proposição anterior que $m \leq n$. Por outro lado, β' gera V e β é LI, assim temos que $n \leq m$. Portanto, $n = m$. ■

Dizemos que o espaço vetorial V possui dimensão finita, quando admite uma base com um número finito de elementos. Como pelo corolário acima, todas as bases de V possuem o mesmo número de elementos, segue que se $\beta = \{v_1, \ldots, v_n\}$ é uma base de V, então a **dimensão** de V como espaço vetorial sobre $\mathbb{K}$, denotado por $dim_{\mathbb{K}}V$, é definido como o número de elementos de uma base de V, ou seja,

$$dim_{\mathbb{K}}V = n.$$

Podemos definir, a dimensão do espaço vetorial nulo como sendo zero.

Exemplo 2.30. *Nos quatro primeiros exemplos abaixo, apresentamos anteriormente uma base para estes espaços vetoriais, portanto já podemos determinar suas dimensões.*

1. *A dimensão de $\mathbb{R}^n$ como espaço vetorial sobre $\mathbb{R}$ e a dimensão de $\mathbb{C}^n$ como espaço vetorial sobre $\mathbb{C}$ é exatamente*

$$dim_{\mathbb{R}}\mathbb{R}^n = n \qquad e \qquad dim_{\mathbb{C}}\mathbb{C}^n = n;$$

Em ambos os casos vimos que $\beta = \{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$ *é uma base de* $\mathbb{K}^n$ *como espaço vetorial sobre* $\mathbb{K}$*, com* $\mathbb{K} = \mathbb{R}$ *ou* $\mathbb{C}$*.*

2. *A dimensão do espaço vetorial* $\mathcal{M}_{n\times m}(\mathbb{R})$ *é dada por*

$$dim_{\mathbb{R}}\mathcal{M}_{n\times m}(\mathbb{R}) = n \cdot m;$$

3. *A dimensão de* $\mathcal{P}_n$ *como espaço vetorial sobre* $\mathbb{R}$ *é* $n+1$*;*

4. *Neste exemplo vamos ver que a dimensão de um espaço vetorial depende do corpo* $\mathbb{K}$*. Provemos que a dimensão do espaço vetorial* $\mathbb{C}$ *como espaço vetorial sobre* $\mathbb{R}$ *é* 2*;*

   Vimos anteriormente que $dim_{\mathbb{C}}\mathbb{C} = 1$ *e a base canônica para tal espaço é* $\gamma = \{1\}$*. Provemos que se enxergarmos* $\mathbb{C}$ *como espaço vetorial sobre* $\mathbb{R}$*, então* $\beta = \{1, i\}$ *é uma base de* $\mathbb{C}$*. De fato,* β *é LI, pois dados* $k_1, k_2 \in \mathbb{R}$ *tais que*

$$k_1 \cdot 1 + k_2 \cdot i = 0 \quad \Rightarrow \quad k_1 \cdot 1 + k_2 \cdot i = 0 + 0i$$

   donde segue da igualdade de números complexos que $k_1 = 0$ *e* $k_2 = 0$*. Portanto,* β *é LI. Mais ainda,* β *gera* $\mathbb{C}$ *já que dado* $z \in \mathbb{C}$ *com* $z = a + bi$ *e* $a, b \in \mathbb{R}$

$$z = a + bi = a \cdot 1 + b \cdot i.$$

   Portanto, β *é uma base de* $\mathbb{C}$ *como espaço vetorial sobre* $\mathbb{R}$*. Logo* $dim_{\mathbb{R}}\mathbb{C} = 2$*.*

5. *Mais geralmente, a dimensão de* $\mathbb{C}^n$ *como espaço vetorial sobre* $\mathbb{R}$ *é* $2n$*.*

   Basta verificar que

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

   é uma base (canônica) deste espaço vetorial visto com escalares em $\mathbb{R}$*.*

Teorema 2.31. *Qualquer conjunto de vetores LI contido em um* $\mathbb{K}$*-espaço vetorial* V*, de dimensão finita, pode ser completado de modo a formar uma base de* V*.*

Demonstração: Suponha que $dim_{\mathbb{K}}V = n$ e sejam $\{v_1,\ldots,v_r\}$ um conjunto LI contido em V. Pelo Teorema 2.28, $r \leq n$. Se $V = [v_1,\ldots,v_r]$, então $v_1,\ldots,v_r$ seria uma base de V e, neste caso, $r = n$. Senão, existe $v_{r+1} \in V$ tal que $v_{r+1} \notin [v_1,\ldots,v_r]$, ou seja, v_{r+1} não é combinação linear de $v_1,\ldots,v_r$. Logo $\{v_1,\ldots,v_r,v_{r+1}\}$ é LI. Se $[v_1,\ldots,v_r,v_{r+1}] = V$ temos a base procurada. Caso contrário, existe $v_{r+2} \notin [v_1,\ldots,v_r,v_{r+1}]$ e assim $\{v_1,\ldots,v_{r+1},v_{r+2}\}$ é LI.

Como não podemos ter mais que n vetores linearmente independentes, pois a dimensão de V é n, este processo termina após um número finito de passos. ■

Corolário 2.32. *Se $dim_{\mathbb{K}}V = n$, então um conjunto $\beta = \{v_1, \ldots, v_n\} \subset V$ é LI se, e somente se, β gera V.*

Demonstração: Seja β um conjunto LI de V. Se β não gera V, pelo teorema anterior podemos completar β de modo a termos uma base de V. Absurdo, pois $dim_{\mathbb{K}}V = n$.

Por outro lado, suponha que β gera V. Donde segue que se β é LD, pela Proposição 2.27, podemos extrair de β uma base de V, o que é um absurdo, uma vez que a dimensão de V é n. ■

Observe que se V é um espaço vetorial de dimensão finita, então qualquer subespaço S de V satisfaz $dim_{\mathbb{K}}S \leq dim_{\mathbb{K}}V$.

Teorema 2.33. *Se S_1, S_2 são subespaços de um 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: Sejam $\beta_1 = \{u_1, \ldots, u_n\}$ e $\beta_2 = \{v_1, \ldots, v_m\}$ bases de S_1 e S_2, respectivamente. Seja $\beta = \beta_1 \cup \beta_2$ e observe que β gera $S_1 + S_2$. ■

Exercício 2.34. *Sejam V um $\mathbb{K}$ espaço vetorial de dimensão n. Se S é um subespaço de V satisfazendo que $dim_{\mathbb{K}}S = n$, prove que $S = V$.*

Exemplo 2.35. *1. Sejam $S_1 = \{(x, 0, 0);\ x \in \mathbb{R}\}$ e $S_2 = \{(x, y, z) \in \mathbb{R}^3;\ y = x + z\}$ subespaços de $\mathbb{R}^3$. Geometricamente, S_1 é o eixo x e S_2 descreve um plano passando pela origem.*

Veja que

$$S_1 = \{x(1, 0, 0);\ x \in \mathbb{R}\} \quad e$$

$$S_2 = \{(x, x + z, z);\ z \in \mathbb{R}\} = \{x(1, 1, 0) + z(0, 1, 1);\ x, z \in \mathbb{R}\}.$$

Além disso, $\beta_1 = \{(1, 0, 0)\}$ é base de S_1, pois β_1 é naturalmente LI e o vetor $(1, 0, 0)$ gera S_1. Analogamente, $\beta_2 = \{(1, 1, 0), (0, 1, 1)\}$ gera S_2 pois seus vetores são escritos como combinação linear dos vetores de β_2. Mais ainda, β_2 é LI, pois seus vetores não são múltiplos. Logo, β_2 é base de S_2. Qualquer que seja $v \in S_1 + S_2$, temos que $v = v_1 + v_2$

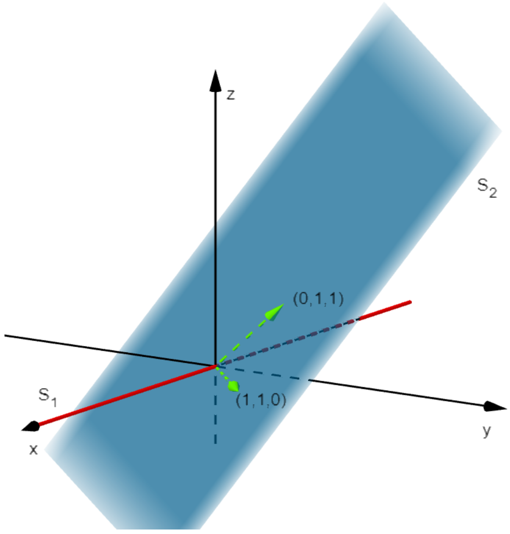

Figura 2.8: Subespaços $S_1 = \{(x, 0, 0);\ x \in \mathbb{R}\}$ e $S_2 = \{(x, y, z) \in \mathbb{R}^3;\ y = x + z\}$

com $v_i \in S_i, i = 1, 2$. *Pelo teorema anterior* $\beta = \beta_1 \cup \beta_2 = \{(1, 0, 0), (1, 1, 0), (0, 1, 1)\}$ *gera* $S_1 + S_2$. *Mais ainda,* β *é LI, pois*

$$\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0.$$

Logo, β *é uma base de* $S_1 + S_2$. *Todavia* $S_1 + S_2$ *é um subespaço de* $\mathbb{R}^3$ *de dimensão* 3, *ou seja,* $S_1 + S_2 = \mathbb{R}^3$. *Pelo teorema acima,*

$$dim_{\mathbb{K}}(S_1 \cap S_2) = dim_{\mathbb{K}} S_1 + dim_{\mathbb{K}} S_2 - dim_{\mathbb{K}}(S_1 + S_2) = 2 + 1 - 3 = 0.$$

Portanto, $S_1 \cap S_2 = \{(0, 0, 0)\}$. *Donde segue que* $\mathbb{R}^3 = S_1 \oplus S_2$.

2. *Sejam*

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

subespaços de $\mathcal{M}_{2\times 2}(\mathbb{R})$. *Verifique que* $\mathcal{M}_{2\times 2}(\mathbb{R}) = W_1 \oplus W_2$.

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* $\mathbb{R}^4$. *Observe que*

$$S_1 = \{(y + t, y, z, t);\ y, z, t \in \mathbb{R}\} = \{y(1, 1, 0, 0) + z(0, 0, 1, 0) + t(1, 0, 0, 1);\ y, z, t \in \mathbb{R}\}.$$

É fácil ver que $\beta_1 = \{(1, 1, 0, 0), (0, 0, 1, 0), (1, 0, 0, 1)\}$ *e* $\beta_2 = \{(0, 0, 1, 0)\}$ *são bases de* S_1

e S_2, respectivamente. Observe que $v = v_1 + v_2 \in S_1 + S_2$ se, e somente se,

$$v = (x, y, z, t) = k_1(1, 1, 0, 0) + k_2(0, 0, 1, 0) + k_3(1, 0, 0, 1) + k_4(0, 0, 1, 0).$$

Portanto, $x = y+t$ o que implica $y+t-x = 0$, logo $S_1+S_2 = \{(x, y, z, t) \in \mathbb{R}^4; x-y-t = 0\} = S_1$, logo $\beta = \beta_1 \cup \beta_2 = \beta_1$ é uma base de $S_1 + S_2$.

Como $S_2 \subset S_1$ segue que $S_1 \cap S_2 = S_2$.

4. *Sejam S_1 o subespaço das matrizes triangulares superiores e S_2 o das matrizes triangulares inferiores. Determine $S_1 \cap S_2$.*

Exemplo 2.36. *Sejam n um inteiro positivo e V o espaço das $n \times n$ matrizes sobre $\mathbb{K}$. Sejam*

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

os subespaços das matrizes simétricas e antissimétricas, respectivamente. Prove que

$$V = W_1 \oplus W_2.$$

Dado $A \in V$ qualquer, A se expressa de maneira única como $A = A_1 + A_2$ com

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

2.3 Mudança de Base

Muitas vezes ao estudarmos um problema em um espaço vetorial consideramos a base canônica, por comodidade ou por simplificar alguns conceitos. Todavia, essa base nem sempre é a melhor opção. Pode ser conveniente considerar uma base diferente da usual, para simplificar um determinado problema ou podemos descrever certas equações, como de curvas ou superfícies, de modo mais simples. Para tanto, vamos definir uma matriz, denominada matriz mudança de base.

Antes porém vamos introduzir uma nova notação: vimos anteriormente que dado uma base $\beta = \{v_1, \ldots, v_n\}$ de um espaço vetorial V sobre um corpo $\mathbb{K}$, qualquer vetor $u \in V$ se escreve de **maneira única** como combinação linear dos vetores da base β, ou seja, existem únicos escalares $k_1, \ldots, k_n \in \mathbb{K}$ tais que

$$u = k_1v_1 + k_2v_2 + \ldots + k_nv_n.$$

Deste modo se pedirmos que β seja uma base ordenada, temos que os escalares $k_1, k_2, \ldots, k_n$ caracterizam completamente o vetor u. Com isso vamos denotar por $[u]_\beta$ a matriz coluna dada pelos escalares que expressam o vetor u na base β, ou seja,

$$[u]_\beta = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.$$

Dizemos que essa é a matriz das coordenadas (ou apenas as coordenadas) do vetor u na base β.

Exemplo 2.37. *Considere $V = \mathbb{R}^2$ com as seguintes bases ordenadas $\beta_1 = \{(1,0),(0,1)\}$, $\beta_2 = \{(-1,4),(2,1)\}$ e $\beta_3 = \{(0,-1),(3,0)\}$. Encontre as coordenadas do vetor $u = (3,5)$ nas respectivas bases dadas acima.*

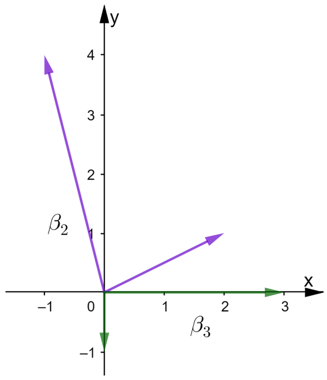

Figura 2.9: Bases $\beta_2 = \{(-1,4),(2,1)\}$ e $\beta_3 = \{(0,-1),(3,0)\}$

Com efeito, sabemos que u se escreve de maneira única em qualquer base de $\mathbb{R}^2$. Assim na base canônica $u = (3,5) = 3(1,0) + 5(0,1)$. Determinemos suas coordenadas nas bases β_2 e β_3.

$$\begin{aligned} (3,5) &= k_1(-1,4) + k_2(2,1) \\ &= (-k_1 + 2k_2,\ 4k_1 + k_2) \end{aligned}$$

Temos assim que resolver

$$\begin{cases} -k_1 + 2k_2 = 3 \\ 4k_1 + k_2 = 5 \end{cases}$$

É fácil ver que a solução deste sistema é $k_1 = \frac{7}{9}$ e $k_2 = \frac{17}{9}$, ou seja,

$$u = (3,5) = \frac{7}{9}(-1,4) + \frac{17}{9}(2,1).$$

De modo análogo $(3,5) = k_1(0,-1) + k_2(3,0) = (3k_2,\ -k_1)$. Donde concluímos que $k_1 = -5$

e $k_2 = 1$. *Portanto*

$$[(3,5)]_{\beta_1} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad [(3,5)]_{\beta_2} = \begin{pmatrix} \frac{7}{9} \\ \frac{17}{9} \end{pmatrix} \quad e \quad [(3,5)]_{\beta_3} = \begin{pmatrix} -5 \\ 1 \end{pmatrix}.$$

Determinemos agora uma matriz que representa a mudança de uma base para outra.

Sejam $\beta = \{u_1, \ldots, u_n\}$ e $\beta' = \{w_1, \ldots, w_n\}$ bases ordenadas de um mesmo espaço vetorial V. Dado um vetor $v \in V$ podemos escrevê-lo na base β por:

$$v = \lambda_1 u_1 + \ldots + \lambda_n u_n.$$

Como β' é base de V, podemos escrever os vetores u_i como combinação linear dos vetores w_i, $i = 1, \ldots, n$.

$$\begin{cases} u_1 = a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ u_2 = a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n \\ \qquad \vdots \\ u_n = a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n. \end{cases} \tag{2.1}$$

Logo,

$$\begin{aligned} v &= \lambda_1 u_1 + \lambda_2 u_2 \ldots + \lambda_n u_n \\ &= \lambda_1(a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n) + \lambda_2(a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n) + \ldots + \\ &\quad \lambda_n(a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n) \\ &= (a_{11}\lambda_1 + a_{12}\lambda_2 + \ldots + a_{1n}\lambda_n)w_1 + (a_{21}\lambda_1 + a_{22}\lambda_2 + \ldots + a_{2n}\lambda_n)w_2 + \ldots + \\ &\quad (a_{n1}\lambda_1 + a_{n2}\lambda_2 + \ldots + a_{nn}\lambda_n)w_n. \end{aligned}$$

Isso significa que as coordenadas de v na base β' é dada por:

$$[v]_{\beta'} = \begin{pmatrix} a_{11}\lambda_1 + a_{12}\lambda_2 + \ldots + a_{1n}\lambda_n \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \ldots + a_{2n}\lambda_n \\ \vdots \\ a_{n1}\lambda_1 + a_{n2}\lambda_2 + \ldots + a_{nn}\lambda_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Denotamos por $[I]^{\beta}_{\beta'}$ a matriz (a_{ij}), chamada a **matriz mudança da base** β para a base β', o qual satisfaz

$$[v]_{\beta'} = [I]^{\beta}_{\beta'} \cdot [v]_{\beta}.$$

Com isso provamos o seguinte teorema.

Teorema 2.38. *Seja V um espaço vetorial sobre $\mathbb{K}$, de dimensão finita n. Sejam β e β' duas bases ordenadas de V, então existe uma única matriz invertível P de ordem n com entradas em $\mathbb{K}$ tal que*

$$[v]_{\beta'} = P \cdot [v]_{\beta} \quad e \quad [v]_{\beta} = P^{-1} \cdot [v]_{\beta'},$$

para todo $v \in V$.

Demonstração: Já provamos que $[v]_{\beta'} = [I]^{\beta}_{\beta'} \cdot [v]_{\beta}$ e que $[v]_{\beta} = [I]^{\beta'}_{\beta} \cdot [v]_{\beta'}$. Logo,

$$[v]_{\beta'} = [I]^{\beta}_{\beta'} \cdot [v]_{\beta} = [I]^{\beta}_{\beta'} \cdot [I]^{\beta'}_{\beta} \cdot [v]_{\beta'}.$$

O qual ocorre se, e somente se, $[I]^{\beta}_{\beta'}[I]^{\beta'}_{\beta} = I$, onde I é a matriz identidade. Portanto,

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

■

Exemplo 2.39. *Sejam $\beta = \{(1,-2),(3,-4)\}$ e $\beta' = \{(1,3),(2,1)\}$ bases de $\mathbb{R}^2$. Determinemos $[I]^{\beta}_{\beta'}$. De fato, escrevendo*

$$(1,-2) = a(1,3) + b(2,1) \qquad e \qquad (3,-4) = c(1,3) + d(2,1),$$

basta resolvermos os sistemas

$$\begin{cases} a + 2b = 1 \\ 3a + b = -2 \end{cases} \qquad e \qquad \begin{cases} c + 2d = 3 \\ 3c + d = -4. \end{cases}$$

Veja que a soluções procuradas são $a = -1, b = 1, c = -\frac{11}{5}$ e $d = \frac{13}{5}$. Portanto, a matriz de mudança da base β para a base β' é

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

CAPÍTULO 3

TRANSFORMAÇÕES LINEARES

Neste capítulo vamos explorar alguns dos conceitos mais interessantes da álgebra linear, o de transformações lineares. O conceito de autovalores e autovetores de uma transformação linear é explorado em diversas áreas da matemática, física, engenharia e áreas correlatas. Vamos considerar neste capítulo espaços vetoriais de dimensão finita.

Sejam V e W espaços vetoriais sobre um corpo $\mathbb{K}$ de dimensão finita.

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:*

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

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

Exemplo 3.2. *Vamos apresentar alguns exemplos de transformações lineares.*

1. *Seja* $T : \mathbb{R}^2 \to \mathbb{R}^3$ *dada por* $T(x,y) = (2x-y,\ x,\ 4y-x)$.

   Dados $u = (x,y)$ *e* $v = (z,w)$ *em* $\mathbb{R}^2$ *temos que*

$$\begin{aligned} T(u+v) = T(x+z, y+w) &= (2(x+z)-(y+w), x+z, 4(y+w)-(x+z)) \\ &= (2x-y+2z-w, x+z, 4y-x+4w-z) \\ &= (2x-y, x, 4y-x) + (2z-w, z, 4w-z) \\ &= T(x,y) + T(z,w) = T(u) + T(v). \end{aligned}$$

E ainda,

$$\begin{aligned} T(ku) &= T(kx, ky) = (2kx - ky, kx, 4ky - kx) = k(2x - y, x, 4y - x) \\ &= kT(x, y) = kT(u). \end{aligned}$$

Verifique que as aplicações abaixo são transformações lineares.

2. *A aplicação $T : V \to W$ dada por $T(v) = 0_W$, para todo $v \in V$ em que 0_W é o elemento neutro de W, é chamada transformação nula.*

3. *A aplicação $I : V \to V$ dada por $I(v) = v$ é chamada transformação identidade.*

Afirmação 3.3. *Se $T : V \to W$ é uma transformação linear, então $T(0_V) = 0_W$.*

De fato,

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V),$$

donde concluímos que $T(0_V) + 0_W = T(0_V) = T(0_V) + T(0_V)$. Como $T(0_V)$ é um vetor de W, ele possui o seu oposto em W. Desse modo,

$$-T(0_V) + T(0_V) + 0_W = -T(0_V) + T(0_V) + T(0_V),$$

portanto, $T(0_V) = 0_W$. ■

Desse modo, se $T(0_V) \neq 0_W$, então T não é uma transformação linear. Entretanto a recíproca da afirmação acima não é verdadeira. Por exemplo, seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (x^2, y)$. Veja que $T(0, 0) = (0, 0)$, todavia T não é linear, pois dado $k \in \mathbb{R}$

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

Observação 3.4. *Note que $T : V \to W$ é uma transformação linear se, e somente se,*

$$T(ku + v) = kT(u) + T(v),$$

para todo $u, v \in V$ e $k \in \mathbb{K}$. De fato, segue da definição de transformação linear que

$$T(ku + v) = T(ku) + T(v) = kT(u) + T(v)$$

para todo $u, v \in V$ e $k \in \mathbb{K}$. Reciprocamente, quaisquer que sejam $u, v \in V$ e $k \in \mathbb{K}$ temos, por *hipótese, que*

$$T(u+v) = T(1 \cdot u + v) = 1 \cdot T(u) + T(v) = T(u) + T(v).$$

E ainda, $T(ku) = T(ku + 0_V) = kT(u) + T(0_V)$. *Provemos que com essa hipótese, é também possível garantir que* $T(0_V) = 0_W$. *Com efeito,*

$$T(0_V) = T(1 \cdot 0_V + 0_V) = 1 \cdot T(0_V) + T(0_V) = T(0_V) + T(0_V)$$

e, pelo mesmo argumento anterior, $T(0_V) = 0_W$. *Desse modo,*

$$T(ku) = kT(u) + T(0_V) = kT(u) + 0_W = kT(u).$$

Portanto, T *é uma transformação linear.*

Exemplo 3.5. *Algumas transformações* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *recebem nomes especiais, verifique que as aplicações abaixo são transformações lineares.*

1. ***Reflexão em torno do eixo Ox***: $T(x, y) = (x, -y)$.

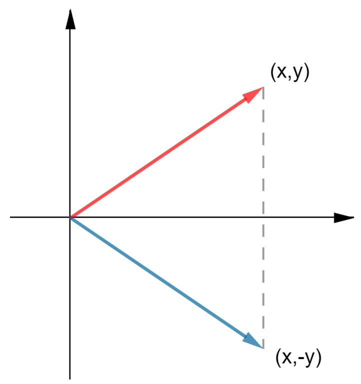

Figura 3.1: Transformação linear $T(x, y) = (x, -y)$

2. ***Reflexão em torno do eixo Oy***: $T(x, y) = (-x, y)$.

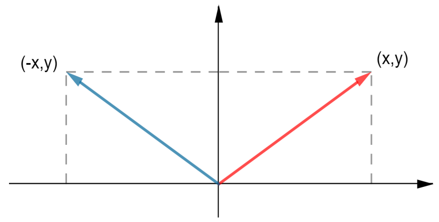

Figura 3.2: Transformação linear $T(x, y) = (-x, y)$

3. ***Reflexão em torno da origem***: $T(x,y) = (-x,-y)$.

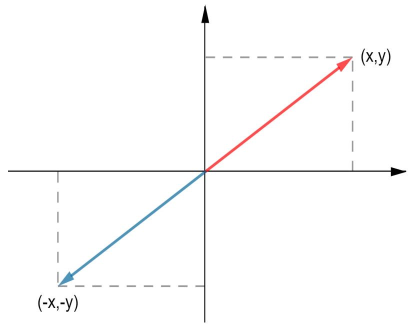

Figura 3.3: Transformação linear $T(x,y) = (-x,-y)$

4. $\begin{cases} \textbf{\textit{Dilatação}}: T(x,y) = k(x,y),\ k \in \mathbb{R}^*,\ |k| > 1; \\ \textbf{\textit{Contração}}: T(x,y) = k(x,y),\ k \in \mathbb{R}^*,\ |k| < 1. \end{cases}$

   Sabemos que $|k| > 1$ *se, e somente se,* $k < -1$ *ou* $k > 1$. *Se* $k > 1$, $T(v)$ *é a dilatação do vetor* $v = (x,y)$ *preservando seu sentido, e se* $k < -1$, *dilatamos o vetor invertendo o sentido do mesmo. De modo similar,* $|k| < 1$ *se, e somente se,* $-1 < k < 1$. *Se* $0 < k < 1$, *então* $T(v)$ *é a contração do vetor* v *preservando seu sentido, enquanto para* $-1 < k < 0$, *contraímos* v *invertendo seu sentido.*

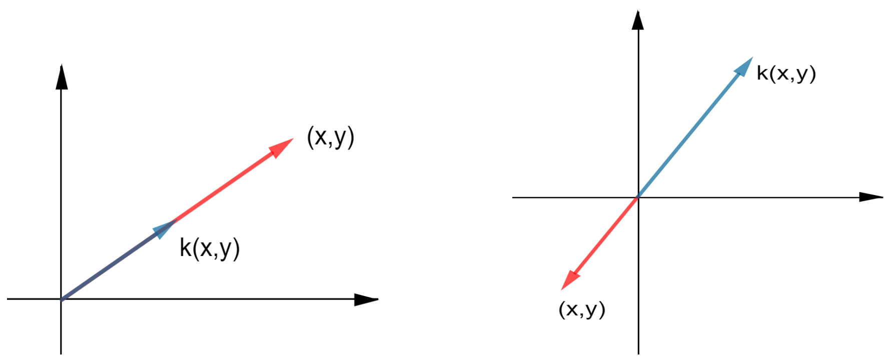

Figura 3.4: Transformação linear $T(x,y) = k(x,y)$

5. *Observe que a* ***translação*** $T(x,y) = (x+a, y+b)$ *com* $(a,b) \neq (0,0)$ *não é uma transformação linear, pois* $T(0,0) = (a,b) \neq (0,0)$.

Veremos que uma transformação linear está completamente determinada se soubermos o que ela faz com os vetores de uma base. Sejam V e W espaços vetoriais sobre $\mathbb{K}$, $T : V \to W$ uma transformação linear e $\beta = \{v_1, \ldots, v_n\}$ uma base de V. Então para todo $v \in V$, dado por $v = k_1v_1 + \ldots + k_nv_n$ com $k_i \in \mathbb{K}$, temos pela linearidade de T que

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

ou seja, se conhecermos a imagem por T dos vetores de uma base, determinamos o valor de T em um vetor arbitrário de V e, portanto, determinamos T.

Teorema 3.6. *Considere V um espaço vetorial sobre um corpo $\mathbb{K}$, de dimensão finita e $\beta = \{v_1, \ldots, v_n\}$ uma base ordenada de V. Sejam W um espaço vetorial sobre $\mathbb{K}$ e $w_1, \ldots, w_n$ vetores arbitrários de W. Então existe exatamente uma única transformação linear $T : V \to W$ tal que $T(v_j) = w_j$ para $j = 1, \ldots, n$.*

Demonstração: Dado $v \in V$ existem únicos escalares $k_1, \ldots, k_n$ tais que

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

Defina $T(v) = k_1 w_1 + \ldots + k_n w_n$.

Primeiramente, T está bem definida, pois para cada $v \in V$, $T(v)$ é um vetor em W. E ainda, cada $v_j \in \beta$ satisfaz $v_j = 0v_1 + \ldots + 0v_{j-1} + 1v_j + 0v_{j+1} + \ldots + 0v_n$ e, por definição,

$$T(v_j) = 0w_1 + \ldots + 0w_{j-1} + 1w_j + 0w_{j+1} + \ldots + 0w_n = w_j,$$

ou seja, $T(v_j) = w_j$ para todo $j = 1, \ldots, n$. Note que T é linear, pois dado $u \in V$ o qual se expressa na base da forma $u = \lambda_1 v_1 + \ldots + \lambda_n v_n$ e $k \in \mathbb{K}$ qualquer

$$kv + u = k(k_1 v_1 + \ldots + k_n v_n) + \lambda_1 v_1 + \ldots + \lambda_n v_n = (kk_1 + \lambda_1)v_1 + \ldots + (kk_n + \lambda_n)v_n.$$

Portanto, por definição de T temos que

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

Por outro lado,

$$kT(v) + T(u) = k(k_1 w_1 + \ldots + k_n w_n) + \lambda_1 w_1 + \ldots + \lambda_n w_n = (kk_1 + \lambda_1)w_1 + \ldots + (kk_n + \lambda_n)w_n.$$

Logo, $T(kv + u) = kT(v) + T(u)$ e, portanto, T é linear.

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

$$U(v) = U(\sum_{i=1}^{n} k_i v_i) = \sum_{i=1}^{n} k_i U(v_i) = \sum_{i=1}^{n} k_i w_i = T(v).$$

Logo $U = T$. ■

Exemplo 3.7. *1. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear tal que $T(2,3) = (-1,5)$ e $T(0,1) = (2,1)$. Vamos observar que pelo resultado anterior, existe uma única transformação linear T satisfazendo as propriedades colocadas.*

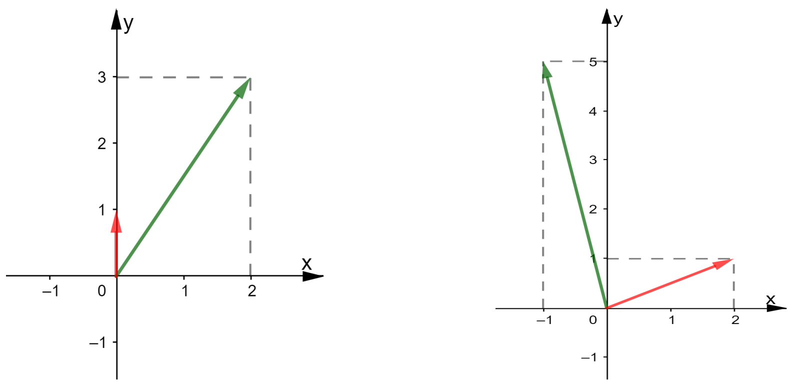

Figura 3.5: Transformação linear satisfazendo $T(2,3) = (-1,5)$ e $T(0,1) = (2,1)$

Como $\{(2,3),(0,1)\}$ é uma base de $\mathbb{R}^2$ temos que todo vetor $v = (x,y) \in \mathbb{R}^2$ se escreve como combinação linear dos vetores da base. Na verdade, dado $(x,y) \in \mathbb{R}^2$ podemos verificar que

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

Aplicando T em ambos os lados da igualdade acima, e como estamos supondo T linear, obtemos

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

Portanto,

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

2. ***Rotação**: Primeiramente, vamos tentar nos convencer que uma rotação em torno da origem, é uma transformação linear. Considere $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma aplicação que dado $v \in \mathbb{R}^2$, então $T(v)$ é o vetor no plano obtido por meio da rotação de v de ângulo θ em torno da origem, no sentido anti-horário. Esse movimento é chamado de movimento rígido do plano, já que preservamos o comprimento de v, mudando apenas sua posição no plano. Veja que a origem é preservada, ou seja, $T(0) = 0$. E mais, dados $u, v \in \mathbb{R}^2$ temos que $u + v$ é o vetor que pode ser representado geometricamente, como o vetor dado pela diago-*

nal do paralelogramo determinado por u *e* v*. Veja que girar de ângulo* θ *os vetores* u, v *e* $u+v$*, é o mesmo que girar o paralelogramo determinado por eles. E portanto, a imagem por* T *é um paralelogramo gerado por* $T(u)$ *e* $T(v)$*, cuja diagonal é* $T(u+v)$*, ou seja,* $T(u+v) = T(u) + T(v)$*. E dado* $k \in \mathbb{R}$ *o vetor* kv *é um múltiplo de* v*. Ao aplicarmos* $T(kv)$*, ou seja, rotacioná-lo de ângulo* θ*, é naturalmente o mesmo que rotacionar o vetor* v*, obtendo* $T(v)$ *e posteriormente multiplicá-lo por* k*, ou seja,* $T(kv) = kT(v)$*. Portanto,* T *é uma transformação linear.*

Para obtermos a transformação de rotação analisemos o que T *faz com uma base de* $\mathbb{R}^2$*. Seja* $\beta = \{e_1, e_2\}$ *a base canônica de* $\mathbb{R}^2$*. Se* T *é uma rotação de ângulo* θ*, com* $0 \leq \theta \leq 2\pi$*, então* T *leva a base* β *nos seguintes vetores:*

$$T(e_1) = (cos\theta, sen\theta) \quad e \quad T(e_2) = (-sen\theta, cos\theta),$$

conforme podemos observar na figura.

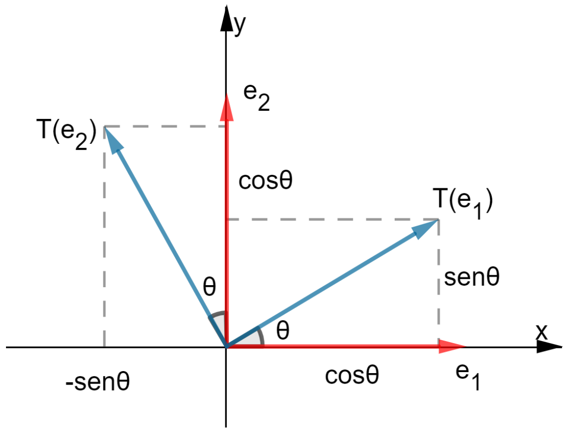

Figura 3.6: Rotação de ângulo θ

Entretanto, qualquer $(x, y) \in \mathbb{R}^2$ *se escreve de modo único da forma*

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

Como T *é linear, temos que*

$$T(x, y) = xT(1, 0) + yT(0, 1) = x(cos\theta, sen\theta) + y(-sen\theta, cos\theta).$$

Portanto,

$$T(x, y) = (xcos\theta - ysen\theta,\ xsen\theta + ycos\theta).$$

- OPERAÇÕES COM TRANSFORMAÇÕES LINEARES

Vamos ver que o conjunto de todas as transformações lineares de V em W herda uma estrutura natural de espaço vetorial. Sejam $T : V \to W$ e $U : V \to W$ transformações lineares e $k \in \mathbb{K}$. Definimos as aplicações de soma e produto por escalar:

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

Dados $T : V \to W$ e $U : W \to Z$ transformações lineares, podemos definir a aplicação de composição de U e T como

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

Exercício 3.8. *(a) Verifique que as aplicações $T+U$, kT e $U \circ T$ definidas acima, são transformações lineares.*

(b) Prove que o conjunto das transformações lineares de V em W, munido da adição e multiplicação por escalar, é um espaço vetorial sobre $\mathbb{K}$.

Denotaremos o espaço vetorial de todas as transformações lineares de V em W por $\mathcal{L}(V, W)$.

Exemplo 3.9. *Considere $T : \mathbb{R}^2 \to \mathbb{R}^3$ e $U : \mathbb{R}^3 \to \mathcal{M}_{2\times 3}(\mathbb{R})$ transformações lineares dadas por*

$$T(x,y) = \left(\frac{x-y}{2},\ 3y,\ y+4x\right) \quad e \quad U(x,y,z) = \begin{pmatrix} 4x & -z & 0 \\ x+y+z & 0 & -y+z \end{pmatrix}.$$

Calculemos $U \circ T : \mathbb{R}^2 \to \mathcal{M}_{2\times 3}(\mathbb{R})$. Com efeito,

$$\begin{aligned} (U\circ T)(x,y) &= U(T(x,y)) = U\left(\tfrac{x-y}{2},\ 3y,\ y+4x\right) \\ &= \begin{pmatrix} 4\left(\frac{x-y}{2}\right) & -(y+4x) & 0 \\ \frac{x-y}{2}+3y+y+4x & 0 & -3y+y+4x \end{pmatrix}. \end{aligned}$$

Portanto,

$$(U\circ T)(x,y) = \begin{pmatrix} 2(x-y) & -y-4x & 0 \\ \frac{9x+7y}{2} & 0 & -2y+4x \end{pmatrix}.$$

Definição 3.10. *Seja V um espaço vetorial sobre $\mathbb{K}$. Uma transformação linear de V em V é dito um operador linear sobre V.*

Vamos denotar por $T^n = T \circ \ldots \circ T$ (n vezes). Além disso, vamos definir por $T^0 = I$ se $T \neq 0$.

Lema 3.11. *Sejam V, W, Z espaços vetoriais sobre o corpo $\mathbb{K}$ e sejam $T_1, T_2 : V \to W$ 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 $U : V \to V$ é um operador linear então $I \circ U = U \circ I = U$ com I o operador identidade.

Demonstração: Fica como exercício a prova desse lema. ∎

3.1 Núcleo, Imagem e Isomorfismos Lineares

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ de dimensão finita.

Definição 3.12. *Seja $T : V \to W$ uma transformação linear. Definimos o **núcleo** de T, denotado por $Nuc(T)$, como o subconjunto de V dado por*

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

*E a **imagem** de T como o subconjunto de W, denotado $Im(T)$, dado por*

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

Na verdade, o núcleo e a imagem de uma transformação linear são mais que subconjuntos, de fato são subespaços vetoriais.

Afirmação 3.13. *Se $T : V \to W$ é uma transformação linear, então $Nuc(T)$ é um subespaço de V e $Im(T)$ é um subespaço de W.*

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

Sejam $u, v \in Nuc(T)$ e $k \in \mathbb{K}$ quaisquer. Assim $T(u) = 0_W = T(v)$ e mais $T(ku + v) = kT(u) + T(v) = k0_W + 0_W = 0_W$. Logo, $ku + v \in Nuc(T)$.

Por outro lado, sejam $z, w \in Im(T)$ e $k \in \mathbb{K}$. Assim existem $u, v \in V$ tais que $z = T(u)$ e $w = T(v)$. E note que $T(ku+v) = kT(u)+T(v) = kz+w$, o que implica que $kz+w \in Im(T)$. Donde segue o resultado. ■

Exemplo 3.14. *1. Seja a transformação linear $T : \mathbb{R}^3 \to \mathbb{R}^2$ dada por*

$$T(x, y, z) = (x+y,\ z-y).$$

Determinemos $Nuc(T)$ e $Im(T)$.

Observe que $T(x, y, z) = (0, 0)$ se, e somente se, $(x+y,\ z-y) = (0, 0)$, ou seja, $x+y = 0$ e $z-y = 0$. Logo, $x = -y$ e $z = y$. Portanto,

$$Nuc(T) = \{(-y, y, y);\ y \in \mathbb{R}\},$$

o qual descreve uma reta em $\mathbb{R}^3$. E $Im(T) = \{(x+y,\ z-y);\ x, y, z \in \mathbb{R}\}$.

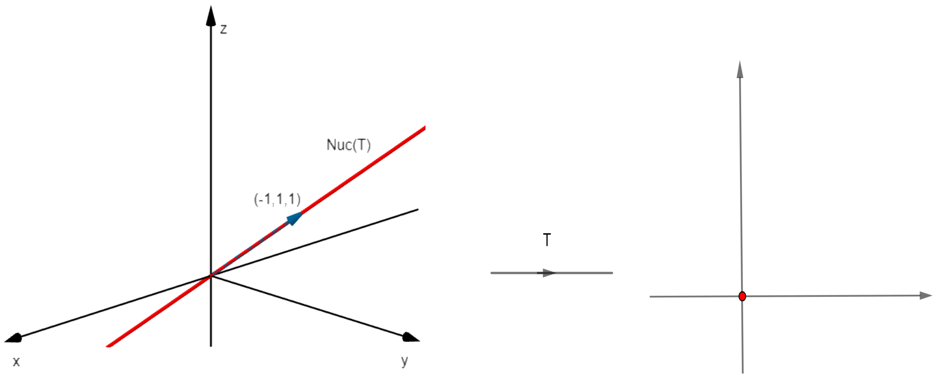

Figura 3.7: $Nuc(T) = \{(-y, y, y);\ y \in \mathbb{R}\}$, com $T(x, y, z) = (x+y,\ z-y)$

2. Seja $T : \mathcal{M}_{2\times 2}(\mathbb{R}) \to \mathcal{M}_{2\times 2}(\mathbb{R})$ dada por

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -2a & -b-c \\ 4a-c+d & 0 \end{pmatrix}.$$

Neste caso,

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -2a & -b-c \\ 4a-c+d & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

se, e somente se, $a = 0, b = -c$ *e* $d = c$*, ou seja,*

$$Nuc(T) = \left\{ \begin{pmatrix} 0 & -c \\ c & c \end{pmatrix} ;\ c \in \mathbb{R} \right\}.$$

Proposição 3.15. *Sejam* V *e* W *espaços vetoriais sobre* $\mathbb{K}$ *e* $T : V \to W$ *uma transformação linear. Se* $\alpha = \{v_1, \ldots, v_n\}$ *é uma base de* V*, então* $\{T(v_1), \ldots, T(v_n)\}$ *gera* $Im(T)$*.*

Demonstração: Seja $w \in Im(T)$. Assim existe $v \in V$ tal que $T(v) = w$. Como α é base de V, podemos escrever $v = k_1v_1 + \ldots + k_nv_n$, com $k_1, \ldots, k_n \in \mathbb{K}$. Assim,

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

Logo, w é uma combinação linear dos vetores $T(v_1), \ldots, T(v_n)$. Portanto $\{T(v_1), \ldots, T(v_n)\}$ gera $Im(T)$. ∎

Segue do resultado acima que $dim_{\mathbb{K}} Im(T) \leq n = dim_{\mathbb{K}} V$, pois de um conjunto que gera $Im(T)$ podemos sempre extrair uma base. O resultado a seguir é um dos principais da Álgebra Linear, pois ele nos dá a relação entre as dimensões do núcleo, da imagem e de V.

Teorema 3.16. ***(Teorema do Núcleo e da Imagem - TNI)*** *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 $\beta = \{v_1, \ldots, v_k\}$ uma base ordenada para $Nuc(T)$. Como $Nuc(T)$ é um subespaço de V, podemos completar a base β de modo a obter uma base de V. Considere assim $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ uma tal base de V. Provemos que $\{T(v_{k+1}), \ldots, T(v_n)\}$ é uma base de $Im(T)$, ou seja, que este conjunto gera $Im(T)$ e que é LI. De fato, dado $w \in Im(T)$, existe $u \in V$ tal que $w = T(u)$. Entretanto, como $u \in V$ ele se escreve de maneira única numa base de V, digamos $u = \lambda_1v_1 + \ldots + \lambda_kv_k + \lambda_{k+1}v_{k+1} + \ldots + \lambda_nv_n$. Deste modo, como $v_1, \ldots, v_k \in Nuc(T)$ temos

$$\begin{aligned} w &= T(u) = T(\lambda_1v_1 + \ldots + \lambda_kv_k + \lambda_{k+1}v_{k+1} + \ldots + \lambda_nv_n) \\ &= \lambda_1T(v_1) + \ldots + \lambda_kT(v_k) + \lambda_{k+1}T(v_{k+1}) + \ldots + \lambda_nT(v_n) \\ &= \lambda_{k+1}T(v_{k+1}) + \ldots + \lambda_nT(v_n). \end{aligned}$$

Logo, $Im(T) = [T(v_{k+1}), \ldots, T(v_n)]$.

Consideremos agora uma combinação linear nula arbitrária

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

com $a_i \in \mathbb{K}$, para $i = k+1, \ldots, n$. Como T é linear segue que $T(a_{k+1}v_{k+1} + \cdots + a_n v_n) = 0$, ou seja, $a_{k+1}v_{k+1} + \cdots + a_n v_n \in Nuc(T)$. Assim, existem $a_1, \ldots, a_k \in \mathbb{K}$ tais que

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

Logo, $-a_1v_1 - \ldots - a_kv_k + a_{k+1}v_{k+1} + \cdots + a_nv_n = 0$. Entretanto, $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ é base de V e, portanto, LI. Donde concluímos que $a_i = 0$ para todo $i = 1, \ldots, n$. Logo, os vetores $T(v_{k+1}), \ldots, T(v_n)$ são LI. Portanto, $dim_{\mathbb{K}} Nuc(T) = k$ e $dim_{\mathbb{K}} Im(T) = n - k$ o que implica que

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

■

Exemplo 3.17.

1. *No Exemplo 3.14 (item 1.) calculamos o núcleo da referida transformação linear, o qual foi descrito como:*

$$Nuc(T) = \{(-y, y, y);\ y \in \mathbb{R}\} = \{y(-1, 1, 1);\ y \in \mathbb{R}\},$$

ou seja, o vetor $(-1, 1, 1)$ gera o núcleo de T, e como o mesmo é LI temos que $\{(-1, 1, 1)\}$ é uma base para o subespaço $Nuc(T)$. Portanto, $dim_{\mathbb{R}} Nuc(T) = 1$. No entanto, segue do Teorema do Núcleo e da Imagem (TNI) que

$$dim_{\mathbb{R}} Im(T) = dim_{\mathbb{R}} \mathbb{R}^3 - dim_{\mathbb{R}} Nuc(T) = 3 - 1 = 2.$$

Como $Im(T)$ é um subespaço de $\mathbb{R}^2$ de dimensão 2, segue que $Im(T) = \mathbb{R}^2$.

2. *No Exemplo 3.14 (item 2.) temos que*

$$Nuc(T) = \left\{ \begin{pmatrix} 0 & -c \\ c & c \end{pmatrix};\ c \in \mathbb{R} \right\} = \left\{ c \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix};\ c \in \mathbb{R} \right\},$$

ou seja, de modo análogo $dim_{\mathbb{R}} Nuc(T) = 1$. Logo, pelo TNI temos que

$$dim_{\mathbb{R}} Im(T) = dim_{\mathbb{R}} \mathcal{M}_{2\times 2}(\mathbb{R}) - dim_{\mathbb{R}} Nuc(T) = 4 - 1 = 3.$$

Definição 3.18. *Seja $T : V \to W$ uma transformação linear, com V de dimensão finita. Definimos como o* **posto** *de T como sendo a dimensão da imagem de T, e por* **nulidade** *de T, como sendo a dimensão do núcleo de T. Notação: posto(T) e nulidade(T).*

Definição 3.19. *Seja $T : V \to W$ uma transformação linear. Dizemos que T é injetora, se para quaisquer $u, v \in V$ com $u \neq v$ tivermos $T(u) \neq T(v)$, ou equivalentemente, se $T(u) = T(v)$, então $u = v$. Dizemos ainda que T é sobrejetora se $Im(T) = W$, ou seja, para todo $w \in W$, existe $v \in V$ tal que $T(v) = w$.*

Exemplo 3.20. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x, y) = (2x, 3y, x + y)$.*

Observe que para todo $(x, y), (z, w) \in \mathbb{R}^2$ tais que $T(x, y) = T(z, w)$, ou seja,

$$(2x, 3y, x + y) = (2z, 3w, z + w).$$

Logo, $2x = 2z$ e $3y = 3w$, ou seja, $x = z$ e $y = w$. Desse modo, $(x, y) = (z, w)$. Portanto, T é injetora.

Veja ainda que

$$Im(T) = \{(2x, 3y, x + y); x, y \in \mathbb{R}\} = \{x(2, 0, 1) + y(0, 3, 1); x, y \in \mathbb{R}\}.$$

Assim $\{(2, 0, 1), (0, 3, 1)\}$ gera $Im(T)$ e como seus vetores são LI, temos que esse conjunto é uma base para $Im(T)$. Portanto,

$$dim_{\mathbb{R}} Im(T) = 2 < 3 = dim_{\mathbb{R}} \mathbb{R}^3.$$

Na verdade, a imagem de T é um plano em $\mathbb{R}^3$.

Proposição 3.21. *Seja $T : V \to W$ uma transformação linear. Então, T é injetora se, e somente se, $Nuc(T) = \{0_V\}$.*

Demonstração: Seja $v \in Nuc(T)$. Assim $T(v) = 0_W$ o que implica que $T(v) = 0_W = T(0_V)$, pois T é linear. Logo, $v = 0_V$ já que T é injetora. Reciprocamente, sejam $u, v \in V$ tais que $T(u) = T(v)$. Assim como T é linear temos que $T(u - v) = T(u) - T(v) = 0_W$. O que significa que $u - v \in Nuc(T)$. Entretanto, por hipótese $Nuc(T) = \{0_V\}$, logo $u - v = 0_V$ e assim $u = v$. ■

Exemplo 3.22. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^n$ com $n \geq 2$ dada por $T(x, y) = (x, 0, \ldots, 0)$. Observe que T não é injetora, pois $(1, 5) \neq (1, 7)$, entretanto $T(1, 5) = (1, 0, \ldots, 0) = T(1, 7)$. Na verdade,*

$v \in Nuc(T)$ se, e somente se, $(x, 0, \ldots, 0) = (0, 0, \ldots, 0)$, ou seja, $x = 0$. Deste modo

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

Mas pelo Teorema do Núcleo e da Imagem, temos que

$$dim_{\mathbb{R}} Im(T) = dim_{\mathbb{R}} \mathbb{R}^2 - dim_{\mathbb{R}} Nuc(T) = 2 - 1 = 1.$$

Portanto, $Im(T) \neq \mathbb{R}^n$, pois $n \geq 2$ e $dim_{\mathbb{R}} Im(T) = 1 < 3 = dim_{\mathbb{R}} \mathbb{R}^3$. Donde concluímos que T não é sobrejetora.

Uma transformação linear $T : V \to W$ é dita **bijetora** quando for injetora e sobrejetora. Transformações lineares bijetoras são também denominadas **isomorfismos**. Nesse caso, dizemos que V e W são espaços vetoriais isomorfos, o qual vamos denotar $V \simeq W$.

Exemplo 3.23.

1. *Os espaços $\mathbb{C}$ e $\mathbb{R}^2$ são $\mathbb{R}$-espaços vetoriais isomorfos. De fato, considere a aplicação $T : \mathbb{C} \to \mathbb{R}^2$ em que para cada $z = x + iy \in \mathbb{C}$ com $x, y \in \mathbb{R}$ definimos*

$$T(x + iy) \;=\; (x, y).$$

Verifique que T é linear, injetora e sobrejetora.

2. *Mais geralmente, os $\mathbb{R}$-espaços vetoriais $\mathbb{C}^n$ e $\mathbb{R}^{2n}$ são isomorfos. De fato, considere a aplicação $T : \mathbb{C}^n \to \mathbb{R}^{2n}$ que a cada $w = (x_1 + iy_1, x_2 + iy_2, \ldots, x_n + iy_n) \in \mathbb{C}^n$ o vetor*

$$T(w) = (x_1,\ y_1,\ x_2,\ y_2, \ldots, x_n,\ y_n) \in \mathbb{R}^{2n}.$$

Como exercício prove que T é linear e bijetora.

3. *Verifique que os espaços $\mathcal{M}_{2\times 2}(\mathbb{R})$ e $\mathbb{R}^4$ são isomorfos, provando que a aplicação*

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

é uma transformação linear bijetora.

Definição 3.24. *Uma aplicação $T : V \to W$ é dita inversível se existir $T' : W \to V$ tal que*

$$T \circ T' = I_W \quad e \quad T' \circ T = I_V,$$

com I_V, I_W as identidades de V e W, respectivamente. A aplicação T' vai ser denotada por T^{-1} e denominada a inversa de T.

Proposição 3.25. *Uma aplicação $T : V \to W$ é bijetora se, e somente se, T for 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 $T^{-1} : W \to V$ tal que $T^{-1}(w) = v$. Observe que

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

ou seja, T^{-1} é a inversa de T.

Reciprocamente, se T é inversível, então T é injetora, pois se $T(u) = T(v)$, então

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

o que implica que $u = v$. E T é sobrejetora, pois dado $w \in W$ tome $v = T^{-1}(w) \in V$. Logo, $T(v) = T(T^{-1}(w)) = w$. ■

Proposição 3.26. *Se $T : V \to W$ for uma transformação linear inversível, então sua inversa $T^{-1} : W \to V$ é uma transformação linear.*

Demonstração: Sejam $u, w \in W$ e $k \in \mathbb{K}$ quaisquer. Como T é inversível existem únicos $a, b \in V$ tais que $u = T(a)$ e $w = T(b)$. Assim $ku + w = kT(a) + T(b) = T(ka + b)$. Logo,

$$T^{-1}(ku + w) = T^{-1}(T(ka + b)) = I_V(ka + b) = ka + b = kT^{-1}(u) + T^{-1}(v).$$

Portanto, T^{-1} é linear. ■

Exemplo 3.27. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x, y) = (2x, -y)$. Verifique que T é inversível e determine T^{-1}.*

Observe que $T(x, y) = (0, 0)$ se, e somente se, $(2x, -y) = (0, 0)$, cuja solução é $x = y = 0$. Logo, $Nuc(T) = \{(0, 0)\}$, ou seja, T é injetora. Mais ainda, pelo TNI

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

Como $Im(T)$ é um subespaço de $\mathbb{R}^2$ de dimensão 2, concluímos que $Im(T) = \mathbb{R}^2$, ou seja, T é sobrejetora. Portanto, T é bijetora e com isso inversível.

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

$$T^{-1}(2,0)=(1,0) \quad e \quad T^{-1}(0,-1)=(0,1).$$

Note que $\{(2,0),(0,-1)\}$ é uma base de $\mathbb{R}^2$. Logo, para todo $(x,y)\in\mathbb{R}^2$ temos que existem $a,b\in\mathbb{R}$ tais que $(x,y)=a(2,0)+b(0,-1)$. Resolvendo esse sistema obtemos $a=\frac{x}{2}$ e $b=-y$. Portanto,

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

Como T^{-1} é linear, obtemos

$$T^{-1}(x,y)=\frac{x}{2}T^{-1}(2,0)-yT^{-1}(0,-1)=\frac{x}{2}(1,0)-y(0,1),$$

donde concluímos que $T^{-1}(x,y)=\left(\frac{x}{2},-y\right)$.

Proposição 3.28. *Sejam $T,T_1:V\to W$ e $T_2:W\to Z$ transformações lineares inversíveis e $k\in\mathbb{K}^*$. Então:*

(a) $(T^{-1})^{-1}=T$;

(b) $(kT)^{-1}=k^{-1}T^{-1}$;

(c) $(T_2\circ T_1)^{-1}=T_1^{-1}\circ T_2^{-1}$.

Demonstração: (a) Basta observar que $T^{-1}\circ T=I_V$ e $T\circ T^{-1}=I_W$, ou seja, a inversa da aplicação T^{-1} é T.

(b) Para todo $v\in V$ e $w\in W$ temos que

$$\begin{aligned}((k^{-1}T^{-1})\circ(kT))(v) &= (k^{-1}T^{-1})(kT(v))=k(k^{-1}T^{-1})(T(v))\\ &= (k\cdot k^{-1})T^{-1}(T(v))=(T^{-1}\circ T)(v)=v\end{aligned}$$

e

$$\begin{aligned}((kT)\circ(k^{-1}T^{-1}))(w) &= (kT)(k^{-1}T^{-1}(w))=k^{-1}((kT)(T^{-1}(w))\\ &= (k^{-1}\cdot k)T(T^{-1}(w))=(T\circ T^{-1})(w)=w.\end{aligned}$$

Portanto, $(kT)^{-1}=k^{-1}T^{-1}$.

(c) De fato, para todo $v\in V$ e $w\in Z$ temos que

$$\begin{aligned}((T_2\circ T_1)\circ(T_1^{-1}\circ T_2^{-1}))(w) &= (T_2\circ T_1)(T_1^{-1}(T_2^{-1}(w)))=T_2(T_1(T_1^{-1}(T_2^{-1}(w)))\\ &= T_2(T_2^{-1}(w))=w.\end{aligned}$$

Analogamente, $(T_1^{-1}\circ T_2^{-1})\circ(T_2\circ T_1)(v)=v$. ∎

Proposição 3.29. *Seja $T : V \to W$ uma transformação linear.*

1. *Se T for injetora, então $dim_{\mathbb{K}}V \leq dim_{\mathbb{K}}W$;*
2. *Se T for sobrejetora, então $dim_{\mathbb{K}}V \geq dim_{\mathbb{K}}W$;*
3. *Se T for inversível, então $dim_{\mathbb{K}}V = dim_{\mathbb{K}}W$.*

Demonstração: Se T é injetora, segue que $Nuc(T) = \{0\}$ e, portanto, $dim_{\mathbb{K}}Nuc(T) = 0$. Pelo Teorema do Núcleo e da Imagem,

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

Se T é sobrejetora, então $Im(T) = W$ donde segue do Teorema do Núcleo e da Imagem que

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

Portanto, $dim_{\mathbb{K}}V \geq dim_{\mathbb{K}}W$.

Se T é inversível o resultado segue dos casos anteriores. ■

Teorema 3.30. *Seja $T : V \to W$ uma transformação linear. Então T é injetora se, e somente se, T leva todo subconjunto LI de V sobre um subconjunto LI de W.*

Demonstração: Seja S um subconjunto LI de V e sejam $v_1, \ldots, v_k$ vetores em S. 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

$$\lambda_1 T(v_1) + \ldots + \lambda_k T(v_k) = 0.$$

O que implica $T(\lambda_1 v_1 + \ldots + \lambda_k v_k) = 0$, ou seja, $\lambda_1 v_1 + \ldots + \lambda_k v_k \in Nuc(T)$. Donde concluímos que $\lambda_1 v_1 + \ldots + \lambda_k v_k = 0_V$, pois $Nuc(T) = \{0_V\}$ já que T é injetora. Como os vetores $v_1, \ldots, v_k$ são LI, segue que $\lambda_i = 0$ para todo $i = 1, \ldots, n$.

Reciprocamente, seja $v \in V$ um vetor não nulo. Então $S = \{v\}$ é um subconjunto LI de V. Mas por hipótese, $T(S) = \{T(v)\}$ é LI. Portanto, $T(v) \neq 0$ o que implica que $Nuc(T) = \{0\}$, ou seja, T é injetora. ■

Teorema 3.31. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$ de dimensão finita tais que $dim_{\mathbb{K}}V = dim_{\mathbb{K}}W$. Se $T : V \to W$ é uma transformação linear, então são equivalentes:*

(i) T é injetora;

(ii) T é sobrejetora;

(iii) T leva base de V em base de W;

(iv) existe pelo menos uma base de V tal que T leva em uma base de W.

Demonstração: $(i) \Rightarrow (ii)$ Como T é injetora temos que $Nuc(T) = \{0\}$. Donde segue do Teorema do Núcleo e da Imagem que $dimV = dimIm(T)$. Mas por hipótese, $dimV = dimW$, logo $dimW = dimIm(T)$. Como $Im(T)$ é um subespaço de W e esses possuem a mesma dimensão, concluímos que $Im(T) = W$ e, portanto, T é sobrejetora.

$(ii) \Rightarrow (i)$ Como T é sobrejetora segue do Teorema do Núcleo e da Imagem que $dimV = dimNuc(T) + dimW$. Logo, $dimNuc(T) = 0$ já que $dimV = dimW$. Assim $Nuc(T) = \{0\}$, ou seja, T é injetora.

$(i) \Rightarrow (iii)$ Seja $\beta = \{v_1, \ldots, v_n\}$ uma base de V. Pelo Teorema anterior $\gamma = \{T(v_1), \ldots, T(v_n)\}$ é um conjunto LI de W. Como $dimW = dimV = n$ temos que γ é base de W.

$(iii) \Rightarrow (iv)$ Imediato.

$(iv) \Rightarrow (i)$ Suponhamos que exista alguma base $\{v_1, \ldots, v_n\}$ de V tal que $\{T(v_1), \ldots, T(v_n)\}$ seja base de W. Como os vetores $T(v_i)$ geram W, é claro que $Im(T) = W$. Seja $v \in Nuc(T)$ dado por $v = k_1v_1 + \ldots + k_nv_n$, com $k_i \in \mathbb{K}, i = 1, \ldots, n$. Logo,

$$0 = T(v) = k_1T(v_1) + \ldots + k_nT(v_n)$$

o que implica que $k_i = 0$ para todo $i = 1, \ldots, n$, ou seja, $v = 0$. Como T é injetora e sobrejetora, segue que T é inversível. ■

Em particular, se T for inversível vimos que $dim_{\mathbb{K}}V = dim_{\mathbb{K}}W$. Portanto todas as condições do resultado anterior são equivalentes.

Teorema 3.32. *Quaisquer dois espaços vetoriais V e W de mesma dimensão são isomorfos.*

Demonstração: Se ambos os espaços V e W forem nulos, não temos nada a provar. Desta maneira, considere V e W espaços vetoriais sobre um corpo $\mathbb{K}$ de dimensão $n \geq 1$, e $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_n\}$ bases de V e W, respectivamente. Pelo Teorema 3.6 existe uma única transformação linear $T : V \to W$ tal que $T(v_i) = w_i$ para todo $i = 1, \ldots, n$. Vamos mostrar que T é sobrejetora. De fato, seja $w = \sum_{i=1}^{n} \lambda_i w_i \in W$ com $\lambda_i \in \mathbb{K}$ e $i = 1, \ldots, n$. Tomando $v = \sum_{i=1}^{n} \lambda_i v_i \in V$ vemos que

$$T(v) = T\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i T(v_i) = \sum_{i=1}^{n} \lambda_i w_i = w.$$

Portanto, T é sobrejetora e como V e W possuem mesma dimensão, segue que T é invertível. Donde concluímos que T um isomorfismo. ■

Corolário 3.33. *Todo espaço vetorial V sobre um corpo $\mathbb{K}$ de dimensão $n \geq 1$ é isomorfo a $\mathbb{K}^n$.*

O corolário acima garante, por exemplo, que

$$\mathcal{M}_{n\times n}(\mathbb{R}) \simeq \mathbb{R}^{n^2}, \quad \mathcal{M}_{n\times m}(\mathbb{R}) \simeq \mathbb{R}^{nm} \quad \text{e} \quad \mathcal{P}_n \simeq \mathbb{R}^{n+1}.$$

Além disso, $\mathbb{C}^n$ como espaço vetorial sobre $\mathbb{R}$ é isomorfo à $\mathbb{R}^{2n}$, etc.

3.2 Matriz de uma Transformação Linear

Sejam V e W espaços vetoriais sobre $\mathbb{K}$ de dimensões n e m, respectivamente. Considere $T : V \to W$ uma transformação linear e $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_m\}$ bases ordenadas de V e W, respectivamente. Sabemos que dado $v \in V$, o mesmo se escreve de modo único na base α da forma $v = k_1v_1 + \ldots + k_nv_n$ com $k_i \in \mathbb{K}$ e $i = 1, \ldots, n$. Assim

$$T(v) = T(k_1v_1 + \ldots + k_nv_n) = k_1T(v_1) + \ldots + k_nT(v_n). \tag{3.1}$$

Como $T(v) \in W$, ou seja, existem $l_1, \ldots, l_m$ em $\mathbb{K}$ tais que $T(v) = l_1w_1 + \ldots + l_mw_m$. Todavia cada $T(v_i) \in W$ para todo $i = 1, \ldots, n$, assim existem $a_{ij} \in \mathbb{K}$ com $i, j = 1, \ldots, n$ tais que

$$\left\{\begin{array}{c} T(v_1) = a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}w_m \\ T(v_2) = a_{12}w_1 + a_{22}w_2 + \ldots + a_{m2}w_m \\ \vdots \\ T(v_n) = a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m. \end{array}\right. \tag{3.2}$$

Substituindo (3.2) em (3.1) temos

$$\begin{aligned} T(v) &= k_1(a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}w_m) + \ldots + k_n(a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m) \\ &= (k_1a_{11} + \ldots + k_na_{1n})w_1 + \ldots + (k_1a_{m1} + \ldots + k_na_{mn})w_m. \end{aligned}$$

Pela unicidade da representação de um vetor em uma base, temos que

$$\left\{\begin{array}{ccc} l_1 & = & k_1a_{11} + \ldots + k_na_{1n} \\ & \vdots & \\ l_m & = & k_1a_{m1} + \ldots + k_na_{mn}. \end{array}\right.$$

Na forma matricial:

$$\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.$$

Denotando a matriz $(a_{ij})_{m\times n}$ por $[T]^{\alpha}_{\beta}$, a relação acima pode ser reescrita da forma

$$[T(v)]_{\beta} = [T]^{\alpha}_{\beta} \cdot [v]_{\alpha}. \tag{3.3}$$

A matriz $[T]^{\alpha}_{\beta}$ é denominada a **matriz da transformação linear** T nas bases α e β.

O que fizemos acima é a prova do seguinte resultado.

Teorema 3.34. *Sejam V, W espaços vetoriais sobre $\mathbb{K}$ com $dim_{\mathbb{K}}V = n$ e $dim_{\mathbb{K}}W = m$. Sejam α e β bases ordenadas de V e W, respectivamente. Para cada transformação linear $T : V \to W$ existe uma $m \times n$ matriz A sobre o corpo $\mathbb{K}$, chamada a matriz de T nas bases α e β tal que*

$$[T(v)]_{\beta} = A \cdot [v]_{\alpha},$$

para todo $v \in V$.

Exemplo 3.35. *Se $I : V \to V$ é a transformação identidade e $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_n\}$ são bases de V, então a matriz de I em relação as bases α e β é dada por:*

$$\begin{cases} I(v_1) = v_1 = a_{11}w_1 + a_{21}w_2 + \ldots + a_{n1}w_n \\ I(v_2) = v_2 = a_{12}w_1 + a_{22}w_2 + \ldots + a_{n2}w_n \\ \qquad\vdots \\ I(v_n) = v_n = a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{nn}w_n, \end{cases}$$

ou seja,

$$[I]^{\alpha}_{\beta} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

a qual é exatamente a matriz mudança da base α para a base β, que satisfaz

$$[v]_{\beta} = [I(v)]_{\beta} = [I]^{\alpha}_{\beta} \cdot [v]_{\alpha}.$$

Se $T : V \to V$ é um operador linear e α é uma base de V, denotamos a matriz $[T]^{\alpha}_{\alpha}$

simplesmente por $[T]_\alpha$.

Exemplo 3.36. *Sejam* $T : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por*

$$T(x, y, z) = (2x + y - z, 3x - 2y + 4z)$$

e $\alpha = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ *e* $\beta = \{(1, 3), (1, 4)\}$ *bases de* $\mathbb{R}^3$ *e* $\mathbb{R}^2$*, respectivamente. Determinemos a matriz de* T *nas bases* α *e* β*. Para tanto, vamos calcular:*

$$\begin{aligned} T(1, 1, 1) &= (2, 5) = a(1, 3) + b(1, 4) \\ T(1, 1, 0) &= (3, 1) = c(1, 3) + d(1, 4) \\ T(1, 0, 0) &= (2, 3) = e(1, 3) + f(1, 4). \end{aligned}$$

Resolvendo os sistemas acima concluímos que $a = 3$, $b = -1$, $c = 11$, $d = -8, e = 5$ *e* $f = -3$*. Logo,*

$$[T]^\alpha_\beta = \begin{pmatrix} 3 & 11 & 5 \\ -1 & -8 & -3 \end{pmatrix}.$$

Agora, note que se tomarmos γ *e* δ *as bases canônicas de* $\mathbb{R}^3$ *e* $\mathbb{R}^2$*, respectivamente, a matriz de* T *é bem mais fácil de ser calculada, pois*

$$\begin{aligned} T(1, 0, 0) &= (2, 3) = 2(1, 0) + 3(0, 1) \\ T(0, 1, 0) &= (1, -2) = 1(1, 0) - 2(0, 1) \\ T(0, 0, 1) &= (-1, 4) = -1(1, 0) + 4(0, 1). \end{aligned}$$

Nesse caso, a matriz de T *nas bases* γ *e* δ *é dada por*

$$[T]^\gamma_\delta = \begin{pmatrix} 2 & 1 & -1 \\ 3 & -2 & 4 \end{pmatrix}.$$

Observe que nas bases canônicas, a matriz de T *pode ser determinada diretamente das coordenadas dessa transformação.*

Vimos anteriormente que dado $T : V \to W$ uma transformação linear, se fixarmos duas bases ordenadas α e β de V e W respectivamente, associamos a essa transformação a matriz $[T]^\alpha_\beta$. Reciprocamente, se tomamos α uma base ordenada de um espaço vetorial V, β uma base ordenada de um espaço vetorial W e dada uma matriz $A = (a_{ij})_{m\times n}$ com entradas em $\mathbb{K}$, então existe uma única transformação linear $T : V \to W$ tal que a matriz $[T]^\alpha_\beta = A$.

De fato, a transformação T está unicamente determinada, pois conhecendo a matriz A sabemos o valor da imagem por T dos vetores da base α.

Exemplo 3.37. *Considere as bases* $\beta = \{(1,1),(0,1)\}$ *de* $\mathbb{R}^2$ *e* $\gamma = \{(0,3,0),(-1,0,0),(0,1,1)\}$ *de* $\mathbb{R}^3$. *Determinemos a transformação linear* $T:\mathbb{R}^2 \to \mathbb{R}^3$ *cuja matriz é*

$$[T]_\gamma^\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \\ -1 & 3 \end{pmatrix}.$$

Por definição,

$$\begin{aligned} T(1,1) &= 0(0,3,0) - 1(-1,0,0) - 1(0,1,1) = (1,-1,-1) \\ T(0,1) &= 2(0,3,0) + 0(-1,0,0) + 3(0,1,1) = (0,9,3). \end{aligned}$$

Entretanto, já vimos que se soubermos o que uma transformação linear faz com uma base, sabemos determiná-la. Fica como exercício determinar a transformação T.

Exercício 3.38. *Sejam* $T_1, T_2 : V \to W$ *transformações lineares com* α *e* β *bases ordenadas de* V *e* W, *respectivamente, e seja* $\lambda \in \mathbb{K}$. *Prove que:*

(a) $[T_1 + T_2]_\beta^\alpha = [T_1]_\beta^\alpha + [T_2]_\beta^\alpha$,

(b) $[\lambda T_1]_\beta^\alpha = \lambda [T_1]_\beta^\alpha$.

Teorema 3.39. *Sejam* V, W *espaços vetoriais sobre* $\mathbb{K}$ *com* $dim_{\mathbb{K}}V = n$ *e* $dim_{\mathbb{K}}W = m$. *Os espaços* $\mathcal{L}(V,W)$ *e* $\mathcal{M}_{m\times n}(\mathbb{K})$ *são isomorfos.*

Demonstração: Sejam α e β bases ordenadas de V e W, respectivamente. Considere a aplicação Ψ que associa a cada transformação linear $T \in \mathcal{L}(V,W)$ sua matriz nas bases α e β, ou seja,

$$\begin{aligned} \Psi : \mathcal{L}(V,W) &\longrightarrow \mathcal{M}_{m\times n}(\mathbb{K}) \\ T &\longmapsto [T]_\beta^\alpha. \end{aligned}$$

Provemos que T é um isomorfismo linear, ou seja, que T é uma bijeção linear.

Dados $T_1, T_2 \in \mathcal{L}(V,W)$ e $\lambda \in \mathbb{K}$ segue do exercício anterior que

$$\Psi(\lambda T_1 + T_2) = [\lambda T_1 + T_2]_\beta^\alpha = \lambda [T_1]_\beta^\alpha + [T_2]_\beta^\alpha = \lambda \Psi(T_1) + \Psi(T_2),$$

isto é, Ψ é linear.

Além disso, Ψ é injetora, pois dado $T \in Nuc(\Psi)$ segue que $\Psi(T) = [T]_\beta^\alpha = 0$. Supondo $\alpha = \{v_1, \ldots, v_n\}$ e $\beta = \{w_1, \ldots, w_n\}$, se a matriz de T nas bases α e β é a matriz nula, significa que $T(v_i) = 0w_1 + \ldots + 0w_n = 0$ para todo $i = 1, \ldots, n$. Portanto, dado qualquer vetor $v = k_1v_1 + \ldots + k_nv_n \in V$ temos que $T(v) = \sum_{i=1}^n k_i T(v_i) = 0$, ou seja, T é a transformação identicamente nula. Logo, $Nuc(\Psi) = \{0\}$.

Falta provarmos que Ψ é sobrejetora. Considere $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$. Pelo Teorema 3.6, existe uma única transformação linear T que leva os vetores da base α nos seguintes vetores:

$$\begin{aligned}
T(v_1) &= a_{11}w_1 + a_{21}w_2 + \ldots + a_{m1}w_m \\
T(v_2) &= a_{12}w_1 + a_{22}w_2 + \ldots + a_{m2}w_m \\
&\vdots \\
T(v_n) &= a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m.
\end{aligned}$$

Por construção, $\Psi(T) = [T]_\beta^\alpha = A$.

Donde concluímos que T é um isomorfismo. ■

Corolário 3.40. *Sejam V e W espaços vetoriais sobre $\mathbb{K}$, com $dim_{\mathbb{K}}V = n$ e $dim_{\mathbb{K}}W = m$. O espaço $\mathcal{L}(V,W)$ possui dimensão finita e mais $dim_{\mathbb{K}}\mathcal{L}(V,W) = nm$.*

Demonstração: Segue do teorema anterior que $\mathcal{L}(V,W) \simeq \mathcal{M}_{m\times n}(\mathbb{K})$ e, portanto,

$$dim_{\mathbb{K}}\mathcal{L}(V,W) = dim\mathcal{M}_{m\times n}(\mathbb{K}) = nm.$$

■

Teorema 3.41. *Seja V um espaço vetorial sobre $\mathbb{K}$ de dimensão n e sejam $\beta = \{v_1, \ldots, v_n\}$ e $\beta' = \{w_1, \ldots, w_n\}$ bases ordenadas de V. Se $T : V \to V$ é um operador linear, então existe uma única $n \times n$ matriz inversível P tal que*

$$[T]_{\beta'} = P^{-1}[T]_\beta P.$$

Demonstração: Vimos no Teorema 2.38 que para todo $v \in V$, existe uma única matriz inversível P tal que

$$[v]_\beta = P[v]_{\beta'}. \tag{3.4}$$

A matriz P é a matriz de mudança da base β para a base β'. Por definição, $[T(v)]_\beta = [T]_\beta[v]_\beta$. Aplicando (3.4) para o vetor $T(v)$ obtemos

$$[T(v)]_\beta = P[T(v)]_{\beta'} \Leftrightarrow [T]_\beta P[v]_{\beta'} = P[T(v)]_{\beta'} \Leftrightarrow P^{-1}[T]_\beta P[v]_{\beta'} = [T(v)]_{\beta'}.$$

Logo, $[T]_{\beta'} = P^{-1}[T]_\beta P$. ■

Definição 3.42. *Sejam A e B matrizes de ordem n com entradas em $\mathbb{K}$. Dizemos que B é semelhante à matriz A, se existe uma matriz inversível P em $\mathcal{M}_{n\times n}(\mathbb{K})$ tal que $B = P^{-1}AP$.*

Pelo teorema acima, as matrizes $[T]_{\beta'}$ e $[T]_{\beta}$ são semelhantes.

Teorema 3.43. *Sejam $T_1 : V \to W$ e $T_2 : W \to Z$ transformações lineares e α, β e γ bases de V, W e Z, respectivamente. Então*

$$[T_2 \circ T_1]_{\gamma}^{\alpha} \;=\; [T_2]_{\gamma}^{\beta} \cdot [T_1]_{\beta}^{\alpha}.$$

Corolário 3.44. *Seja $T : V \to W$ uma transformação linear inversível e α e β bases de V e W, respectivamente. Então $T^{-1} : W \to V$ satisfaz*

$$[T^{-1}]_{\alpha}^{\beta} = ([T]_{\beta}^{\alpha})^{-1}.$$

Demonstração: Segue do teorema acima que

$$I_V = [I_V]_{\alpha}^{\alpha} = [T^{-1} \circ T]_{\alpha}^{\alpha} = [T^{-1}]_{\alpha}^{\beta} \cdot [T]_{\beta}^{\alpha}.$$

Analogamente, $I_W = [I_W]_{\beta}^{\beta} = [T \circ T^{-1}]_{\beta}^{\beta} = [T]_{\beta}^{\alpha} \cdot [T^{-1}]_{\alpha}^{\beta}$.

Portanto, a inversa da matriz $[T]_{\beta}^{\alpha}$ é $[T^{-1}]_{\alpha}^{\beta}$. ∎

Corolário 3.45. *Sejam $T : V \to W$ uma transformação linear com $dim_{\mathbb{K}}V = dim_{\mathbb{K}}W$ e α e β bases ordenadas de V e W, respectivamente. Então, T é inversível se, e somente se,*

$$det([T]_{\beta}^{\alpha}) \neq 0.$$

Demonstração: Primeiramente, se $dim_{\mathbb{K}}V = dim_{\mathbb{K}}W = n$, então $[T]_{\beta}^{\alpha}$ é uma matriz quadrada de ordem n. Se T é inversível, segue do Corolário 3.44 que a matriz $[T]_{\beta}^{\alpha}$ também é inversível e, portanto, seu determinante é diferente de zero.

Reciprocamente, suponha que $det([T]_{\beta}^{\alpha}) \neq 0$, ou seja, a matriz $[T]_{\beta}^{\alpha}$ admite uma inversa. Entretanto, os espaços vetoriais $\mathcal{M}_{n\times n}(\mathbb{K})$ e $\mathcal{L}(W, V)$ são isomorfos. Desse modo, fixado as bases α e β, existe uma única transformação linear $S : W \to V$ associada à matriz $([T]_{\beta}^{\alpha})^{-1}$, ou seja, tal que $[S]_{\alpha}^{\beta} = ([T]_{\beta}^{\alpha})^{-1}$. Logo,

$$[S]_{\alpha}^{\beta} \cdot [T]_{\beta}^{\alpha} = ([T]_{\beta}^{\alpha})^{-1} \cdot [T]_{\beta}^{\alpha} = I.$$

Mas pelo Teorema 3.43, $[S \circ T]_{\beta}^{\beta} = [S]_{\alpha}^{\beta} \cdot [T]_{\beta}^{\alpha} = I$. Provemos que $S \circ T = I$. Se denotamos

$\beta = \{v_1, \ldots, v_n\}$, por definição de matriz de uma transformação linear, $(S \circ T)(v_i) = v_i$ para todo $i = 1, \ldots, n$. Assim, para todo $v = k_1v_1 + \ldots + k_nv_n \in V$ com $k_i \in \mathbb{K}$, $i = 1, \ldots, n$ temos que

$$(S \circ T)(v) = (S \circ T)\left(\sum_{i=1}^{n} k_iv_i\right) = \sum_{i=1}^{n} k_i(S \circ T)(v_i) = \sum_{i=1}^{n} k_iv_i = v,$$

ou seja, $S \circ T = I$. Analogamente, $T \circ S = I$. Portanto, T é inversível. ■

Corolário 3.46. *Um operador linear $T : V \to V$ é inversível se, e somente se, o determinante da matriz de T é diferente de zero, qualquer que seja a base escolhida.*

Demonstração: Sejam β e β' bases de V e $[T]_\beta$ e $[T]_{\beta'}$ as matrizes de T nas bases β e β', respectivamente. Sabemos que $[T]_\beta$ e $[T]_{\beta'}$ são semelhantes, ou seja, existe uma matriz inversível P tal que $[T]_{\beta'} = P^{-1}[T]_\beta P$. Note que

$$\begin{aligned} det([T]_{\beta'}) &= det(P^{-1}[T]_\beta P) = det(P^{-1}) \cdot det([T]_\beta) \cdot det(P) \\ &= \tfrac{1}{det(P)} \cdot det([T]_\beta) \cdot det(P) = det([T]_\beta). \end{aligned}$$

Portanto, T é inversível se, e somente se, $det([T]_\beta) = det([T]_{\beta'}) \neq 0$. Logo, o determinante da matriz de T é não nulo, qualquer que seja a base escolhida. ■

Como o corolário acima garante que um operador linear $T : V \to V$ é inversível se, e somente se, $det([T]_\beta) \neq 0$ independente da base escolhida, quando não houver risco de confusão ou a propriedade a ser explorada não depender da base escolhida, vamos omitir a base na matriz de T, representando sua matriz simplesmente por $[T]$.

Exemplo 3.47. *Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear dada por $T(x, y) = (3x - y, x + 2y)$. Verifiquemos se T é inversível e, em caso afirmativo, vamos calcular a inversa de T via sua matriz em uma dada base.*

A matriz de T na base canônica é dada por:

$$[T] = \begin{pmatrix} 3 & -1 \\ 1 & 2 \end{pmatrix}$$

Como $det([T]) = 6 + 1 = 7 \neq 0$, o resultado anterior nos garante que T é inversível. Mais ainda,

$$[T^{-1}] = ([T])^{-1} = \frac{1}{7} \cdot \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}.$$

Assim, da relação (3.3) temos que na base canônica, as coordenadas do vetor $T^{-1}(x,y)$ é dada por

$$[T^{-1}(x,y)] = [T^{-1}] \cdot [(x,y)] = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{3}{7} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

Logo,

$$T^{-1}(x,y) = \left(\frac{2}{7}x + \frac{1}{7}y,\ -\frac{1}{7}x + \frac{3}{7}y\right).$$

É comum cometermos um certo um abuso de notação, principalmente quando trabalhamos com a base canônica. Muitas vezes misturamos as coordenadas de um vetor em uma dada base (que é uma matriz) com o vetor, como por exemplo: Se a matriz de uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$ é dada por

$$[T] = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

então escrevemos

$$\begin{aligned} T(x_1, \ldots, x_n) &= [T] \cdot [(x_1, \ldots, x_n)] = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \ldots, a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n). \end{aligned}$$

3.3 Autovalores e Autovetores de um Operador Linear

Os conceitos explorados nesta seção são de suma importância na matemática pura e aplicada. Neste último caso, diversos problemas na natureza são descritos por meio de uma equação, chamada equação diferencial, e calcular os autovalores e autovetores da parte linear de tal equação, se existirem, nos dão informações sobre as soluções de uma tal equação. O processo de diagonalização de um operador linear, ou mais geralmente de uma matriz, também é algo muito interessante, visto que uma matriz na forma diagonal é muito simples de se trabalhar e com isso podemos simplificar o problema a ser estudado.

Daqui em diante assumimos que $T : V \to V$ é um operador linear sobre um espaço vetorial n-dimensional V sobre um corpo $\mathbb{K}$.

Definição 3.48. *Um **autovalor** de um operador linear $T : V \to V$ é um escalar $\lambda \in \mathbb{K}$ tal que existe um vetor não nulo $v \in V$ com $T(v) = \lambda v$. Se λ é um autovalor de T, então todo $v \in V$*

tal que $T(v) = \lambda v$ é denominado um ***autovetor*** *de T associado ao autovalor λ.*

Observação 3.49.

1. *Se v é um autovetor de T associado à λ, então kv é um autovetor de T associado à λ, para todo $k \in \mathbb{K}^*$, pois*

$$T(kv) = kT(v) = k(\lambda v) = \lambda(kv).$$

2. *Se λ é um autovalor de T, denotamos por V_λ o conjunto de todos os autovetores de T associados à λ, juntamente com o vetor nulo, o qual é denominado o* ***autoespaço*** *de V correspondente ao autovalor λ, isto é,*

$$V_\lambda = \{v \in V;\ T(v) = \lambda v\}.$$

O leitor pode verificar que V_λ é um subespaço de V.

Teorema 3.50. *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 $(T - \lambda I)$ é singular (não inversível);*
3. *$det([T - \lambda I]) = 0$.*

Demonstração: Se λ é um autovalor de T, então existe $v \neq 0$ tal que $T(v) = \lambda v$. Assim,

$$\begin{aligned} T(v) = \lambda v \quad &\Leftrightarrow \quad T(v) = \lambda I(v) \\ &\Leftrightarrow \quad (T - \lambda I)(v) = 0. \end{aligned}$$

Logo, $v \in Nuc(T - \lambda I)$, com $v \neq 0$. Portanto, $T - \lambda I$ não é injetivo, consequentemente não inversível. Entretanto, segue do Corolário 3.46 que $T - \lambda I$ é não inversível se, e somente se, sua matriz associada $[T - \lambda I]$ é não inversível o qual ocorre se, e somente se, $det([T - \lambda I]) = 0$. ∎

Se $T : V \to V$ é um operador linear com $dim_{\mathbb{R}} V = n$, observe que

$$det([\lambda I - T]) = det(-[T - \lambda I]) = (-1)^n det([T - \lambda I]).$$

Assim, λ é um autovalor de T se, e somente se, $det([T - \lambda I]) = 0$, ou seja, se, e somente se, $det([\lambda I - T]) = 0$. Deste modo considere o polinômio

$$p(x) = det([xI - T]),$$

denominado o **polinômio característico** de T.

Note que, as raízes do polinômio característico são os autovalores de T e que p é um polinômio mônico de grau n.

Observe que podemos definir o conceito de autovalores e de polinômio característico de uma matriz quadrada arbitrária A de ordem n. Mais precisamente, um escalar λ é um autovalor de A se a matriz $(\lambda I - A)$ é não inversível, com I sendo a matriz identidade. E o polinômio característico de A é definido como sendo

$$p(x) = det(xI - A).$$

O polinômio característico foi definido como o determinante da matriz de T em alguma base. Mas será que esse conceito depende da base escolhida? E os autovalores de T? Sabemos que dadas duas bases β e β' de V, as matrizes $[T]_\beta$ e $[T]_{\beta'}$ são semelhantes. O próximo resultado vai nos auxiliar a responder as questões acima.

Proposição 3.51. *Matrizes semelhantes possuem o mesmo polinômio característico.*

Demonstração: Sejam A e B matrizes semelhantes de ordem n, ou seja, existe uma matriz inversível P de ordem n tal que $B = P^{-1}AP$. Logo

$$det(xI - B) = det(xI - P^{-1}AP) = det(P^{-1}PxI - P^{-1}AP).$$

Observe que dado uma matriz C qualquer de ordem n vale que $CI = IC$. Assim $xIC = xCI = CxI$, ou seja, xI comuta com qualquer matriz C. Portanto,

$$\begin{aligned} det(xI - B) &= det(P^{-1}xIP - P^{-1}AP) = det(P^{-1}(xI - A)P) \\ &= det(P^{-1})det(xI - A)det(P) = \tfrac{1}{det(P)}det(xI - A)det(P) \\ &= det(xI - A). \end{aligned}$$

■

Portanto, o polinômio característico não depende da base para o qual associamos a matriz do operador T. E como $[xI - T] = x[I] - [T]$ faz sentido definir o polinômio característico de T como o polinômio característico de qualquer $n \times n$ matriz que represente T em relação à alguma base ordenada de V.

Para determinar os autovetores de T vamos denotar as coordenadas de um vetor $v \in V$ na base canônica por $[v]$. Assim, v é um autovetor de T associado a um autovalor λ se $T(v) = \lambda v$, ou seja, $(\lambda I - T)(v) = 0$. Em notação matricial obtemos

$$[\lambda I - T] \cdot [v] = 0.$$

O número de vezes que $(x - \lambda)$ aparece como um fator do polinômio característico de T é denominado **multiplicidade algébrica** de λ, cuja notação é $m_a(\lambda)$. A dimensão do autoespaço V_λ é denominado **multiplicidade geométrica** de λ, cuja notação é $m_g(\lambda)$.

Exemplo 3.52. *Sejam $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por $T(x, y) = (3x, 8x - y)$ e $A = [T]$ a matriz de T na base canônica de $\mathbb{R}^2$.*

$$xI - A = \begin{pmatrix} x-3 & 0 \\ -8 & x+1 \end{pmatrix}.$$

O polinômio característico de T é $p(x) = det(xI - A) = (x-3)(x+1)$. Portanto, os autovalores de T são $\lambda_1 = 3$ e $\lambda_2 = -1$. Nesse caso, as multiplicidades algébricas dos autovalores são: $m_a(3) = 1 = m_a(-1)$.

Os autovetores $v = (x, y)$ associados ao autovalor $\lambda_1 = -1$ são:

$$(-I - A)[v] = \begin{pmatrix} -4 & 0 \\ -8 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Assim, obtemos $-4x = 0$ o que implica $x = 0$. Logo, o autoespaço associado ao autovalor $\lambda_1 = -1$ é dado por

$$V_{-1} = \{(0, y);\ y \in \mathbb{R}\} = \{y(0, 1);\ y \in \mathbb{R}\}.$$

Agora, os autovetores $v = (x, y)$ associados ao autovalor $\lambda_2 = 3$ são:

$$(3I - A)v = \begin{pmatrix} 0 & 0 \\ -8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

O que nos dá a equação $-8x + 4y = 0$, ou seja, $y = 2x$. Logo, o autoespaço associado ao autovalor $\lambda_2 = 3$ é dado por

$$V_3 = \{(x, 2x);\ x \in \mathbb{R}\} = \{x(1, 2);\ x \in \mathbb{R}\}.$$

As multiplicidades geométricas dos autovalores são: $m_g(3) = 1 = m_g(-1)$, pois $dim_{\mathbb{R}} V_{-1} = dim_{\mathbb{R}} V_3 = 1$.

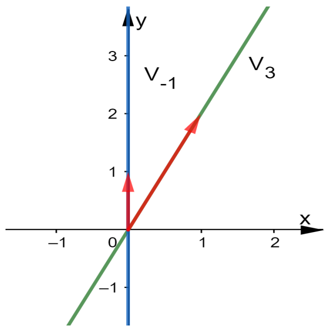

Figura 3.8: Autoespaços V_{-1} e V_3 de $T(x,y) = (3x, 8x-y)$

3.4 Diagonalização de Operadores Lineares

Nessa subseção vamos apresentar uma condição para que um operador linear seja diagonalizável, ou seja, que possamos representá-lo por meio de uma matriz diagonal.

Definição 3.53. *Seja T um operador linear sobre um espaço vetorial V de dimensão finita. Dizemos que T é diagonalizável se existe uma base de V formada por autovetores de T.*

Seja $\beta = \{v_1, v_2, \ldots, v_n\}$ uma base de autovetores de T, com v_i um autovetor associado a um autovalor λ_i, para $i \in \{1, \ldots, n\}$. Então

$$\begin{aligned} T(v_1) &= \lambda_1 v_1 = \lambda_1 v_1 + 0v_2 + \ldots + 0v_n \\ T(v_2) &= \lambda_2 v_2 = 0v_1 + \lambda_2 v_2 + \ldots + 0v_n \\ &\vdots \\ T(v_n) &= \lambda_n v_n = 0v_1 + 0v_2 + \ldots + \lambda_n v_n. \end{aligned}$$

Portanto, a matriz de T na base β é dada por:

$$[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Proposição 3.54. *Autovetores associados a autovalores distintos são LI.*

Demonstração: Sejam $v_1, v_2 \in V$ autovetores de T associados aos autovalores distintos λ_1 e λ_2, respectivamente. Sejam $k_1, k_2 \in \mathbb{K}$ tais que $k_1 v_1 + k_2 v_2 = 0$. Aplicando o operador linear $T - \lambda_2 I$ de ambos os lados, obtemos

$$\begin{aligned} (T-\lambda_2 I)(k_1 v_1 + k_2 v_2) = 0 \quad &\Leftrightarrow \quad k_1\lambda_1 v_1 + k_2\lambda_2 v_2 - k_1\lambda_2 v_1 - k_2\lambda_2 v_2 = 0 \\ &\Leftrightarrow \quad k_1(\lambda_1 - \lambda_2)v_1 = 0. \end{aligned}$$

Como $\lambda_1 \neq \lambda_2$ e $v_1 \neq 0$ temos que $k_1 = 0$. Aplicando $T - \lambda_1 I$ na equação $k_1v_1 + k_2v_2 = 0$ obtemos

$$(T - \lambda_1 I)(k_1v_1 + k_2v_2) = 0 \quad \Leftrightarrow \quad k_1\lambda_1v_1 + k_2\lambda_2v_2 - k_1\lambda_1v_1 - k_2\lambda_1v_2 = 0$$

implicando $k_2(\lambda_2 - \lambda_1)v_2 = 0$. Como $\lambda_1 \neq \lambda_2$ e $v_2 \neq 0$ temos que $k_2 = 0$. ∎

Corolário 3.55. *Se $dimV = n$ e $T : V \to V$ é um operador linear com n autovalores distintos, então T é diagonalizável.*

Demonstração: De fato, suponha que T possua n autovalores distintos, digamos $\lambda_1, \ldots, \lambda_n$. Sejam $v_1, \ldots, v_n$ autovetores associados associados a $\lambda_1, \ldots, \lambda_n$, respectivamente. Pelo resultado anterior, $\beta = \{v_1, \ldots, v_n\}$ é um conjunto LI. Como temos n vetores linearmente independentes em um espaço vetorial n dimensional, segue do Corolário 2.32 que β gera V e, portanto, é uma base de autovetores de V. Concluímos assim que T é diagonalizável. ∎

Exemplo 3.56. *Um operador T pode não ter autovalores. Por exemplo, seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ cuja matriz na base canônica é dada por*

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

O polinômio característico de T é

$$p(x) = det(xI - A) = \begin{vmatrix} x & 1 \\ -1 & x \end{vmatrix} = x^2 + 1$$

que não possui raízes reais. Logo, o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ não é diagonalizável sobre $\mathbb{R}$.

No entanto, se $V = \mathbb{C}^2$ é visto como um espaço vetorial sobre $\mathbb{C}$ e $U : \mathbb{C}^2 \to \mathbb{C}^2$ possui a mesma matriz A, então temos o mesmo polinômio característico que, nesse caso, se fatora da forma

$$p(x) = x^2 + 1 = (x - i)(x + i),$$

ou seja, U possui autovalores i e $-i$, que são distintos. Como $dim_{\mathbb{C}}\mathbb{C}^2 = 2$, o corolário anterior nos garante que $U : \mathbb{C}^2 \to \mathbb{C}^2$ é diagonalizável.

Mesmo assim vamos determinar os autoespaços associados a cada autovalor. Os autovetores

$v=(x,y)$ associados ao autovalor $\lambda_1 = i$ são:

$$(iI-A)[v] = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

O que implica

$$\begin{cases} ix + y = 0 \\ -x + iy = 0 \end{cases}$$

e, portanto, $y=-ix$. Assim o autoespaço associado ao autovalor $\lambda_1 = i$ é

$$V_i = \{(x,-ix);\ x \in \mathbb{C}\} = \{x(1,-i);\ x \in \mathbb{C}\}.$$

Analogamente para o autovalor $\lambda_2 = -i$ temos que

$$(iI-A)[v] = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Donde obtemos

$$\begin{cases} -ix + y = 0 \\ -x - iy = 0. \end{cases}$$

Logo, $y = ix$. Assim o autoespaço associado ao autovalor $\lambda_2 = -i$ é

$$V_{-i} = \{(x,ix);\ x \in \mathbb{C}\} = \{x(1,i);\ x \in \mathbb{C}\}.$$

Portanto, $\beta = \{(1,-i),\ (1,i)\}$ é uma base de autovetores de U de modo que

$$[U]_\beta = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Exemplo 3.57. *Seja $T:\mathbb{R}^3 \to \mathbb{R}^3$ cuja matriz na base canônica é dada por*

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

O operador T é diagonalizável?

Veja que o polinômio característico de T é o determinante de uma matriz que podemos dividir *em um bloco de ordem 2 e um de ordem 1:*

$$\begin{aligned} p(x) &= \begin{vmatrix} x-2 & 1 & 0 \\ 1 & x-2 & 0 \\ 0 & 0 & x-1 \end{vmatrix} \\ &= [(x-2)^2 - 1] \cdot (x-1) \\ &= (x-1)(x^2 - 4x + 3) \\ &= (x-1)^2(x-3), \end{aligned}$$

pois as raízes do polinômio $x^2 - 4x + 3$ *são* 1 *e* 3.

Neste caso, as multiplicidades algébricas de cada autovalor são:

$$m_a(1) = 2 \quad e \quad m_a(3) = 1.$$

Calculemos os autovetores associados a cada autovalor.

- $\lambda_1 = 1$

  Seja $v = (x, y, z) \in \mathbb{R}^3$, *deste modo:*

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

  Assim temos que encontrar a solução da equação $-x + y = 0$, *ou seja,* $y = x$. *Deste modo, o autoespaço associado ao autovalor* $\lambda_1 = 1$ *é*

$$V_1 = \{(x, x, z);\ x, z \in \mathbb{R}\} = \{x(1,1,0) + z(0,0,1);\ x, z \in \mathbb{R}\}.$$

  Assim $\beta_1 = \{(1,1,0), (0,0,1)\}$ *é uma base de* V_1, *com isso a multiplicidade geométrica do autovalor* 1 *é* $m_g(1) = dim_{\mathbb{R}} V_1 = 2$.

- $\lambda_2 = 3$

  Seja $v = (x, y, z) \in \mathbb{R}^3$, *deste modo:*

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Assim temos que encontrar a solução do sistema

$$\begin{cases} x + y = 0 \\ 2z = 0, \end{cases}$$

ou seja, $y = -x$ *e* $z = 0$*. Deste modo, o autoespaço associado a* $\lambda_2 = 3$ *é*

$$V_3 = \{(x, -x, 0);\ x \in \mathbb{R}\} = \{x(1, -1, 0);\ x \in \mathbb{R}\}.$$

Facilmente vemos que $\beta_2 = \{(1, -1, 0)\}$ *é uma base de* V_3*, donde concluímos que a multiplicidade geométrica do autovalor* 3 *é* $m_g(3) = 1$.

Como $\beta = \beta_1 \cup \beta_2 = \{(1, 1, 0), (0, 0, 1), (1, -1, 0)\}$ *é uma base de autovetores de* $\mathbb{R}^3$*, concluímos que* T *é um operador diagonalizável. Mais ainda,*

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Exercício 3.58. *Verifique quais dos operadores abaixo são diagonalizáveis.*

1. $T : \mathbb{R}^3 \to \mathbb{R}^3$ *cuja matriz é dada por*

$$[T] = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

2. $T : \mathbb{R}^3 \to \mathbb{R}^3$ *cuja matriz é dada por*

$$[T] = \begin{pmatrix} 3 & -3 & -4 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{pmatrix}.$$

3. $T : \mathbb{R}^3 \to \mathbb{R}^3$ *cuja matriz é dada por*

$$[T] = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & -2 \\ 3 & 2 & 1 \end{pmatrix}.$$

4. $T : \mathbb{C}^4 \to \mathbb{C}^4$ cuja matriz é dada por

$$[T] = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Uma propriedade interessante e útil do polinômio característico é dada pelo Teorema de Cayley-Hamilton.

Teorema 3.59. ***(Teorema de Cayley-Hamilton)*** *Seja T um operador linear sobre um espaço vetorial n-dimensional V. Se p é o polinômio característico de T, então $p(T) = 0$, ou ainda, se $A = [T]$, então $p(A) = 0$.*

Exemplo 3.60. *De modo a ilustrar o que afirma o Teorema de Cayley-Hamilton, retomamos o Exemplo 3.57 em que o polinômio característico é dado por $p = (x-1)^2(x-3)$. Verifique que a matriz A do operador nesse exemplo satisfaz*

$$p(A) = (A - I)^2 \cdot (A - 3I) = 0,$$

em que I é a matriz identidade.

APÊNDICE A

ALGUNS CONCEITOS ALGÉBRICOS

Esse apêndice tem por objetivo servir de consulta para uma revisão sobre o conjunto dos números complexos $\mathbb{C}$ e de resultados referentes ao cálculo de raízes de polinômios de qualquer grau. Iniciamos introduzindo os principais conceitos associados ao conjunto dos números complexos, como as operações de soma e produto de dois números complexos, bem como o módulo de um número complexo.

A segunda parte deste apêndice é dedicada a um importantíssimo teorema na matemática, chamado Teorema Fundamental da Álgebra, que garante que todo polinômio com coeficientes em $\mathbb{C}$ possui uma raiz complexa. Como consequência desse resultado, um polinômio com coeficientes em $\mathbb{C}$ possui tantas raízes quanto o grau desse polinômio. Nesta direção, vamos apresentar alguns resultados que garantem a existência de tais raízes e outros que nos ajudam a tentar determiná-las.

A.1 Números Complexos

Vamos fazer uma breve revisão de conceitos básicos sobre o conjunto dos números complexos, conjunto este que certamente provocou uma revolução na matemática, quebrando preconceitos e cujas aplicações na matemática, física, engenharia entre tantas outras áreas, mostram sua importância, que muitas vezes não é explorada com a devida atenção.

Historicamente, o conjunto dos números complexos surgem a partir da busca de raízes de polinômios de graus 3 e não de polinômios de grau 2, como muitas vezes nos é passado.

Vamos definir o **conjunto dos números complexos**, denotado por $\mathbb{C}$, como o conjunto dos elementos da forma $a + bi$, com $a, b \in \mathbb{R}$ e tal que $i^2 = -1$.

Observe que tomando $b = 0$, o conjunto dos números reais está contido em $\mathbb{C}$.

Dizemos que dois números complexos $a+bi$ e $c+di$ são iguais se, e somente se, $a = c$ e $b = d$, ou seja, um número complexo $z = a + bi$ está completamente determinado por a e b. Dizemos que a é a parte real de z, que será denotado por $Re(z)$, e b é a parte imaginária de z, o qual denotamos por $Im(z)$.

Vamos definir as seguintes operações em $\mathbb{C}$. Dados $z = a + bi$ e $w = c + di$ em $\mathbb{C}$, definimos a soma $z + w$ e o produto $z \cdot w$ por:

$$z + w = a + bi + c + di = a + c + (b + d)i \quad \text{e}$$

$$z \cdot w = (a + bi) \cdot (c + di) = ac - bd + (ad + bc)i.$$

William Rowan Hamilton (1805-1865) foi um matemático, físico e astrônomo irlandês que deu uma elegante abordagem dos números complexos como pares de números reais, dando uma interpretação geométrica para este conceito. O conjunto dos números complexos estão em correspondência biunívoca com pares ordenados de números reais, ou seja, como $z = a + bi$ está completamente determinado por a e b, podemos associar o par $(a, b) \in \mathbb{R}^2$ e vice-versa.

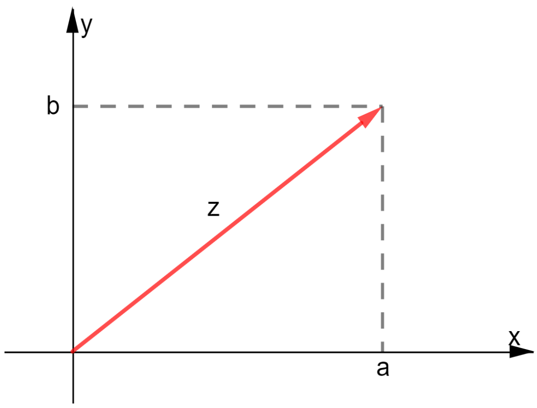

Figura A.1: Representação geométrica do número complexo $z = a + bi$

Vejamos algumas propriedades sobre as operações definidas acima.

- **<u>Adição</u>**

  Dados $z_1, z_2, z_3 \in \mathbb{C}$ as seguintes propriedades são satisfeitas:

  A.1 Associatividade: $z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3$;

  A.2 Comutatividade: $z_1 + z_2 = z_2 + z_1$;

  A.3 A adição admite um único elemento neutro, $0 = 0 + 0i$, ou seja, $0 + z = z$ para todo $z \in \mathbb{C}$;

A.4 Para todo $z = a + bi \in \mathbb{C}$ existe um único $w \in \mathbb{C}$ tal que $z + w = 0$. Dizemos que w é o oposto de z e denotamos $w = -z = -a - bi$.

- **Diferença**

  Sejam $z_1 = a + bi$ e $z_2 = c + di$ em $\mathbb{C}$. Definimos a diferença entre z_1 e z_2, denotada por $z_1 - z_2$, como

$$z_1 - z_2 = z_1 + (-z_2).$$

- **Multiplicação**

  Dados $z_1, z_2, z_3 \in \mathbb{C}$ as seguintes propriedades são satisfeitas:

  M.1 Associatividade: $z_1 \cdot (z_2 \cdot z_3) = (z_1 \cdot z_2) \cdot z_3$;

  M.2 Comutatividade: $z_1 \cdot z_2 = z_2 \cdot z_1$;

  M.3 Distributividade: $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$;

  M.4 A multiplicação admite a existência de um elemento identidade $1 = 1 + 0i$, ou seja, um elemento com a propriedade de que $1 \cdot z = z$ para todo $z \in \mathbb{C}$. O elemento identidade é único.

Os conceitos de módulo e conjugado de um número complexo serão essenciais para definirmos o inverso de um número complexo.

Dado um número complexo $z = a + bi$ vamos definir $\overline{z} = a - bi$ como o conjugado de z. O conjugado $\overline{z}$ de z é representado pelo simétrico de z em relação ao eixo Ox.

Vamos definir o módulo de $z = a + bi$ como sendo o número real não negativo dado por

$$|z| = \sqrt{a^2 + b^2}.$$

Geometricamente, $|z|$ mede a distância da origem 0 até z, isto é, mede o módulo do vetor (a, b) que representa o número complexo $z = a + bi$ (veja Figura A.1).

Temos a seguinte relação entre módulo e conjugado de um número complexo:

$$z \cdot \overline{z} = (a + bi) \cdot (a - bi) = a^2 + b^2 = |z|^2.$$

Se $z \neq 0$, vamos denotar por z^{-1} ou $\frac{1}{z}$ o número complexo $\frac{1}{|z|^2} \cdot \overline{z}$. E veja que

$$z \cdot z^{-1} = z \cdot \frac{1}{|z|^2} \cdot \overline{z} = \frac{1}{|z|^2} \cdot z \cdot \overline{z} = \frac{1}{|z|^2} \cdot |z|^2 = 1,$$

ou seja, $z^{-1} = \frac{1}{|z|^2} \cdot \overline{z}$ é de fato o **inverso** de z.

Desta forma, dado dois números complexos z_1 e z_2 com $z_2 \neq 0$, podemos definir o quociente de z_1 por z_2 como sendo

$$\frac{z_1}{z_2} = z_1 \cdot \frac{1}{z_2}.$$

Exemplo A.1. *Calculemos o inverso do número complexo* $z = 2 - 3i$. *Veja que* $|z|^2 = 2^2 + (-3)^2 = 13$. *Deste modo*

$$z^{-1} = \frac{1}{|z|^2} \cdot \overline{z} = \frac{1}{13} \cdot (2 + 3i) = \frac{2}{13} + \frac{3}{13}i.$$

Podemos facilmente verificar as seguintes propriedades:

Proposição A.2. *Sejam* z_1 *e* z_2 *números complexos, então*

a) $(\overline{z_1 + z_2}) = \overline{z_1} + \overline{z_2}$;

b) $(\overline{z_1 \cdot z_2}) = \overline{z_1} \cdot \overline{z_2}$.

Corolário A.3. *Dados* $z_1, z_2 \in \mathbb{C}$ *então* $|z_1 z_2| = |z_1| \cdot |z_2|$.

Demonstração: Da relação $z \cdot \overline{z} = |z|^2$ e pela proposição acima temos

$$\begin{aligned} |z_1 z_2|^2 &= z_1 z_2 \cdot \overline{z_1 z_2} = z_1 z_2 \cdot \overline{z_1} \cdot \overline{z_2} \\ &= z_1 \overline{z_1} \cdot z_2 \overline{z_2} = |z_1|^2 \cdot |z_2|^2 = (|z_1| \cdot |z_2|)^2. \end{aligned}$$

Extraindo a raiz quadrada de ambos os lados, obtemos que $|z_1 z_2| = |z_1| \cdot |z_2|$. ■

A.2 Teorema Fundamental da Álgebra

O matemático alemão Carl Friedrich Gauss (1777-1855) é considerado o maior matemático de todos os tempos. Dentre tantas de suas contribuições para a matemática, o Teorema Fundamental da Álgebra é um dos resultados mais importantes da teoria das equações algébricas. Apesar de ser um resultado da Álgebra, Gauss utilizou propriedades topológicas da reta e do plano para dar uma demonstração deste resultado.

É importante ressaltar que a busca de raízes complexas de um polinômio com coeficientes complexos (em particular, com coeficientes em $\mathbb{R}$) é essencial, visto que um polinômio com coeficientes reais pode não admitir nenhuma raiz real, como podemos facilmente verificar que o polinômio $x^2 + 1$ não possui nenhuma raiz real, ou seja, não existe solução real para a equação

$$x^2 + 1 = 0.$$

Dado um polinômio

$$p = a_n x^n + \ldots + a_1 x + a_0,$$

com $a_i \in \mathbb{C}$, dizemos que um número $k \in \mathbb{C}$ é uma raiz de p, se k satisfaz a equação $p = 0$, ou seja,

$$a_n k^n + \ldots + a_1 k + a_0 = 0.$$

Teorema A.4 (**Teorema Fundamental da Álgebra**). *Toda equação polinomial dada da forma*

$$a_n x^n + \ldots + a_1 x + a_0 = 0,$$

com $a_i \in \mathbb{C}$ para $i = 0, 1, \ldots, n$ possui uma raiz complexa.

Corolário A.5. *Toda equação polinomial de grau n, possui exatamente n raízes complexas, contando suas multiplicidades.*

Demonstração: Seja um polinômio $p = a_n x^n + \ldots + a_1 x + a_0$, com $a_i \in \mathbb{C}$ para $i = 0, 1, \ldots, n$ de grau n. Pelo Teorema Fundamental da Álgebra a equação $p = 0$ admite uma raiz complexa, digamos r_1. Assim pelo teorema de divisão de polinômios, podemos escrever $p = (x - r_1)p_1$ para algum polinômio p_1 de grau $n - 1$. Aplicando novamente o Teorema Fundamental da Álgebra, sabemos que $p_1 = 0$ admite uma raiz complexa, digamos r_2. Logo, $p = (x - r_1)(x - r_2)p_2$, para algum polinômio p_2 de grau $n - 2$. Continuando este processo, concluímos que

$$p = c(x - r_1)(x - r_2) \cdots (x - r_t),$$

onde c é uma constante e $r_1, \ldots, r_t$ são as raízes complexas de p, contadas com suas multiplicidades. ∎

O Teorema Fundamental da Álgebra nos diz que um polinômio de grau n admite em $\mathbb{C}$, exatamente n raízes complexas, entretando não nos dá uma pista de como obtê-las. Na verdade, determinar as raízes de um polinômio é algo bem complicado.

O resultado abaixo nos auxilia na busca de raízes racionais de polinômios com coeficientes racionais.

Primeiramente considere o seguinte polinômio com coeficientes racionais

$$q = \frac{b_n}{c_n} x^n + \ldots + \frac{b_1}{c_1} x + \frac{b_0}{c_0},$$

com $c_i \neq 0$ para todo $i = 0, 1, \ldots, n$. Multiplicando o polinômio acima por $c_n \cdots c_1 c_0$ temos

$$p = (c_n \cdots c_1 c_0) q = a_n x^n + \ldots + a_1 x + a_0$$

com $a_i = c_n \cdots c_{i-1} c_{i+1} \cdots c_1 c_0 \cdot b_i \in \mathbb{Z}$. Observe que α é raiz de q se, e somente se, α é raiz de p. Deste modo, com o intuito de procurar raízes de polinômios com coeficientes racionais, basta analisar o caso de polinômios com coeficientes inteiros. Para o que segue, dados dois números inteiros a e b, dizemos que a divide b se existe $\lambda \in \mathbb{Z}$ tal que

$$b = \lambda \cdot a,$$

ou seja, a é um divisor de b. Neste caso, denotamos $a|b$. Por exemplo, $5|75$, pois $75 = 15 \cdot 5$.

Se a não dividir b utilizamos a notação $a \nmid b$. Por exemplo, $3 \nmid 14$, pois não existe $\lambda \in \mathbb{Z}$ satisfazendo $14 = \lambda \cdot 3$.

Proposição A.6. *Seja $p = a_n x^n + \ldots + a_1 x + a_0$ um polinômio com coeficientes inteiros. Se $\frac{\alpha}{\beta} \in \mathbb{Q}$ com $mdc(\alpha, \beta) = 1$ é raiz de p, então $\alpha | a_0$ e $\beta | a_n$.*

Demonstração: Suponha que $\frac{\alpha}{\beta} \in \mathbb{Q}$ raiz de p. Vamos denotar por $p(\frac{\alpha}{\beta})$ o elemento

$$p\left(\frac{\alpha}{\beta}\right) = a_n \left(\frac{\alpha}{\beta}\right)^n + \ldots + a_1 \left(\frac{\alpha}{\beta}\right) + a_0.$$

Assim, por hipótese, $0 = p(\frac{\alpha}{\beta}) = a_n(\frac{\alpha}{\beta})^n + \ldots + a_1(\frac{\alpha}{\beta}) + a_0$. Deste modo,

$$0 = \beta^n \cdot p\left(\frac{\alpha}{\beta}\right) = a_n \alpha^n + a_{n-1} \alpha^{n-1} \beta \ldots + a_1 \alpha \beta^{n-1} + a_0 \beta^n. \tag{A.1}$$

Logo, $-a_n \alpha^n = \beta(a_{n-1} \alpha^{n-1} \ldots + a_1 \alpha \beta^{n-2} + a_0 \beta^{n-1})$, ou seja, β divide $a_n \alpha^n$. Como $mdc(\alpha, \beta) = 1$ temos que $\beta \nmid \alpha$ donde segue que $\beta | a_n$. Analogamente, segue de (A.1) que

$$-a_0 \beta^n = \alpha(a_n \alpha^{n-1} + a_{n-1} \alpha^{n-2} \beta + \ldots + a_1 \beta^{n-1}).$$

Logo, $\alpha | a_0 \beta^n$, mas como α e β são primos entre si, segue que $\alpha | a_0$. ■

Exemplo A.7. *1. Determinemos as soluções da equação $q = 0$ com $q = 2x^3 - 8x^2 + 5x + 1$.*

Pelo resultado anterior, se $\frac{\alpha}{\beta}$ for uma raiz racional de q, então $\alpha | 1$ e $\beta | 2$, ou seja, as possibilidades para α são ± 1 e para β são ± 1 e ± 2. Deste modo, as possíveis raízes *racionais de q são: ± 1 e $\pm\frac{1}{2}$. Note que 1 é raiz de q, pois*

$$2 \cdot 1^3 - 8 \cdot 1^2 + 5 \cdot 1 + 1 = 0.$$

Assim, dividindo q pelo polinômio $x-1$, obtemos $q = (x-1)(2x^2 - 6x - 1)$. Agora basta determinar as raízes do polinômio de grau dois $r = 2x^2 - 6x - 1$, que são $\frac{3\pm\sqrt{11}}{2}$. Portanto,

$$\begin{aligned} q &= 2x^3 - 8x^2 + 5x + 1 \;=\; (x-1)(2x^2 - 6x - 1) \\ &= (x-1) \cdot \left[2\left(x - \tfrac{3+\sqrt{11}}{2}\right)\left(x - \tfrac{3-\sqrt{11}}{2}\right)\right] \\ &= 2(x-1)\left(x - \tfrac{3+\sqrt{11}}{2}\right)\left(x - \tfrac{3-\sqrt{11}}{2}\right). \end{aligned}$$

2. *Observe que para o polinômio $p = 6x^3 - 5x^2 - 3x - 2$, os divisores de $a_0 = -2$ são ± 1 e ± 2 e os divisores de $a_3 = 6$ são $\pm 1, \pm 2, \pm 3$ e ± 6. Assim, as possíveis raízes racionais de p, se existirem, são: $\pm 1, \pm 2, \pm\frac{1}{2}, \pm\frac{1}{3}, \pm\frac{1}{6}$ e $\pm\frac{2}{3}$. Verifique se algum destes valores é uma raiz do polinômio p.*

3. *Aplique os resultados acima para determinar as soluções da equação $p = 0$ em que*

$$p = x^5 - 10x^4 + 54x^3 - 132x^2 + 137x - 50.$$

Pelo resultado acima se existirem raízes racionais elas serão os divisores de -50, ou seja, $\pm 1, \pm 2, \pm 5, \pm 10, \pm 25$ e ± 50. Verifique que

$$p = (x-1)^2(x-2)(x-(3+4i))(x-(3-4i)).$$

O próximo resultado nos diz que as raízes complexas de um polinômio com coeficientes reais satisfazem uma propriedade especial.

Corolário A.8. *Seja $p = a_n x^n + \ldots + a_1 x + a_0$ um polinômio com coeficientes reais. Se $z = a + bi \in \mathbb{C}$ é uma raiz de p, então $\overline{z}$ também o é.*

Demonstração: De fato, se z é raiz de p, então $a_n z^n + \ldots + a_1 z + a_0 = 0$. O que implica que $\overline{a_n z^n + \ldots + a_1 z + a_0} = \overline{0}$. Donde segue de uma proposição acima que

$$\overline{a_n}\,\overline{z^n} + \ldots + \overline{a_1 z} + \overline{a_0} = \overline{0}.$$

Todavia $a_i \in \mathbb{R}$ para todo $i = 0, 1, \ldots, n$. Logo,

$$a_n \overline{z}^n + \ldots + a_1 \overline{z} + a_0 = 0,$$

ou seja, $\overline{z}$ é uma raiz de p. ■

Podemos observar nos exemplos anteriores a validade do resultado anterior.

Corolário A.9. *Toda equação polinomial com coeficientes reais de grau ímpar, tem pelo menos uma raiz real.*

Demonstração: Segue diretamente do corolário anterior que as raízes complexas, com parte imaginária não nula, de um polinômio p aparecem aos pares, ou seja, se $z = a + bi$ é raiz de p, com $b \neq 0$, então $\overline{z} = a - bi$ também é raiz de p. Como o grau do polinômio é ímpar, segue o resultado.

■

Observe que o resultado acima não é válido para polinômios com coeficientes complexos.

Exemplo A.10. 1. *O polinômio $p = x - i$ possui uma única raiz $x = i$.*

2. *Dado $q = x^2 - ix + 1 = 0$ é fácil verificar que*

$$x = \frac{i \pm \sqrt{(-i)^2 - 4}}{2} = \frac{i \pm \sqrt{-5}}{2} = \frac{i \pm i\sqrt{5}}{2}$$

são as raízes de q, ou seja, $x_1 = \frac{i}{2}(1 + \sqrt{5})$ e $x_2 = \frac{i}{2}(1 - \sqrt{5})$. Mas observe que uma raiz não é a conjugada da outra.

A.3 Corpos

Neste texto consideramos espaços vetoriais sobre o conjunto $\mathbb{R}$ ou sobre $\mathbb{C}$. No entanto, podemos considerar espaços vetoriais tomando os escalares em um conjunto com boas propriedades. Na verdade todos os resultados anteriores são válidos para um espaço vetorial sobre um conjunto chamado corpo. No que segue vamos definir o que denominamos de corpo, o qual é uma estrutura algébrica muito importante.

Definição A.11. *Um conjunto $\mathbb{K}$ munido de uma operação de adição $(+)$ e uma operação de multiplicação $(\cdot)$, será chamado de* ***corpo****, se as seguintes condições são satisfeitas:*

A1 Associatividade: $a + (b + c) = (a + b) + c$, para todos $a, b, c \in \mathbb{K}$;

A2 Comutatividade: $a + b = b + a$, para todos $a, b \in \mathbb{K}$;

A3 A adição admite um elemento neutro $0 \in \mathbb{K}$, ou seja, $0 + a = a$ para todo $a \in \mathbb{K}$;

A4 *Para todo* $a \in \mathbb{K}$ *existe* $-a \in \mathbb{K}$ *de modo que* $a + (-a) = 0$. *Dizemos que* $-a$ *é o oposto do elemento* a*;*

M1 *Associatividade:* $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, *para todos* $a, b, c \in \mathbb{K}$*;*

M2 *Comutatividade:* $a \cdot b = b \cdot a$, *para todos* $a, b \in \mathbb{K}$*;*

M3 *Distributividade:* $a \cdot (b + c) = a \cdot b + a \cdot c$, *para todos* $a, b, c \in \mathbb{K}$*;*

M4 *A multiplicação possui um elemento identidade* $1 \in \mathbb{K} \setminus \{0\}$, *ou seja,* $a \cdot 1 = a$, *para todo* $a \in \mathbb{K}$*;*

M5 *Para todo* $a \in \mathbb{K} \setminus \{0\}$, *existe* $a^{-1} \in \mathbb{K}$ *tal que* $a \cdot a^{-1} = 1$. *Dizemos que* a^{-1} *é o inverso de* a.

Os elementos neutro e identidade de $\mathbb{K}$ *são únicos.*

Os conjuntos $\mathbb{Q}, \mathbb{R}$ e $\mathbb{C}$ são exemplos de corpos e estes foram os corpos que exploramos nestas notas.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] Anton, H.; Rorres, C., *Álgebra Linear com Aplicações.* $8^{a.}$ Edição. Bookman. Porto Alegre, 2001.

[2] Boldrini, J.L.; Costa, S.I.R.; Figueiredo, V.L.; Wetzler, H.G., *Álgebra Linear.* $3^{a.}$ Edição. Editora Harbra Ltda. São Paulo, 1986.

[3] Boyer, C.B., *História da Matemática.* Ed. Edgard Blucher Ltda. São Paulo, 1974.

[4] Callioli, C.A.; Domingues, H.H.; Costa, R.C.F., *Álgebra Linear e Aplicações.* $6^{a.}$ Edição. Editora Atual, São Paulo, 1991.

[5] Coelho, F.U., *Um Curso de Álgebra Linear*, Editora EDUSP, 2001.

[6] Eves, H., *Introdução à História da Matemática.* Unicamp, Campinas, 1995.

[7] Lipschutz, S., *Álgebra Linear.* $3^{a.}$ Edição, Makron Books, São Paulo, 1994.

[8] Lima, E.L., *Geometria Analítica e Álgebra Linear.* Coleção Matemática Universitária. SBM. Rio de Janeiro, 2001.

[9] Poole, D., *Álgebra Linear.* Thomson, São Paulo, 2006.