

Illum Escola de Ciência

# Notas de Matemática: Álgebra Linear

Prof. Dr. Vinícius Francisco Wasques  
Prof. Dr. Mario Martinez

28 de julho de 2022, Campinas

# Sumário

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Sistemas Lineares</b>                              | <b>2</b>  |
| 1.1      | Escalonamento   | 4         |
| <b>2</b> | <b>Introdução aos Espaços Vetoriais</b>               | <b>13</b> |
| 2.1      | Representação de vetores                              | 13        |
| 2.2      | Representação de pontos e vetores                     | 13        |
| 2.3      | Operações entre vetores                               | 15        |
| 2.4      | Distâncias e ângulos                                  | 18        |
| <b>3</b> | <b>Combinação Linear</b>                              | <b>22</b> |
| 3.1      | Conjunto gerador                                      | 24        |
| 3.2      | Bases e coordenadas                                   | 26        |
| <b>4</b> | <b>Transformações Lineares</b>                        | <b>28</b> |
| 4.1      | Transformação Linear                                  | 28        |
| 4.2      | Núcleo e Imagem de uma Transformação Linear           | 34        |
| 4.3      | Operadores lineares                                   | 37        |
| <b>5</b> | <b>Autovalores e autovetores</b>                      | <b>40</b> |
| 5.1      | Autovalores e autovetores                             | 40        |
| <b>6</b> | <b>Operadores lineares auto-adjuntos e ortogonais</b> | <b>47</b> |
| 6.1      | Operadores ortogonais                                 | 48        |
| 6.2      | Operadores auto-adjuntos                              | 48        |

# Introdução

Este texto traz um breve resumo do conteúdo de **Álgebra Linear**, estudado pela turma 22 da Ilum Escola de Ciência. Esses tópicos abordam problemas e conceitos na área de Geometria Analítica e Álgebra Linear.

Inicialmente esse material servirá como guia e também um resumo do conteúdo já visto pelos alunos. Os exemplos e exercícios propostos aqui são, em sua maioria, os mesmos discutidos em sala de aula. Este texto é baseado, principalmente, nas seguintes referências [Martínez \(2004\)](#); [Boldrini \(1984\)](#); [Callioli, Domingues and Costa \(1983\)](#).

# Capítulo 1

## Sistemas Lineares

Começaremos este texto tratando de sistemas lineares. Muitos dos problemas que veremos neste curso recaem em resoluções de sistemas lineares. Além disso, muitos problemas da realidade podem ser modelados como sistemas de equações lineares. Por exemplo, uma indústria de alimentos deve fabricar uma ração para um certo tipo de animal, com determinados requisitos em termos nutricionais. Cada quilograma de ração deve conter  $P$  gramas de proteína,  $G$  gramas de gordura e  $C$  gramas de carboidratos. A ração é elaborada usando um conjunto de ingredientes disponíveis no mercado, digamos, arroz, soja, ossos, miúdos de boi, banha de porco e peixe.

As quantidades, em gramas, de cada um dos ingredientes são denotados por  $x_1, x_2, x_3, x_4, x_5, x_6$ . Isto é,  $x_1$  é a quantidade de arroz,  $x_2$  a quantidade de soja e assim por diante. Cada ingrediente tem os três nutrientes em diferentes proporções. Portanto, as equações que representam a necessidade de que as exigências nutricionais sejam cumpridas são:

$$\begin{aligned} p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4 + p_5x_5 + p_6x_6 &= P, \\ g_1x_1 + g_2x_2 + g_3x_3 + g_4x_4 + g_5x_5 + g_6x_6 &= G, \\ c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 + c_6x_6 &= C. \end{aligned} \tag{1.1}$$

em que cada  $p_i$  representa a fração de proteína,  $g_i$  é a fração de gordura e  $c_i$  a fração de carboidratos, para cada  $i = 1, \dots, 6$ . Além disso:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1000,$$

pois a soma das quantidades dos ingredientes deve ser 1000 gramas.

Em outras palavras, esse sistema é composto por quatro equações e seis incógnitas ( $x_1, x_2, x_3, x_4, x_5$  e  $x_6$ ). Encontrando os valores das incógnitas que satisfaçam as quatro equações do sistema acima, e fabricarmos cada quilograma de ração misturando  $x_1$  gramas do ingrediente 1,  $x_2$  gramas do ingrediente 2, e assim por diante, obteremos um produto aceitável do ponto de vista nutricional. Dessa forma, um *sistema linear* é um conjunto de equações lineares, podendo ter várias incógnitas e várias equações que devem ser satisfeitas simultaneamente.

O sistema (1.1) não tem uma única solução, isto é, os ingredientes podem ser misturados em diferentes proporções para obter os resultados adequados. Mas como chegamos a esta conclusão? Vamos apresentar uma abordagem para solucionar um sistema linear de  $m$  equações com  $n$  incógnitas. Frisamos que os números de equações e de incógnitas são arbitrários. Nada impede, por exemplo que um sistema de equações esteja constituído por única equação e várias incógnitas.

Consideremos o seguinte exemplo:

$$x_1 - x_2 + 3x_3 + 2x_4 = 1 \quad (1.2)$$

$$-x_2 + 5x_3 + 3x_4 = 0 \quad (1.3)$$

Trata-se de um sistema com 2 equações e 4 incógnitas. Colocando  $x_2$  em evidência na segunda equação, obtemos

$$x_2 = 5x_3 + 3x_4. \quad (1.4)$$

Substituindo na primeira equação, temos:

$$x_1 - 5x_3 - 3x_4 + 3x_3 + 2x_4 = 1,$$

ou seja:

$$x_1 = 2x_3 + x_4 + 1. \quad (1.5)$$

Por (1.4) e (1.5) vemos que, dando valores arbitrários a  $x_3$  e  $x_4$  podemos calcular  $x_1$  e  $x_2$  e, assim, obter diferentes soluções do sistema original. Por exemplo, se  $x_3 = x_4 = 0$ , temos  $x_1 = 1, x_2 = 0$ . Portanto, a "4-upla"  $(x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0)$  é uma solução do sistema original, o que pode ser facilmente verificado substituindo em (1.2) e (1.3). Se  $x_3 = 1, x_4 = 2$  obtemos  $x_1 = 5, x_2 = 11$ . Com efeito, é fácil ver que a 4-upla  $(x_1 = 5, x_2 = 11, x_3 = 1, x_4 = 2)$  é uma solução do sistema original.

Dessa maneira, podemos obter infinitas soluções deste sistema. Como existem dois parâmetros aos quais podemos dar livremente valores para obter diferentes soluções, dizemos que o sistema tem dois *graus de liberdade*<sup>1</sup>.

Uma maneira simples de representar *todas* as soluções do sistema é mediante sua *solução geral*. Por (1.4) e (1.5), a solução geral deste sistema é

$$\begin{pmatrix} 2x_3 + x_4 + 1 \\ 5x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix}$$

Para expressar esta solução geral de maneira ainda mais clara, vamos usar as regras:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} a + \alpha \\ b + \beta \\ c + \gamma \\ d + \delta \end{pmatrix} \quad (1.6)$$

e

$$\lambda \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \\ \lambda c \\ \lambda d \end{pmatrix}. \quad (1.7)$$

<sup>1</sup>O grau de liberdade (número de variáveis livres) de um sistema linear escalonado é o número de variáveis menos o número de linhas não nulas

Então, temos:

$$\begin{aligned} \begin{pmatrix} 2x_3 + x_4 + 1 \\ 5x_3 + 3x_4 \\ x_3 \\ x_4 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2x_3 \\ 5x_3 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} x_4 \\ 3x_4 \\ 0 \\ x_4 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ 5 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Esta é uma segunda forma de expressar a solução geral do sistema.

**Observação 1.0.1** *É também comum representar o sistema (1.1) da seguinte forma:*

$$\begin{cases} x_1 - x_2 + 3x_3 + 2x_4 = 1 \\ -x_2 + 5x_3 + 3x_4 = 0 \end{cases}$$

*O símbolo de chaves que aparece antes das equações representa que ambas equações fazem parte do mesmo problema, devendo ser satisfeitas simultaneamente.*

**Exercício 1.0.1** *Escreva a solução geral dos seguintes sistemas, indicando quais são as variáveis livres e quantos graus de liberdade existem.*

1.

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

2.

$$x + y + z + w = 10$$

## 1.1 Escalonamento

Obter a solução geral do sistema (1.2–1.3) foi bastante fácil. O que facilitou o processo foi o fato de que a variável  $x_1$  não aparecia na equação (1.3). Sempre que, em cada equação de um sistema linear, apareça uma variável que não aparece em nenhuma das equações seguintes, o sistema será fácil de resolver. Sistemas com essas características se dizem *sistemas lineares escalonados*<sup>2</sup>.

Quando um sistema não é naturalmente escalonado a melhor estratégia para resolvê-lo é transformá-lo em um sistema equivalente escalonado. Esse processo se chama *Eliminação Gaussiana*. A ferramenta fundamental para poder fazer isso é dada pelo seguinte teorema.

**Teorema 1.1.1** *Suponha que, em um sistema linear, substituimos uma de suas equações por “ela mesma vezes um número diferente de zero mais um múltiplo de outra”. Então, o sistema resultante tem as mesmas soluções que o original.*

<sup>2</sup>Sejam cuidadosos: dizemos que uma variável “aparece” em uma equação quando o coeficiente pela qual está multiplicada é diferente de zero. Por exemplo, a variável  $x_1$  aparece na equação  $3x_1 + x_2 - x_3 = 8$  mas “não aparece” na equação  $0x_1 + x_2 - x_3 = 9$ .

Vejamos como aplicar o **Teorema 1.1.1** no seguinte exemplo.

**Exemplo 1.1.1** Considere o sistema linear de 3 equações com 4 incógnitas abaixo.

$$\begin{cases} 3x_1 + 2x_2 - x_3 - x_4 = 4 \\ x_1 - x_2 + 2x_3 - 3x_4 = 2 \\ 3x_1 + 3x_2 - 3x_3 - 2x_4 = 1. \end{cases}$$

A variável  $x_1$  aparece na primeira equação. Para que deixe de aparecer na segunda equação podemos substituir esta equação por “3 vezes ela mesma menos a primeira equação”. Imediatamente, substituindo a terceira equação por “ela mesma menos a primeira equação” obtemos o sistema equivalente

$$\begin{cases} 3x_1 + 2x_2 - x_3 - x_4 = 4 \\ -5x_2 + 7x_3 - 8x_4 = 2 \\ x_2 - 2x_3 - x_4 = -3. \end{cases}$$

Neste novo sistema, a variável  $x_2$  aparece na segunda equação. Para “eliminar”  $x_2$  da terceira equação, podemos substituir a terceira equação do novo sistema por “5 vezes ela mesma mais a segunda equação”. Assim, obtemos o sistema equivalente:

$$\begin{cases} 3x_1 + 2x_2 - x_3 - x_4 = 4 \\ -5x_2 + 7x_3 - 8x_4 = 2 \\ -3x_3 - 13x_4 = -13 \end{cases} .$$

Este sistema já é escalonado. Portanto, pode ser resolvido pelo processo mostrado anteriormente. Mais precisamente, podemos:

- (a) Pôr em evidência  $x_3$  na terceira equação e substituir nas duas primeiras equações;
- (b) Pôr em evidência  $x_2$  na segunda equação e substituir na primeira equação;
- (c) Pôr em evidência  $x_1$  na primeira equação.

Desta maneira,  $x_1$ ,  $x_2$  e  $x_3$  ficarão em função de  $x_4$ .

Às vezes, no processo de escalonamento, pode ser completamente eliminada uma equação ou pode ser detectado que o sistema não tem nenhuma solução. Vejamos exemplos de ambas situações.

**Exemplo 1.1.2** Considere o sistema

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ 3x_1 - x_2 - 6x_3 = 2 \\ 8x_1 + x_2 - 7x_3 = 3 \end{cases} .$$

Substituindo a segunda equação por “ela mesma vezes 5 menos a primeira vezes 3” e a terceira equação por “ela mesma vezes cinco menos a primeira vezes 8” obtemos

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ -11x_2 - 27x_3 = 7 \\ -11x_2 - 27x_3 = 7. \end{cases}$$

Agora, para eliminar  $x_2$  da terceira equação substituímos esta equação por “ela mesma menos a segunda”. Assim, obtemos

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ -11x_2 - 27x_3 = 7 \\ 0 = 0. \end{cases}$$

A terceira equação ficou sem incógnitas e, certamente, pode ser eliminada, pois a proposição  $0 = 0$  é uma tautologia<sup>3</sup> que não oferece nenhuma informação. Portanto, o sistema sem essa equação já ficou escalonado.

**Exemplo 1.1.3** Considere o sistema

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ 3x_1 - x_2 - 6x_3 = 2 \\ 8x_1 + x_2 - 7x_3 = 9 \end{cases} .$$

Substituindo, como no exemplo anterior, a segunda equação por “ela mesma vezes 5 menos a primeira vezes 3” e a terceira equação por “ela mesma vezes cinco menos a primeira vezes 8” obtemos

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ -11x_2 - 27x_3 = 7 \\ -11x_2 - 27x_3 = 37 \end{cases} .$$

Finalmente, para eliminar  $x_2$  da terceira equação substituímos esta equação por “ela mesma menos a segunda”. Assim, obtemos

$$\begin{cases} 5x_1 + 2x_2 - x_3 = 1 \\ -11x_2 - 27x_3 = 7 \\ 0 = 30. \end{cases}$$

Este sistema é equivalente ao colocado inicialmente e, obviamente, não tem nenhuma solução pois a “equação”  $0 = 30$  não se cumpre nunca.

O processo de escalonamento pode ser descrito, em geral, pelo seguinte *algoritmo*.

### Algoritmo 1.1

Suponha que o sistema está formado por  $m$  equações.

Para  $i = 1, \dots, m - 1$ , execute o Passo 1. **Passo 1.**

1. Se em nenhuma das equações  $i, \dots, m$  aparece nenhuma variável, termine a execução do **Algoritmo 1.1**. Nesse caso, o escalonamento do sistema está finalizado.
2. Se na equação  $i$  não aparece nenhuma variável mas em uma equação posterior aparece alguma, permuta a equação  $i$  com esta equação (que agora passará a ser a nova equação  $i$ ).
3. Escolha uma das variáveis que aparecem na (talvez nova) equação  $i$ . Chamemos  $x$  a essa variável. Para cada  $j = i + 1, \dots, m$ , substitua a equação  $j$  por um múltiplo diferente de zero dela mesma mais um múltiplo da equação  $i$ , de maneira a eliminar a variável  $x$  da equação  $j$ .

<sup>3</sup>Na lógica proposicional, uma tautologia é uma fórmula proposicional que é verdadeira para todos os possíveis valores de suas variáveis proposicionais.



Ao terminar a execução do Algoritmo 1.1 o sistema terá ficado escalonado. Se alguma equação do sistema ficou da forma  $0 = c$  mas  $c \neq 0$  o sistema original não tem solução. Se isto não ocorreu, eliminando as equações da forma  $0 = 0$ , poderemos encontrar a solução geral do sistema seguindo o procedimento explicado.

Do ponto de vista geométrico, um sistema linear poder ser interpretado da seguinte forma. Considere o sistema

$$\begin{cases} 2x + y = 1 \\ x - y = 2 \end{cases} \quad (1.8)$$

A primeira equação também pode ser escrita como  $y = -2x + 1$ . Sendo assim, poderíamos associar a primeira equação à seguinte função  $y = f(x) = -2x + 1$ . De modo similar, a segunda equação pode ser associada a seguinte função  $y = g(x) = x - 2$ . Essas duas retas podem ser vistas na Figura 1.1.

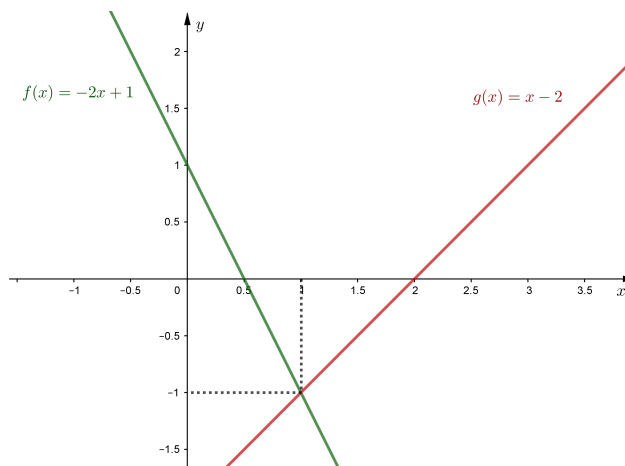


Figura 1.1: Funções  $f$  e  $g$  que representam a primeira e a segunda equação do sistema (1.9), respectivamente. O ponto de intersecção representa a solução do sistema.

O ponto de intersecção entre as duas retas representam a solução do sistema. Perceba que ao escalonar o sistema (1.9), obtemos o sistema equivalente

$$\begin{cases} 2x + y = 1 \\ 3y = -3 \end{cases}, \quad (1.9)$$

encontrando como solução  $x = 1$  e  $y = -1$ , que são exatamente as coordenadas do ponto de intersecção entre as curvas  $f$  e  $g$  ilustradas na Figura 1.1.

**Exercício 1.1.1** Qual seria a interpretação geométrica de um sistema linear de duas equações e duas incógnitas, para o caso em que não existe solução? E para o caso que existem infinitas soluções?

**Exercício 1.1.2** Analise geometricamente todos os casos possíveis de solução para um sistema linear de três equações e três incógnitas. Faça o mesmo para um sistema com duas equações e três incógnitas.

**Exercício 1.1.3** Resolva, usando escalonamento, os sistemas lineares

$$\begin{cases} 3x + 2y - z = 3 \\ 5x - y + 3z = 1 \end{cases}$$

e

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

**Exercício 1.1.4** *Invente sistemas lineares de  $2 \times 2$  com infinitas soluções e sem solução. Verifique sua invenção usando o algoritmo de escalonamento.*

**Exercício 1.1.5** *Faça como no exercício anterior com sistemas de diferentes dimensões.*

**Exercício 1.1.6** *Encontre o conjunto de soluções dos seguintes sistemas lineares escalonados. Escreva a “solução geral” de cada um destes sistemas. Interprete geometricamente. Identifique parâmetros da solução geral e graus de liberdade. Em cada caso onde haja infinitas soluções calcule explicitamente pelo menos duas. Desenhe tudo o que for possível.*

1. 
$$\begin{cases} x_1 + 2x_2 = 4 \\ x_2 = 6 \end{cases}$$

2. 
$$3x_1 - x_2 = 1$$

3. 
$$\begin{cases} x_1 + x_2 = 8 \\ x_1 = 3 \end{cases}$$

4. 
$$x_1 = 5$$

5. 
$$\begin{cases} x + y = 2 \\ y = 9 \end{cases}$$

6. 
$$\begin{cases} x_1 + x_2 - x_3 = 3 \\ x_2 + x_3 = 5 \\ x_3 = 8 \end{cases}$$

7. 
$$\begin{cases} x_1 - x_2 - x_3 = 100 \\ x_2 = 5 \\ x_3 = 6 \end{cases}$$

8. 
$$\begin{cases} x_1 - x_2 - x_3 = 7 \\ x_1 + x_2 = 987 \\ x_1 = 11 \end{cases}$$

9. 
$$\begin{cases} 3x_1 + x_2 - x_3 = 4 \\ 2x_2 + 5x_3 = 10 \end{cases}$$

10.

$$x_1 - x_2 - x_3 = 2$$

11.

$$\begin{cases} x + y + z + w = 4 \\ y + z + w = 3 \\ z + w = 2 \\ w = 1 \end{cases}$$

12.

$$\begin{cases} 3x_1 + x_2 - x_3 - x_4 = 2 \\ x_2 - x_3 + x_4 = 1 \\ x_3 + 5x_4 = 6 \end{cases}$$

13.

$$\begin{cases} 2x - y + z - w = 1 \\ x + y + z = 4 \end{cases}$$

14.

$$4x_1 - x_2 - x_3 - x_4 = 6$$

15.

$$\begin{cases} 2x_1 - x_2 - x_3 - x_4 - x_5 - x_6 = 8 \\ x_3 - 3x_4 + 4x_6 = 2001 \end{cases}$$

**Exercício 1.1.7** Explique a razão pela qual cada um dos sistemas do exercício anterior são escalonados. Explique a razão pela qual todo sistema escalonado é “fácil de resolver”.

**Exercício 1.1.8** Sistematize em forma de algoritmo o procedimento para encontrar a solução geral de um sistema escalonado.

**Exercício 1.1.9** Transforme os sistemas abaixo em sistemas escalonados. Depois, encontre o conjunto de soluções. Escreva a “solução geral” de cada um destes sistemas. Interprete geometricamente. Identifique parâmetros da solução geral e graus de liberdade. Em cada caso onde haja infinitas soluções calcule explicitamente pelo menos duas. Desenhe tudo o que for possível.

(a)

$$3x_1 + 2x_2 = 4$$

$$3x_1 + 3x_2 = 10$$

(b)

$$3x_1 - x_2 = 1$$

(c)

$$x_1 + x_2 = 8$$

$$2x_1 + x_2 = 11$$

(d)

$$x_1 = 5$$

(e)

$$x + y = 2$$

$$x - y = -7$$

(f)

$$x_1 + x_2 - x_3 = 3$$

$$x_1 - 2x_3 = -2$$

$$2x_1 + x_2 - 2x_3 = 9$$

(g)

$$x_1 - x_2 - x_3 = 100$$

$$x_1 - x_3 = 105$$

$$2x_1 - x_2 - 2x_3 = 211$$

(h)

$$x_1 - x_2 - x_3 = 7$$

$$2x_1 - x_3 = 994$$

$$4x_1 - x_2 - 2x_3 = 1112$$

(i)

$$3x_1 + x_2 - x_3 = 4$$

$$3x_1 + 3x_2 + 4x_3 = 14$$

(j)

$$x_1 - x_2 - x_3 = 2$$

(k)

$$x + y + z + w = 4$$

$$x + 2y + 2z + 2w = 7$$

$$2x + 3y + 4z + 4w = 13$$

$$4x + 6y + 7z + 8w = 25$$

(l)

$$3x_1 + x_2 - x_3 - x_4 = 2$$

$$3x_1 + 2x_2 - 2x_3 = 3$$

$$6x_1 + 3x_2 - 2x_3 + 4x_4 = 11$$

(m)

$$2x - y + z - w = 1$$

$$x + y + z + w = 4$$

(n)

$$4x_1 - x_2 - x_3 - x_4 = 6$$

(o)

$$2x_1 - x_2 - x_3 - x_4 - x_5 - x_6 = 8$$

$$x_1 + x_2 + x_3 - 3x_4 + x_5 + 4x_6 = 38$$

(p)

$$2x_1 + x_2 = 3$$

$$x_1 + 2x_2 = 3$$

$$x_1 + x_2 = 4$$

(q)

$$2x_1 + x_2 = 3$$

$$x_1 + 2x_2 = 3$$

$$x_1 + x_2 = 2$$

**Exercício 1.1.10** *Um sistema linear*

$$a_{i,1}x_1 + \dots + a_{i,n}x_n = b_i, \quad i = 1, \dots, n$$

se diz homogêneo se  $b_i = 0$  para todo  $i = 1, \dots, n$ . Que aspecto deve ter a solução geral de um sistema homogêneo?

**Exercício 1.1.11** *O exame de sangue do paciente A estabelece que ele tem 37 mg/dl de HDL ("colesterol bom"), 86.2 mg/dl de LDL ("colesterol ruim"), 144 mg/dl de triglicerídeos e 152 mg/dl de Colesterol Total. Já o exame do paciente B diz que ele tem 53 mg/dl de HDL ("colesterol bom"), 132 mg/dl de LDL ("colesterol ruim"), 82 mg/dl de triglicerídeos e 201 mg/dl de Colesterol Total. Por último, para o paciente C, o relatório do laboratório dá por resultado 50 mg/dl de HDL ("colesterol bom"), 100 mg/dl de LDL ("colesterol ruim"), 150 mg/dl de triglicerídeos e 180 mg/dl de Colesterol Total.*

*O colesterol total, o colesterol bom e os triglicerídeos são efetivamente medidos no laboratório, mas*

*o colesterol ruim é estimado usando uma fórmula (chamada "fórmula de Friedwald"). Qual você acha que é essa fórmula?*

**Exercício 1.1.12** *Neste capítulo começamos mostrando que o problema de fabricação de uma ração animal envolve a resolução de um sistema linear de equações. Entretanto, esse problema prático não acaba aí, pois muitas soluções desse sistema não servem (aquelas que têm componentes negativas) e algumas soluções do sistema linear são melhores que outras porque os ingredientes podem ser adquiridos com um custo menor. Usando estas observações, formule o problema da ração de uma maneira mais realista. Para o caso de dois ingredientes e um nutriente, interprete geometricamente a nova formulação. Faça o mesmo para os casos de um nutriente e três ingredientes e para o caso de dois nutrientes e três ingredientes.*

## Capítulo 2

# Introdução aos Espaços Vetoriais

### 2.1 Representação de vetores

René Descartes (1596-1650) e Pierre de Fermat (1601-1655) identificaram o plano físico com os pares de números reais e o espaço físico tridimensional com as 3-uplas de números reais. Com isso foi possível conectar as propriedades da geometria euclidiana com cálculos puramente algébricos [Martínez \(2004\)](#).

Chamamos de *vetor* uma grandeza que possui três características: direção, sentido e intensidade (tamanho). Por exemplo, o vetor ilustrado na Figura 2.1 representa um segmento orientado e definido por um par ordenado, em que  $A$  é a origem e  $B$  é a extremidade do vetor.

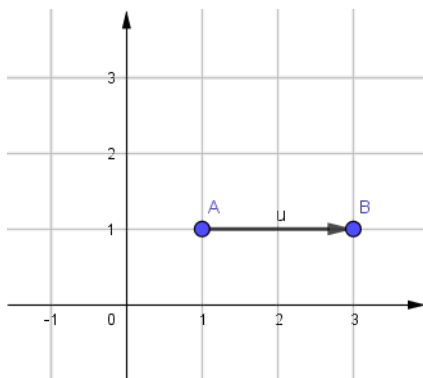


Figura 2.1: Representação de um vetor  $\vec{u}$ , cujo ponto inicial é  $A = (1, 1)$  e o ponto final é  $B = (3, 1)$ . O vetor  $\vec{u}$  é caracterizado pela direção horizontal, sentido para direita e tamanho 2.

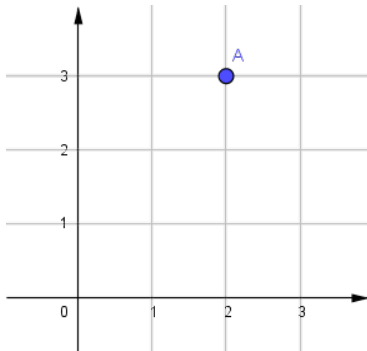
### 2.2 Representação de pontos e vetores

Chamamos  $\mathbb{R}^n$  ao conjunto das  $n$ -uplas de números reais. Por exemplo, o  $\mathbb{R}^2$  é composto por 2-uplas (ou também chamadas de pares) denotadas por  $(x, y)$ , em que cada coordenada  $x$  e  $y$  são números reais. Já o  $\mathbb{R}^3$  é composto por 3-uplas (ou também chamadas de triplas ou ternas) denotadas por  $(x, y, z)$ , em que cada coordenada  $x$ ,  $y$  e  $z$  são números reais. Esses pares ou triplas também podem ser denotadas

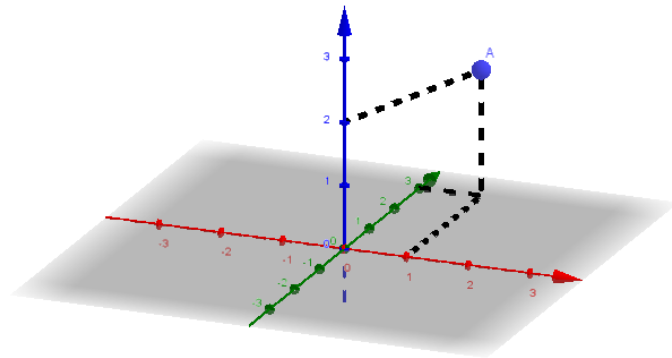
em sua forma matricial, isto é, como vetores coluna:

$$\begin{bmatrix} x \\ y \end{bmatrix} \text{ e } \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Esses pontos podem ser visualizados como na Figura 2.2.



(a) Representação de um ponto com duas coordenadas  $(x, y) = (2, 3)$



(b) Representação de um ponto com três coordenadas  $(x, y, z) = (1, 3, 2)$

Figura 2.2: Representação de pontos com diferentes coordenadas

A 2-upla

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

é chamada de *origem* de  $\mathbb{R}^2$ . De modo similar,  $(0, 0, 0)$  é a origem de  $\mathbb{R}^3$ , e assim por diante. De um modo geral, por simplificação, vamos denotar a origem de  $\mathbb{R}^n$  por 0.

Como é possível perceber na Figura 2.1, um vetor pode ser descrito pelo seu ponto inicial ( $A$ ) e seu ponto final ( $B$ ). Note que os vetores dados na Figura 2.3 são todos equivalentes, pois possuem a mesma direção, sentido e tamanho.

Sendo assim, todo vetor que possui mesma direção, sentido e tamanho, independentemente do ponto inicial e ponto final, são vistos como “iguais”. Esses vetores são chamados de vetores *equipolentes*. Portanto, para não haver distinção entre eles, associamos todos a um único vetor, aquele que possui ponto inicial localizado na origem (veja Figura 2.4).

Assim, quando falarmos “considere o vetor  $(1, 2)$ ” estaremos nos referindo ao vetor cujo ponto final é  $(1, 2)$  e o ponto inicial é o  $(0, 0)$ , isto é, estaremos nos referindo a uma grandeza “vetorial”. Por outro lado, se falarmos “considere o ponto  $(1, 2)$ ” estaremos nos referindo apenas ao ponto, ou seja, estaremos falando sobre uma grandeza “escalar”.



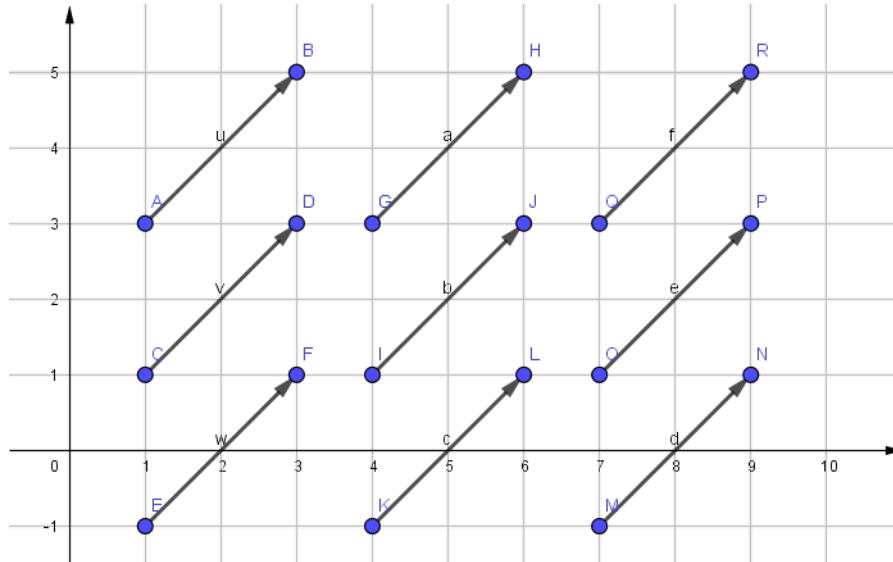


Figura 2.3: Representação de vetores chamados de equipolentes. Todos os vetores possuem mesma direção, sentido e tamanho.

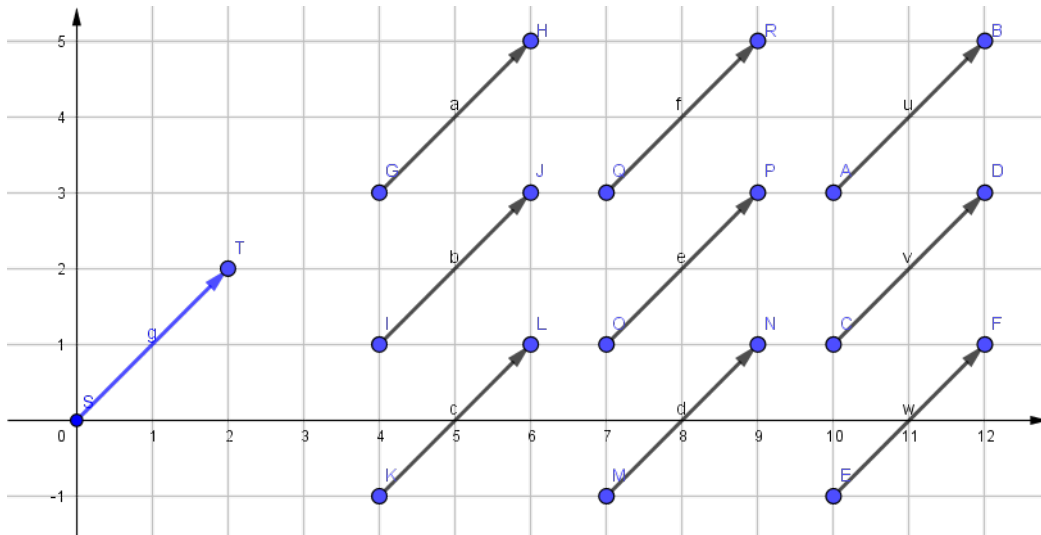


Figura 2.4: Vetores equipolentes associados a um único vetor, cujo ponto inicial é a origem (vetor em azul).

## 2.3 Operações entre vetores

Elementos de  $\mathbb{R}^n$  se somam efetuando a operação  $+$  em cada coordenada, isto é, a soma de  $(1, 2)$  com  $(3, 4)$  é dada por  $(1, 2) + (3, 4) = (1 + 3, 2 + 4) = (4, 6)$ , ou ainda na forma matricial

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 + 3 \\ 2 + 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

Da mesma forma faríamos para vetores com três ou mais coordenadas. Como estamos representando vetores por  $n$ -uplas, a mesma ideia de soma será generalizada para vetores. Por exemplo, para somar os vetores  $\vec{u} = (1, 2)$  e  $\vec{v} = (3, 1)$ , poderíamos somar suas respectivas coordenadas, obtendo  $\vec{w} = (1, 2) + (3, 1) = (1 + 3, 2 + 1) = (4, 3)$  que pode ser visto na Figura 2.5.

Agora, o que esse processo significa geometricamente? Para somar dois vetores, é preciso fixar o

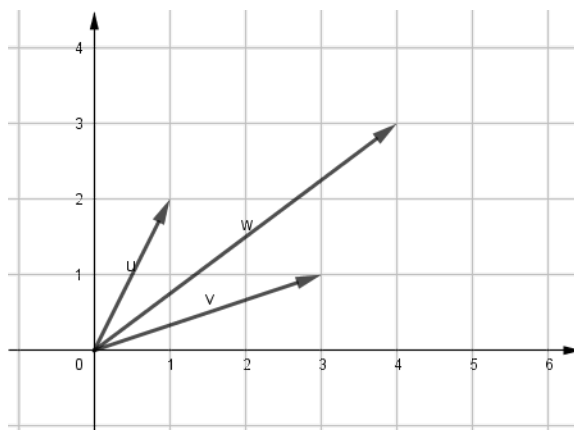


Figura 2.5: Soma entre os vetores  $\vec{u} = (1, 2)$  e  $\vec{v} = (3, 1)$ , resultando no vetor  $\vec{w}$ .

primeiro vetor  $\vec{u}$  e colocar o ponto inicial do vetor  $\vec{v}$  “em cima” do ponto final de  $\vec{u}$ , e assim, o caminho total (do ponto inicial de  $\vec{u}$  até o ponto final de  $\vec{v}$ ) representaria a soma  $\vec{u} + \vec{v}$ . Como estamos adotando o critério de que todos os vetores têm ponto inicial na origem, transladamos  $\vec{v}$  até o ponto final e com isso podemos efetuar a soma (veja a Figura 2.6)

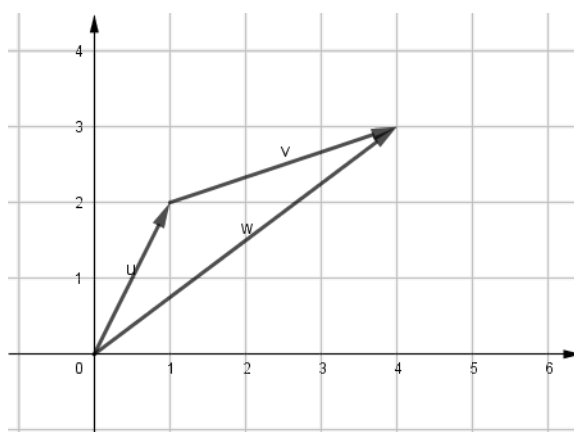


Figura 2.6: Soma entre os vetores  $\vec{u} = (1, 2)$  e  $\vec{v} = (3, 1)$ , resultando no vetor  $\vec{w}$  através da regra do paralelogramo.

Para fazer o contrário, isto é,  $\vec{v} + \vec{u}$  bastaria inverter o processo mencionado acima. Independente da ordem escolhida, teríamos o mesmo resultado como poder ser observado na Figura 2.7. Perceba que os vetores  $\vec{u}$  e  $\vec{v}$  formaram um paralelogramo. A soma entre os vetores, através dessa regra, é chamada de *regra do paralelogramo*, e como podemos observar, ela é equivalente a soma coordenada por coordenada.

Podemos também definir a multiplicação de um número escalar por um vetor, multiplicando esse valor escalar em cada coordenada do vetor. Por exemplo, dado o vetor  $\vec{u} = (2, 3)$ , temos que  $2(1, 3) = (2 \cdot 1, 2 \cdot 3) = (2, 6)$ . Por outro lado,  $\frac{1}{2}(1, 3) = (\frac{1}{2} \cdot 1, \frac{1}{2} \cdot 3) = (\frac{1}{2}, \frac{3}{2})$ . Através da Figura 2.8 percebemos que a multiplicação por 2 “esticou” o vetor  $\vec{u}$ , enquanto que a multiplicação por  $\frac{1}{2}$  “encolheu” o vetor  $\vec{u}$ .

Isso ocorre em geral. A multiplicação de um vetor por um escalar maior que 1 esticará o vetor, enquanto que a multiplicação por um escalar entre 0 e 1 encolherá esse vetor. Observe que a multiplicação pelo valor 1 preservará o tamanho do vetor e a multiplicação por 0 resultará na obtenção da origem.

**Exercício 2.3.1** Considere o vetor  $\vec{u} = (1, 3)$ . O que aconteceria com este vetor caso multiplicássemos pelo número  $-2$ ? E se multiplicássemos pelo valor  $-\frac{1}{2}$ ? O que você concluiria a partir de suas observações?

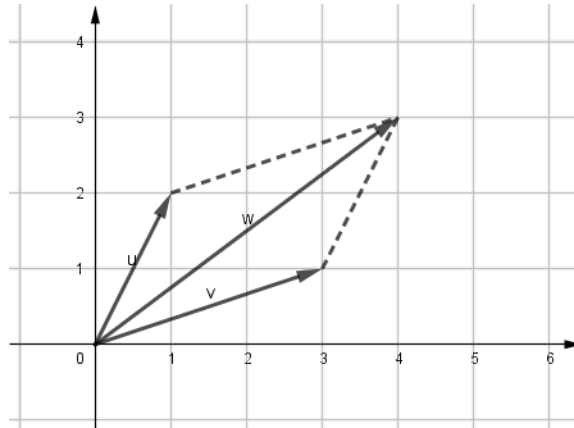


Figura 2.7: Soma entre os vetores  $\vec{u} = (1, 2)$  e  $\vec{v} = (3, 1)$ , resultando no vetor  $\vec{w}$  através da regra do paralelogramo. As linhas pontilhadas representam as translações dos vetores  $\vec{u}$  e  $\vec{v}$  para aplicar a regra do paralelogramo

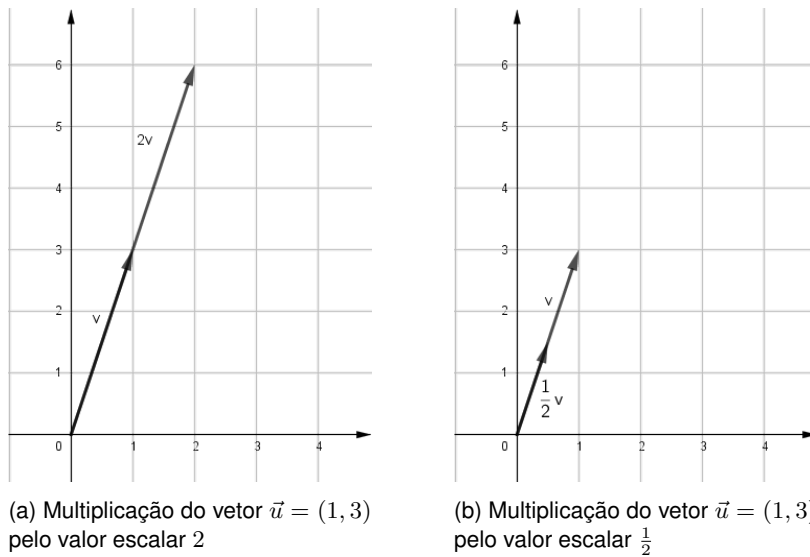


Figura 2.8: Representação de pontos com diferentes coordenadas

As operações definidas aqui satisfazem as seguinte propriedades:

1. **Associatividade:**  $(u + v) + w = u + (v + w)$ , para todo  $u, v, w \in \mathbb{R}^n$ ;
2. **Comutatividade:**  $u + v = v + u$ , para todo  $u, v \in \mathbb{R}^n$ ;
3. **Elemento neutro:** Existe  $0 = (0, \dots, 0) \in \mathbb{R}^n$ , chamado de vetor nulo, tal que  $0 + u = u$ , para todo  $u \in \mathbb{R}^n$ ;
4. **Elemento oposto:** Existe  $-u \in \mathbb{R}^n$ , chamado de vetor oposto, tal que  $(-u) + u = 0$ , para todo  $u \in \mathbb{R}^n$ ;
5. **Distributividade escalar:**  $\alpha(u + v) = \alpha v + \alpha u$ , para todo  $u, v \in \mathbb{R}^n$  e  $\alpha \in \mathbb{R}$ ;
6. **Distributividade vetor:**  $(\alpha + \beta)v = \alpha v + \beta v$ , para todo  $v \in \mathbb{R}^n$  e  $\alpha, \beta \in \mathbb{R}$ ;
7. **Associatividade escalar:**  $(\alpha\beta)v = \alpha(\beta v)$ , para todo  $v \in \mathbb{R}^n$  e  $\alpha, \beta \in \mathbb{R}$ ;
8. **Elemento identidade:**  $1.v = v$ , para todo  $v \in \mathbb{R}^n$ ;

Quando um espaço  $U$  satisfizer as 8 condições mencionadas acima, diremos que  $U$  é um espaço vetorial. Por exemplo,  $\mathbb{R}$ ,  $\mathbb{R}^2$ ,  $\mathbb{R}^3 \dots \mathbb{R}^n$  são espaços vetoriais. Além destes, temos que os números complexos ( $\mathbb{C}$ ), polinômios de ordem  $n$  ( $\mathbb{P}_n(\mathbb{R})$ ) e matrizes ( $\mathbb{M}_{m \times n}(\mathbb{R})$ ) são outros exemplos de espaços vetoriais.

Às vezes não precisamos estudar o comportamento de um espaço  $U$  como um todo, mas sim uma parte (subconjunto) dele que possuem propriedades tão boas quanto as de espaços vetoriais. Esses subconjuntos  $W \subseteq U$  são chamados de *subespaços vetoriais* e requerem as seguintes propriedades:

1. O vetor nulo do espaço  $U$  também pertence ao subconjunto  $W$ ;
2. Quaisquer dois elementos  $u, v \in W$ , temos que  $u + v \in W$ ;
3. Para qualquer  $u \in W$  e  $\lambda \in \mathbb{R}$ , temos que  $\lambda u \in W$ .

**Exemplo 2.3.1** Considere  $U = \mathbb{R}^2$  e  $W \subseteq U$  dada por  $W = [(1, 0)]$ . Isto é,  $W$  é a reta da abscissa. Veja que  $(0, 0)$  pertence a  $W$ , pois  $(0, 0) = 0(1, 0)$ . Tomando dois elementos em  $W$ , isto é,  $u = (a, 0)$  e  $v = (b, 0)$  temos que  $u + v = (a, 0) + (b, 0) = (a + b, 0) \in W$ . Por fim, tomando qualquer  $(a, 0) \in W$  e  $\lambda \in \mathbb{R}$ , temos que  $\lambda(a, 0) = (\lambda a, 0) \in W$ . Portanto,  $W$  é um subespaço vetorial de  $\mathbb{R}^2$ .

**Exercício 2.3.2** Como devem ser geometricamente os subespaços de  $\mathbb{R}^2$ ?

**Exercício 2.3.3** Determine (geometricamente) todos os subespaços vetoriais de  $\mathbb{R}^3$ .

## 2.4 Distâncias e ângulos

A *norma* (ou comprimento) de um vetor  $x \in \mathbb{R}^n$  se define da seguinte forma:

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Por exemplo, o tamanho (norma) do vetor  $(1, 2)$  é dado por  $\|(1, 2)\| = \sqrt{1^2 + 2^2} = \sqrt{5}$ . Para três dimensões teríamos que, por exemplo,  $\|(1, -1, 0)\| = \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2}$  é o tamanho do vetor  $(1, -1, 0)$ . Pelo Teorema de Pitágoras, esta medida é a distância euclidiana entre o ponto  $x$  e a origem. Veja Figura 2.9.

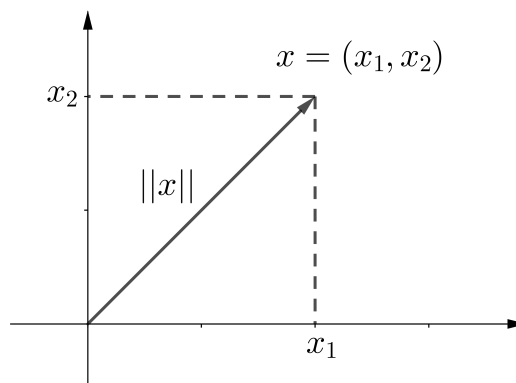


Figura 2.9: A norma de um vetor  $x = (x_1, x_2)$

**Observação 2.4.1** Observe que a norma é uma generalização do módulo, isto é, quando estamos trabalhando com apenas uma variável, temos que  $\|x\| = \sqrt{x^2} = |x|$ .

Uma propriedade interessante de norma é que, para qualquer  $x \in \mathbb{R}^n$  e  $\lambda \in \mathbb{R}$ , temos

$$\|\lambda x\| = |\lambda| \|x\|.$$

Através da norma é possível calcular a *distância* entre dois vetores, que é definida da seguinte forma:  $dist(x, y) = \|x - y\|$  (veja a Figura 2.10). Por exemplo, a distância entre os dois vetores  $x = (1, 2)$  e  $y = (2, -3)$  é dada por

$$\|x - y\| = \|(1, 2) - (2, -3)\| = \|(-1, 5)\| = \sqrt{(-1)^2 + 5^2} = \sqrt{26}$$

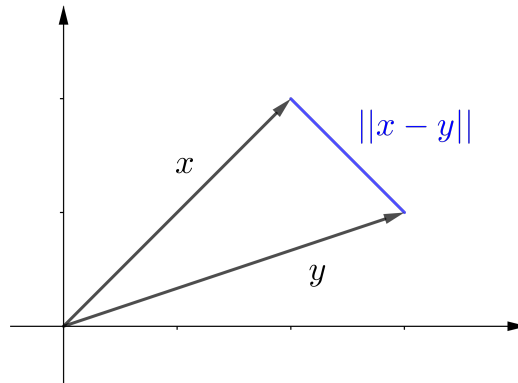


Figura 2.10: A distância entre vetores  $x$  e  $y$

Chamamos de *produto escalar* (ou de modo geral, *produto interno*) entre os vetores  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$  da seguinte forma:

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n.$$

Essa operação possui as seguintes propriedades [Boldrini \(1984\)](#)

**Proposição 2.4.1** Para quaisquer  $x, y, z \in \mathbb{R}^n$ ,  $\alpha, \beta \in \mathbb{R}$ , tem-se:

$$\langle x, y \rangle = \langle y, x \rangle,$$

$$\langle \alpha x, y \rangle = \langle x, \alpha y \rangle = \alpha \langle x, y \rangle,$$

$$\langle x + y, z \rangle = \langle z, x + y \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

A *Desigualdade Cauchy-Schwarz* é uma propriedade importante que relaciona produto escalar e normas. Para quem tiver curiosidade sobre como é obtida tal relação, sugerimos a seguinte referência [Martínez \(2004\)](#).

**Teorema 2.4.1** Para quaisquer  $x, y \in \mathbb{R}^n$ , tem-se que

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \tag{2.1}$$

Da relação (2.1) podemos concluir que

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \|y\|} \leq 1.$$

e assim, a quantidade  $\frac{\langle x, y \rangle}{\|x\| \|y\|}$  mede alguma propriedade significativa em relação à posição relativa dos vetores  $x$  e  $y$ . De fato, veremos agora que essa propriedade é o *cosse*no do ângulo entre os vetores não-nulos  $x$  e  $y$ .

Para ver isto, desenhe os vetores não-nulos  $x$  e  $y$  de maneira que formem um ângulo agudo. Agora, considere todos os pontos da forma  $tx$  com  $t \geq 0$ . Esses pontos formam a semirreta com origem em  $O$  e que contém o vetor  $x$ . Desenhe, nesta semirreta, o ponto que fica mais perto de  $y$ . Chamemos  $P$  a esse ponto. Claramente, o segmento que une  $y$  com  $P$  deve ser perpendicular à semirreta, portanto, os pontos  $y$ ,  $P$  e a origem  $O$  formam um triângulo retângulo. A hipotenusa mede  $\|y\|$  e o “cateto adjacente” mede  $\|P\|$ . Logo, o que entendemos por *cosse*no do ângulo entre  $x$  e  $y$  (da trigonometria plana) é o quociente  $\|P\|/\|y\|$ . (Como o ângulo é agudo, o sinal é positivo.)

Agora vamos procurar uma fórmula analítica para  $P$ . Este ponto deve ser aquele elemento da forma  $tx$ , com  $t \geq 0$ , cuja distância a  $y$  é a menor possível. Portanto, se trata de encontrar  $t \geq 0$  tal que a expressão

$$\|y - tx\|$$

seja mínima. Desenvolvendo, temos:

$$\|y - tx\|^2 = \|y\|^2 + t^2\|x\|^2 - 2t\langle x, y \rangle.$$

Logo, encontrar  $t$  tal que minimize  $\|y - tx\|$  equivale a encontrar  $t$  minimizando a quantidade

$$t^2\|x\|^2 - 2t\langle x, y \rangle.$$

Agora,

$$\begin{aligned} t^2\|x\|^2 - 2t\langle x, y \rangle &= t^2\|x\|^2 - 2t\langle x, y \rangle + \langle x, y \rangle^2 / \|x\|^2 - \langle x, y \rangle^2 / \|x\|^2 \\ &= (t\|x\| - \langle x, y \rangle / \|x\|)^2 - \langle x, y \rangle^2 / \|x\|^2. \end{aligned}$$

O valor de  $t$  que minimiza esta expressão é aquele que faz com que o termo elevado ao quadrado se anule, ou seja:

$$t = \frac{\langle x, y \rangle}{\|x\|^2}.$$

Como nosso ângulo é agudo, este  $t$  é maior que zero, portanto  $\langle x, y \rangle > 0$ . Assim, temos:

$$P = \frac{\langle x, y \rangle}{\|x\|^2} x,$$

portanto

$$\|P\| = \frac{\langle x, y \rangle}{\|x\|}.$$

Mas o *cosse*no trigonométrico do ângulo é  $\|P\|/\|y\|$ , portanto esse *cosse*no é igual a

$$\frac{\langle x, y \rangle}{\|x\| \|y\|}$$

como suspeitávamos.

Fica a cargo da leitora mostrar que o mesmo raciocínio geométrico, com pequenas variações, pode ser reproduzido se o ângulo entre  $x$  e  $y$  é obtuso e também se o ângulo é reto. Portanto, de agora em diante, a quantidade

$$\frac{\langle x, y \rangle}{\|x\| \|y\|}$$

será adotada como definição de cosseno do ângulo entre os vetores não-nulos  $x$  e  $y$ , isto é,

$$\frac{\langle x, y \rangle}{\|x\| \|y\|} = \cos(\theta) \quad \text{ou equivalentemente} \quad \langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos(\theta)$$

Assim, os vetores serão *ortogonais* se seu produto escalar é nulo.

## Capítulo 3

# Combinação Linear

Dizemos que o vetor  $v \in \mathbb{R}^n$  é *combinação linear* de  $v_1, \dots, v_m \in \mathbb{R}^n$  se existem escalares  $t_1, \dots, t_m$  tais que

$$v = \lambda_1 v_1 + \dots + \lambda_m v_m.$$

Por exemplo, o vetor  $(1, 2)$  é combinação linear dos vetores  $(1, 0)$  e  $(0, 1)$  uma vez que

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

ou ainda, o mesmo vetor  $(1, 2)$  pode ser escrito como combinação linear dos vetores  $(1, 1)$  e  $(1, 0)$  já que

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

Dado um conjunto  $\{v_1, \dots, v_m\} \subset \mathbb{R}^n$ , se algum deles é combinação linear dos outros ou se o único elemento do conjunto é o vetor nulo, dizemos que eles são *linearmente dependentes*. Em caso contrário, dizemos que  $v_1, \dots, v_m$  são *linearmente independentes*.

**Exemplo 3.0.1** Considere o conjunto  $U = \{(1, 2), (1, 1), (1, 0)\}$ . Como visto anteriormente, temos que  $(1, 2)$  é escrito como combinação linear de  $(1, 1)$  e  $(1, 0)$ . Portanto, o conjunto  $U$  é linearmente dependente.

É comum se referir aos conjuntos linearmente dependentes por conjuntos L.D. e os conjuntos linearmente independentes por conjuntos L.I.

**Exemplo 3.0.2** Considere o conjunto  $U = \{(1, 1), (1, 0)\}$ . Perceba que não existe um escalar  $\lambda$  de tal forma que  $(1, 1) = \lambda(1, 0)$ , pois caso contrário teríamos

$$(1, 1) = \lambda(1, 0) \tag{3.1}$$

$$= (\lambda, 0) \tag{3.2}$$

$$\Rightarrow \begin{cases} 1 = \lambda \\ 1 = 0 \end{cases} \tag{3.3}$$

$$\tag{3.4}$$

que é um sistema impossível. Portanto, o conjunto  $U$  é L.I.

Uma outra forma de analisar se um conjunto de vetores  $v_1, \dots, v_m$  é L.D. ou L.I. se dá a partir da



seguinte combinação linear:

$$\lambda_1 v_1 + \dots + \lambda_m v_m = 0$$

Se a combinação linear dada acima só for verdade se **todos** os escalares forem iguais a 0, então o conjunto de vetores é L.I., caso contrário, isto é, se nem todos os escalares forem iguais a 0, então o conjunto de vetores são L.D. Perceba que esse resultado é equivalente a definição de combinação linear. Por exemplo, se  $v_1$  e  $v_2$  são L.D. então um é combinação linear do outro, ou seja,  $v_1 = \lambda_2 v_2$  implicando que  $v_1 - \lambda_2 v_2 = 0$ , e portanto segue que existem escalares não nulos tal que a combinação linear entre eles é igual a 0. Por outro lado, se existem escalares não nulos tais que  $\lambda_1 v_1 + \lambda_2 v_2 = 0$ , então isolando  $v_1$  temos que  $v_1 = -\frac{\lambda_2}{\lambda_1} v_2$ , ou seja,  $v_1$  é múltiplo de  $v_2$  e portanto são L.D.

**Exemplo 3.0.3** Considere os conjuntos  $U = \{(4, 1), (2, 3), (6, 4)\}$ ,  $V = \{(4, 1), (2, 3), (10, 5)\}$  e  $W = \{(4, 1), (2, 3)\}$ .

Veja que  $U$  é um conjunto L.D. já que

$$\lambda_1(4, 1) + \lambda_2(2, 3) + \lambda_3(6, 4) = (0, 0)$$

implica no seguinte sistema linear

$$\begin{cases} \lambda_1 4 + \lambda_2 2 + \lambda_3 6 = 0 \\ \lambda_1 1 + \lambda_2 3 + \lambda_3 4 = 0 \end{cases}$$

que é um sistema possível e indeterminado. Portanto,  $U$  é um conjunto L.D.

De modo similar, constatamos que o conjunto  $V$  é L.D. porque fazendo a combinação linear

$$\lambda_1(4, 1) + \lambda_2(2, 3) + \lambda_3(10, 5) = (0, 0)$$

obtemos o seguinte sistema linear

$$\begin{cases} \lambda_1 4 + \lambda_2 2 + \lambda_3 10 = 0 \\ \lambda_1 1 + \lambda_2 3 + \lambda_3 5 = 0 \end{cases}$$

que é um sistema possível e indeterminado.

Por fim, vamos analisar o conjunto  $W$ . A combinação linear

$$\lambda_1(4, 1) + \lambda_2(2, 3) = (0, 0)$$

equivale ao seguinte sistema linear

$$\begin{cases} \lambda_1 4 + \lambda_2 2 = 0 \\ \lambda_1 1 + \lambda_2 3 = 0 \end{cases}$$

que é um sistema possível e determinado, cuja única solução é dada por  $\lambda_1 = \lambda_2 = 0$ . Portanto, o conjunto  $W$  é L.I.

A representação geométrica desses vetores podem ser vistas na Figura 3.1. Na figura temos que  $u = (2, 3)$ ,  $v = (4, 1)$ ,  $b = (6, 4)$  e  $e = (10, 5)$ . É possível perceber que o vetor  $b$  pode ser obtido através de uma combinação linear dos vetores  $u$  e  $v$  fazendo  $b = u + v$  (a partir da regra do paralelogramo). De modo similar, o vetor  $e$  pode ser obtido através de uma combinação linear dos vetores  $u$  e  $v$ , "esticando" primeiro o vetor  $v$  para depois somá-lo ao vetor  $u$ , isto é,  $e = u + 2v$ .

**Deixaremos como exercício resolver os sistemas lineares obtidos neste exemplo e chegar nas conclusões mencionadas.**

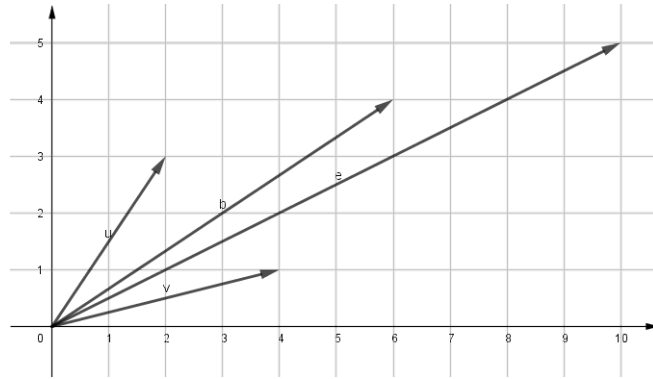


Figura 3.1: Conjunto de vetores  $U$ ,  $V$  e  $W$ , em que o único conjunto L.I. é o conjunto  $W$ .

**Exercício 3.0.1** Verifique quais dos vetores abaixo são L.D ou L.I:

1.  $(1, 1)$
2.  $(0, 0)$
3.  $(1, 2)$  e  $(-2, 3)$
4.  $(1, 1)$  e  $(-1, 1)$
5.  $(0, 0)$  e  $(1, 1)$
6.  $(1, 0, 1)$  e  $(1, 1, 1)$
7.  $(1, 0, 1)$  e  $(1, 0, 2)$
8.  $(1, 0, 1)$ ,  $(1, 1, 1)$  e  $(-1, 0, 2)$
9.  $(1, 0, 1)$ ,  $(1, 1, 1)$  e  $(2, -1, 3)$
10.  $(1, 0, 0)$ ,  $(1, 1, 1)$  e  $(2, 3, 0)$

### 3.1 Conjunto gerador

Chamamos de *conjunto gerador* de  $u_1, \dots, u_m$  (denotado por  $[u_1, \dots, u_m]$ ) o conjunto formado por todas as combinações lineares entre esses vetores, isto é,

$$[u_1, \dots, u_m] = \{u \in U : u = \lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_m u_m, \quad \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}\}.$$

**Exemplo 3.1.1** Considere  $u = (1, 0)$ . Então o conjunto gerador de  $u$  é dado por

$$[u] = \{(x, y) \in U : (x, y) = \lambda(1, 0)\}$$

isto é, o gerador de  $u$  produz elementos da forma  $(\lambda, 0)$  (veja a reta azul na Figura 3.2).

Considere agora  $v = (0, 1)$ . Então o conjunto gerador de  $v$  é dado por

$$[v] = \{(x, y) \in U : (x, y) = \lambda(0, 1)\}$$

isto é, o gerador de  $v$  produz elementos da forma  $(0, \lambda)$  (veja a reta vermelha na Figura 3.2).

Já o conjunto gerador de  $u$  e  $v$  é dado por

$$[u, v] = \{(x, y) \in U : (x, y) = \lambda_1(1, 0) + \lambda_2(0, 1)\}$$

isto é, o gerador de  $u$  e  $v$  produz elementos da forma  $(\lambda_1, \lambda_2)$ . Ou seja, o gerador  $[u, v]$  produz todo o espaço  $\mathbb{R}^2$  (veja o plano roxo na Figura 3.2).

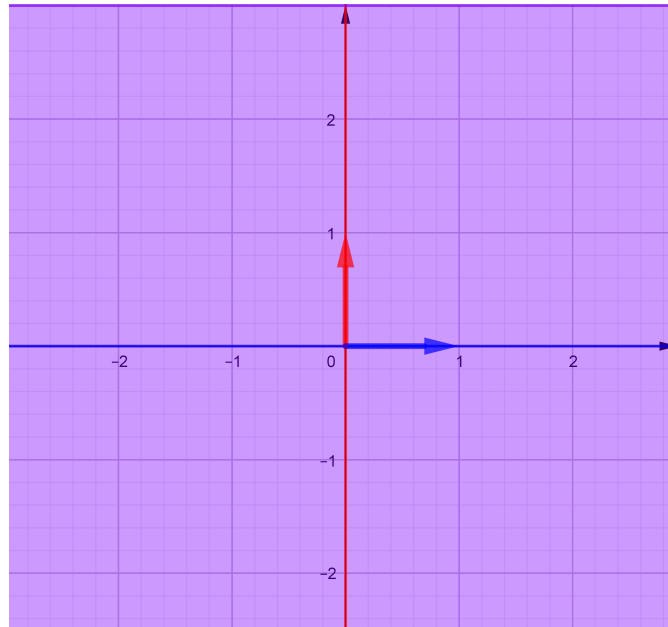


Figura 3.2: Representação dos conjuntos geradores  $[(1, 0)]$  (representado pela reta azul),  $[(0, 1)]$  (representado pela reta vermelha) e  $[(1, 0), (0, 1)]$  (representado pelo plano em roxo), em que o vetor  $(1, 0)$  está representado pela “seta” azul e  $(0, 1)$  pela seta vermelha.

**Observação 3.1.1** Perceba que o conjunto gerador abaixo

$$[(1, 1), (1, 0)] = \{(x, y) \in U : (x, y) = \lambda_1(1, 1) + \lambda_2(1, 0)\}$$

também gera todo o conjunto  $\mathbb{R}^2$ , uma vez que todo elemento  $(x, y)$  de  $\mathbb{R}^2$  pode ser escrito como combinação linear de  $(1, 1)$  e  $(1, 0)$ . Isso significa que um mesmo espaço pode ser gerado por diferentes conjuntos geradores. Para visualizar que de fato um elemento  $(x, y)$  qualquer pode ser escrito como combinação linear de  $(1, 1)$  e  $(1, 0)$ , considere o seguinte raciocínio

$$\begin{aligned} (x, y) &= \lambda_1(1, 1) + \lambda_2(1, 0) \Leftrightarrow \\ (x, y) &= (\lambda_1, \lambda_1) + (\lambda_2, 0) \Leftrightarrow \\ (x, y) &= (\lambda_1 + \lambda_2, \lambda_1) \Leftrightarrow \\ &\Leftrightarrow \begin{cases} x = \lambda_1 + \lambda_2 \\ y = \lambda_1 \end{cases} \end{aligned}$$

concluindo que  $\lambda_2 = y$  e  $\lambda_1 = x - y$ .

No **Exemplo 3.1.1** vimos que o vetor  $(1, 0)$  gera apenas uma reta, enquanto que  $(0, 1)$  gera uma outra reta que só possui o  $(0, 0)$  como elemento comum. Ao combinar os dois vetores  $(1, 0)$  e  $(0, 1)$  conseguimos gerar todo o espaço  $\mathbb{R}^2$ . Isso significa que qualquer elemento que tomarmos de  $\mathbb{R}^2$ , por exemplo

$(a, b) \in \mathbb{R}^2$ , e considerarmos o conjunto  $U = \{(1, 0), (0, 1), (a, b)\}$ , então teremos que  $U$  é L.D. Isso quer dizer que estamos considerando no conjunto  $U$  elementos que não nos trarão nenhuma novidade, com respeito a  $(1, 0)$  e  $(0, 1)$ . Com esse comentário, vamos introduzir o conceito de *base*.

## 3.2 Bases e coordenadas

A partir da observação feita anteriormente, podemos concluir que em  $\mathbb{R}^2$  não pode haver um conjunto linearmente independente com mais de 2 vetores. De modo similar, concluiríamos que em  $\mathbb{R}^3$  não poderia haver um conjunto L.I. com mais de três vetores. Generalizando, temos que, em  $\mathbb{R}^n$ , não pode haver um conjunto linearmente independente com mais de  $n$  elementos.

Ainda mais, percebemos que para gerar o espaço  $\mathbb{R}^2$  precisamos de 2 vetores. De modo similar, concluiríamos que para gerar  $\mathbb{R}^3$  precisaríamos de 3 vetores, e assim por diante.

Dizemos que  $B$  é uma *base* para um espaço  $V$  se o conjunto  $U$  for L.I. e gerar todos os elementos de  $U$ , isto é,  $[B] = U$ . Em  $\mathbb{R}^n$  a definição de base é equivalente a ter  $n$  vetores L.I. (pois são necessário  $n$  vetores para gerar o espaço).

**Exemplo 3.2.1** O conjunto  $B = \{(1, 0), (0, 1)\}$  é uma base para  $\mathbb{R}^2$ , uma vez que é L.I. e gera  $\mathbb{R}^2$ . A base  $B$  é chamada de base canônica. Note que  $C = \{(1, 1), (1, 0)\}$  também é uma base para  $\mathbb{R}^2$ . Com isso podemos concluir que bases de espaços vetoriais não são únicas.

Um outro resultado interessante que temos é que se um conjunto  $B = \{u_1, \dots, u_n\}$  é L.I., então todo vetor  $u \in \mathbb{R}^n$  pode ser escrito de maneira única como combinação linear de  $u_1, \dots, u_n$ , isto é,  $u$  se escreve como  $t_1 u_1 + \dots + t_n u_n$  de maneira única. Esses coeficientes  $t_1, \dots, t_n$  se chamam *coordenadas* de  $u$  na base  $B = \{u_1, \dots, u_n\}$ .

**Exercício 3.2.1** Verifique quais dos conjuntos abaixo formam uma base para  $\mathbb{R}^2$ :

1.  $(1, 2)$  e  $(-2, 3)$
2.  $(1, 1)$  e  $(-1, 1)$
3.  $(0, 0)$  e  $(1, 1)$
4.  $(1, 1)$  e  $(0, -1)$
5.  $(1, 2)$  e  $(3, -2)$
6.  $(1, 2)$  e  $(3, 6)$
7.  $(1, 2)$  e  $(-2, -1)$
8.  $(1, 2)$  e  $(\frac{1}{2}, 1)$

**Exercício 3.2.2** Determine as coordenadas do vetor  $(1, 3)$  com respeito a cada base encontrada no exercício anterior.

**Exercício 3.2.3** Verifique quais dos conjuntos abaixo formam uma base para  $\mathbb{R}^3$ :

1.  $(1, 0, 1)$  e  $(1, 1, 1)$

2.  $(1, 0, 1)$  e  $(1, 0, 2)$

3.  $(1, 0, 1)$ ,  $(1, 1, 1)$  e  $(-1, 0, 2)$

4.  $(1, 0, 1)$ ,  $(1, 1, 1)$  e  $(2, -1, 3)$

5.  $(1, 0, 0)$ ,  $(1, 1, 1)$  e  $(2, 3, 0)$

6.  $(1, 1, 0)$ ,  $(0, -1, 1)$  e  $(2, 1, 1)$

**Exercício 3.2.4** *Determine as coordenadas do vetor  $(1, -1, 2)$  com respeito a cada base encontrada no exercício anterior.*

## Capítulo 4

# Transformações Lineares

### 4.1 Transformação Linear

*Funções lineares* são as associações mais simples que podemos pensar. Por exemplo, imagine que de um quilograma de soja são extraídos 0,2 litros de óleo. Desse modo, de uma produção de  $x$  kg de soja, seriam extraídos  $0,2x$  litros de óleo, e com isso, poderíamos escrever a função  $T(x) = 0,2x$ , em que  $T$  é a quantidade em litros de óleo de soja e  $x$  é a quantidade de soja em kg.

Agora, vejamos algumas características interessantes desse simples exemplo. Primeiro note que para calcular uma produção de  $(x + y)$  kg de soja, podemos tanto multiplicar  $(x + y)$  pelo coeficiente de produção 0,2, como calcular as produções separadamente e depois somá-las, isto é,  $T(x+y) = 0,2(x+y) = 0,2x + 0,2y = T(x) + T(y)$ . Em segundo lugar, veja que se a quantidade de soja for multiplicada por um fator  $\lambda$  (representando um escalar), a produção de óleo também será afetada por esse fator, isto é,  $T(\lambda x) = 0,2(\lambda x) = \lambda 0,2x = \lambda T(x)$ .

Essas duas propriedades caracterizam o que chamamos de *Transformações Lineares*. Isto é, uma transformação linear  $T : U \rightarrow V$  é uma função que leva elementos de  $U$  em elementos de  $V$ , que satisfaz as duas propriedades abaixo:

1.  $T(u + v) = T(u) + T(v)$ ,  $\forall u, v \in U$ ;
2.  $T(\lambda u) = \lambda T(u)$ ,  $\forall u \in U$  e  $\forall \lambda \in \mathbb{R}$ .

**Exemplo 4.1.1** *Boldrini (1984)* A quantidade em litros de óleo extraída por quilograma de cereal, segundo um determinado processo, pode ser descrita pela tabela

|          | Soja | Milho | Algodão | Amendoim |
|----------|------|-------|---------|----------|
| Óleo (l) | 0,2  | 0,06  | 0,13    | 0,32     |

A quantidade total de óleo produzido por  $x$  kg de soja,  $y$  kg de milho,  $z$  kg de algodão e  $w$  kg de amendoim é dada por  $T(x, y, z, w) = 0,2x + 0,06y + 0,13z + 0,32w$ , que cumpre as propriedades da soma e multiplicação:

1.

$$\begin{aligned} T((x_1, y_1, z_1, w_1) + (x_2, y_2, z_2, w_2)) &= T((x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)) \\ &= 0, 2(x_1 + x_2) + 0, 06(y_1 + y_2) + 0, 13(z_1 + z_2) + 0, 32(w_1 + w_2) \\ &= 0, 2x_1 + 0, 06y_1 + 0, 13z_1 + 0, 32w_1 \\ &+ 0, 2x_2 + 0, 06y_2 + 0, 13z_2 + 0, 32w_2 \\ &= T(x_1, y_1, z_1, w_1) + T(x_2, y_2, z_2, w_2) \end{aligned}$$

2.

$$\begin{aligned} T(\lambda(x_1, y_1, z_1, w_1)) &= T((\lambda x_1, \lambda y_1, \lambda z_1, \lambda w_1)) \\ &= 0, 2(\lambda x_1) + 0, 06(\lambda y_1) + 0, 13(\lambda z_1) + 0, 32(\lambda w_1) \\ &= \lambda(0, 2x_1 + 0, 06y_1 + 0, 13z_1 + 0, 32w_1) \\ &= \lambda T(x_1, y_1, z_1, w_1) \end{aligned}$$

Portanto,  $T$  é uma transformação linear de  $\mathbb{R}^4$  em  $\mathbb{R}$ .

É também comum representar tais aplicações nas suas formas “matriciais”, isto é, a aplicação do exemplo anterior pode ser denotada por:

$$T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0, 2x + 0, 06y + 0, 13z + 0, 32w,$$

e portanto, as propriedades verificadas acima ficam:

$$T \left( \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{pmatrix} \right) = T \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} + T \begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{pmatrix}$$

e

$$T \begin{pmatrix} \lambda x_1 \\ \lambda y_1 \\ \lambda z_1 \\ \lambda w_1 \end{pmatrix} = \lambda T \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix}$$

**Exemplo 4.1.2** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = 2x$ . Como  $T(x + y) = 2(x + y) = 2x + 2y = T(x) + T(y)$ , e também,  $T(\lambda x) = 2(\lambda x) = \lambda 2x = \lambda T(x)$ , então segue que  $T$  é uma transformação linear de  $\mathbb{R}$  em  $\mathbb{R}$ .

O exemplo acima exemplifica que aplicações que vão de  $\mathbb{R}$  em  $\mathbb{R}$ , descritas por  $T(x) = \alpha x$  são sempre transformações lineares. No entanto, aplicações do tipo “afim” não cumprem essa propriedade. Vejamos:

**Exemplo 4.1.3** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = 2x + 3$ . Por um lado, temos que  $T(x + y) = 2(x + y) + 3$ . Por outro lado, temos que  $T(x) + T(y) = 2x + 3 + 2y + 3 = 2x + 2y + 6$ , que

claramente são diferentes. Dessa forma, não temos que a propriedade da soma é cumprida, e assim,  $T$  não é uma transformação linear.

Além de funções afins, temos que funções polinomiais (diferentes das lineares) não tem a característica de uma transformação linear. Vejamos:

**Exemplo 4.1.4** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = x^2$ . Por um lado, temos que  $T(x + y) = (x + y)^2 = x^2 + 2xy + y^2$ . Por outro lado, temos que  $T(x) + T(y) = x^2 + y^2$ . Dessa forma, a aplicação  $T$  não é uma transformação linear.

Uma pergunta que surge é a seguinte: Como podemos identificar possíveis aplicações que não são transformações lineares, sem tem que de fato verificar as propriedades? Uma possível forma de abordar essa questão é que se  $T$  é uma transformação linear, então devemos ter que  $T(0) = 0$ . Isso ocorre porque  $T(0) = T(0 + 0) = T(0) + T(0)$ , e assim,  $T(0) = T(0) + T(0)$ , implicando que  $T(0) = 0$ .

Perceba que a aplicação dada no **Exemplo 4.1.8** não satisfaz  $T(0) = 0$ , de fato,  $T(0) = 2 \cdot 0 + 3 = 3 \neq 0$ . Isso implica que  $T$  não é uma transformação linear. Porém, veja que no **Exemplo 4.1.9** a aplicação  $T(x) = x^2$  satisfaz a condição  $T(0) = 0$ , no entanto, vimos que não é uma transformação linear. Portanto, o único resultado que temos é que se uma aplicação não satisfaz a condição  $T(0) = 0$ , então  $T$  não é uma transformação linear. Agora, caso  $T$  satisfaça a condição  $T(0) = 0$ , então não necessariamente podemos concluir que  $T$  é linear, faltando analisar se de fato  $T$  cumpre as propriedades da soma e multiplicação por escalar.

Veja que também podemos ter aplicações que levamos elementos de um espaço  $U$  em outro espaço  $V$  distintos.

**Exemplo 4.1.5** Considere a aplicação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  dada por  $T(x, y) = (2x, 0, x + y)$ . Assim,

$$\begin{aligned} T((x_1, y_1) + (x_2, y_2)) &= T(x_1 + x_2, y_1 + y_2) = (2(x_1 + x_2), 0, (x_1 + x_2) + (y_1 + y_2)) \\ &= (2x_1, 0, x_1 + y_1) + (2x_2, 0, x_2 + y_2) \\ &= T(x_1, y_1) + T(x_2, y_2) \end{aligned}$$

e

$$\begin{aligned} T(\lambda(x_1, y_1)) &= T(\lambda x_1, \lambda y_1) = (\lambda 2x_1, 0, \lambda x_1 + \lambda y_1) \\ &= \lambda(2x_1, 0, x_1 + y_1) \\ &= \lambda T(x_1, y_1) \end{aligned}$$

Portanto,  $T$  é uma transformação linear. Vamos fazer também esse exemplo com a notação matricial, para



ilustrar os diferentes tipos de notação.

$$\begin{aligned}
 T\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= T\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} \\
 &= \begin{pmatrix} 2(x_1 + x_2) \\ 0 \\ (x_1 + x_2) + (y_1 + y_2) \end{pmatrix} \\
 &= \begin{pmatrix} 2x_1 + 2x_2 \\ 0 \\ (x_1 + x_2) + (y_1 + y_2) \end{pmatrix} \\
 &= \begin{pmatrix} 2x_1 \\ 0 \\ x_1 + y_1 \end{pmatrix} + \begin{pmatrix} 2x_2 \\ 0 \\ x_2 + y_2 \end{pmatrix} \\
 &= T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}
 \end{aligned}$$

e

$$T\begin{pmatrix} \lambda x_1 \\ \lambda y_1 \end{pmatrix} = \begin{pmatrix} \lambda 2x_1 \\ 0 \\ \lambda x_1 + \lambda y_1 \end{pmatrix} = \lambda \begin{pmatrix} 2x_1 \\ 0 \\ x_1 + y_1 \end{pmatrix} = \lambda T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

*Funções lineares* são as associações mais simples que podemos pensar. Por exemplo, imagine que de um quilograma de soja são extraídos 0,2 litros de óleo. Desse modo, de uma produção de  $x$  kg de soja, seriam extraídos  $0,2x$  litros de óleo, e com isso, poderíamos escrever a função  $T(x) = 0,2x$ , em que  $T$  é a quantidade em litros de óleo de soja e  $x$  é a quantidade de soja em kg.

Agora, vejamos algumas características interessantes desse simples exemplo. Primeiro note que para calcular uma produção de  $(x + y)$  kg de soja, podemos tanto multiplicar  $(x + y)$  pelo coeficiente de produção 0,2, como calcular as produções separadamente e depois somá-las, isto é,  $T(x+y) = 0,2(x+y) = 0,2x + 0,2y = T(x) + T(y)$ . Em segundo lugar, veja que se a quantidade de soja for multiplicada por um fator  $\lambda$  (representando um escalar), a produção de óleo também será afetada por esse fator, isto é,  $T(\lambda x) = 0,2(\lambda x) = \lambda 0,2x = \lambda T(x)$ .

Essas duas propriedades caracterizam o que chamamos de *Transformações Lineares*. Isto é, uma transformação linear  $T : U \rightarrow V$  é uma função que leva elementos de  $U$  em elementos de  $V$ , que satisfaz as duas propriedades abaixo:

1.  $T(u + v) = T(u) + T(v)$ ,  $\forall u, v \in U$ ;
2.  $T(\lambda u) = \lambda T(u)$ ,  $\forall u \in U$  e  $\forall \lambda \in \mathbb{R}$ .

**Exemplo 4.1.6** *Boldrini (1984)* A quantidade em litros de óleo extraída por quilograma de cereal, segundo um determinado processo, pode ser descrita pela tabela

|          | Soja | Milho | Algodão | Amendoim |
|----------|------|-------|---------|----------|
| Óleo (l) | 0,2  | 0,06  | 0,13    | 0,32     |

A quantidade total de óleo produzido por  $x$  kg de soja,  $y$  kg de milho,  $z$  kg de algodão e  $w$  kg de amendoim é dada por  $T(x, y, z, w) = 0,2x + 0,06y + 0,13z + 0,32w$ , que cumpre as propriedades da soma e multiplicação:

1.

$$\begin{aligned}T((x_1, y_1, z_1, w_1) + (x_2, y_2, z_2, w_2)) &= T((x_1 + x_2, y_1 + y_2, z_1 + z_2, w_1 + w_2)) \\&= 0, 2(x_1 + x_2) + 0, 06(y_1 + y_2) + 0, 13(z_1 + z_2) + 0, 32(w_1 + w_2) \\&= 0, 2x_1 + 0, 06y_1 + 0, 13z_1 + 0, 32w_1 \\&+ 0, 2x_2 + 0, 06y_2 + 0, 13z_2 + 0, 32w_2 \\&= T(x_1, y_1, z_1, w_1) + T(x_2, y_2, z_2, w_2)\end{aligned}$$

2.

$$\begin{aligned}T(\lambda(x_1, y_1, z_1, w_1)) &= T((\lambda x_1, \lambda y_1, \lambda z_1, \lambda w_1)) \\&= 0, 2(\lambda x_1) + 0, 06(\lambda y_1) + 0, 13(\lambda z_1) + 0, 32(\lambda w_1) \\&= \lambda(0, 2x_1 + 0, 06y_1 + 0, 13z_1 + 0, 32w_1) \\&= \lambda T(x_1, y_1, z_1, w_1)\end{aligned}$$

Portanto,  $T$  é uma transformação linear de  $\mathbb{R}^4$  em  $\mathbb{R}$ .

É também comum representar tais aplicações nas suas formas “matriciais”, isto é, a aplicação do exemplo anterior pode ser denotada por:

$$T \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 0, 2x + 0, 06y + 0, 13z + 0, 32w,$$

e portanto, temos:

$$T \left( \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{pmatrix} \right) = T \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} + T \begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{pmatrix}$$

e

$$T \begin{pmatrix} \lambda x_1 \\ \lambda y_1 \\ \lambda z_1 \\ \lambda w_1 \end{pmatrix} = \lambda T \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix}$$

**Exemplo 4.1.7** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = 2x$ . Como  $T(x + y) = 2(x + y) = 2x + 2y = T(x) + T(y)$ , e também,  $T(\lambda x) = 2(\lambda x) = \lambda 2x = \lambda T(x)$ , então segue que  $T$  é uma transformação linear de  $\mathbb{R}$  em  $\mathbb{R}$ .

O exemplo acima exemplifica que aplicações que vão de  $\mathbb{R}$  em  $\mathbb{R}$ , descritas por  $T(x) = \alpha x$  são sempre transformações lineares. No entanto, aplicações do tipo “afim” não cumprem essa propriedade. Vejamos:

**Exemplo 4.1.8** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = 2x + 3$ . Por um lado, temos que  $T(x + y) = 2(x + y) + 3$ . Por outro lado, temos que  $T(x) + T(y) = 2x + 3 + 2y + 3 = 2x + 2y + 6$ , que claramente são diferentes. Dessa forma, não temos que a propriedade da soma é cumprida, e assim,  $T$  não é uma transformação linear.

Além de funções afins, temos que funções polinomiais (diferentes das lineares) não têm a característica de uma transformação linear. Vejamos:

**Exemplo 4.1.9** Considere a aplicação  $T : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $T(x) = x^2$ . Por um lado, temos que  $T(x + y) = (x + y)^2 = x^2 + 2xy + y^2$ . Por outro lado, temos que  $T(x) + T(y) = x^2 + y^2$ . Dessa forma, a aplicação  $T$  não é uma transformação linear.

Uma pergunta que surge é a seguinte: Como podemos identificar possíveis aplicações que não são transformações lineares, sem ter que de fato verificar as propriedades? Uma possível forma de abordar essa questão é que se  $T$  é uma transformação linear, então devemos ter que  $T(0) = 0$ . Isso ocorre porque  $T(0) = T(0 + 0) = T(0) + T(0)$ , e assim,  $T(0) = T(0) + T(0)$ , implicando que  $T(0) = 0$ .

Perceba que a aplicação dada no **Exemplo 4.1.8** não satisfaz  $T(0) = 0$ , de fato,  $T(0) = 2 \cdot 0 + 3 = 3 \neq 0$ . Isso implica que  $T$  não é uma transformação linear. Porém, veja que no **Exemplo 4.1.9** a aplicação  $T(x) = x^2$  satisfaz a condição  $T(0) = 0$ , no entanto, vimos que não é uma transformação linear. Portanto, o único resultado que temos é que se uma aplicação não satisfaz a condição  $T(0) = 0$ , então  $T$  não é uma transformação linear. Agora, caso  $T$  satisfaça a condição  $T(0) = 0$ , então não necessariamente podemos concluir que  $T$  é linear, faltando analisar se de fato  $T$  cumpre as propriedades da soma e multiplicação por escalar.

Veja que também podemos ter aplicações que levam elementos de um espaço  $U$  em outro espaço  $V$  distintos.

**Exemplo 4.1.10** Considere a aplicação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  dada por  $T(x, y) = (2x, 0, x + y)$ . Assim,

$$\begin{aligned} T((x_1, y_1) + (x_2, y_2)) &= T(x_1 + x_2, y_1 + y_2) = (2(x_1 + x_2), 0, (x_1 + x_2) + (y_1 + y_2)) \\ &= (2x_1, 0, x_1 + y_1) + (2x_2, 0, x_2 + y_2) \\ &= T(x_1, y_1) + T(x_2, y_2) \end{aligned}$$

e

$$\begin{aligned} T(\lambda(x_1, y_1)) &= T(\lambda x_1, \lambda y_1) = (\lambda 2x_1, 0, \lambda x_1 + \lambda y_1) \\ &= \lambda(2x_1, 0, x_1 + y_1) \\ &= \lambda T(x_1, y_1) \end{aligned}$$

Portanto,  $T$  é uma transformação linear. Vamos fazer também esse exemplo com a notação matricial, para

ilustrar os diferentes tipos de notação.

$$\begin{aligned}
 T\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= T\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} \\
 &= \begin{pmatrix} 2(x_1 + x_2) \\ 0 \\ (x_1 + x_2) + (y_1 + y_2) \end{pmatrix} \\
 &= \begin{pmatrix} 2x_1 + 2x_2 \\ 0 \\ (x_1 + y_1) + (x_2 + y_2) \end{pmatrix} \\
 &= \begin{pmatrix} 2x_1 \\ 0 \\ x_1 + y_1 \end{pmatrix} + \begin{pmatrix} 2x_2 \\ 0 \\ x_2 + y_2 \end{pmatrix} \\
 &= T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + T\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}
 \end{aligned}$$

e

$$T\begin{pmatrix} \lambda x_1 \\ \lambda y_1 \end{pmatrix} = \begin{pmatrix} \lambda 2x_1 \\ 0 \\ \lambda x_1 + \lambda y_1 \end{pmatrix} = \lambda \begin{pmatrix} 2x_1 \\ 0 \\ x_1 + y_1 \end{pmatrix} = \lambda T\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

## 4.2 Núcleo e Imagem de uma Transformação Linear

Já vimos quais características uma aplicação  $T : U \rightarrow V$  precisa ter, para ser chamada de transformação linear. Agora, vamos estudar algumas ferramentas que envolvem transformações lineares. A primeira delas é chamada de *núcleo*. O núcleo (denotado por  $\text{Ker}(T)$ ) de uma transformação linear  $T : U \rightarrow V$  é definido como um subconjunto de  $U$  (domínio), de todos os elementos que são levados no 0 de  $V$ , em outras palavras,

$$\text{Ker}(T) = \{x \in U : T(x) = 0\}.$$

**Exemplo 4.2.1** Considere a transformação linear  $T(x, y) = (0, y - 2x, 0)$ . Logo, o núcleo de  $T$  é dado por

$$(x, y) \in \text{Ker}(T) \iff T(x, y) = (0, 0, 0) \iff (0, y - 2x, 0) = (0, 0, 0)$$

implicando que  $y - 2x = 0$ , e assim,  $y = 2x$ . Isso significa que o núcleo da transformação  $T$  é formado por elementos do tipo  $(x, y) = (x, 2x)$ , isto é,

$$\text{Ker}(T) = \{(x, 2x) : x \in \mathbb{R}\}.$$

Note que  $(1, 2)$  é um gerador do núcleo, já que  $(x, 2x) = x(1, 2)$  para todo  $x \in \mathbb{R}$ . Nesse caso, também podemos escrever  $\text{Ker}(T) = \{(x, 2x) : x \in \mathbb{R}\} = [(1, 2)]$ . A Figura 4.1 ilustra a transformação  $T$  levando os elementos que se encontram na reta  $y = 2x$  no  $(0, 0, 0)$ .

Uma pergunta natural que surge é a seguinte, por que é importante estudar o núcleo de uma transformação? A partir do núcleo é possível estudar, por exemplo, decomposição de espaços e obter algumas informações sobre transformações lineares. Existe um resultado na matemática que diz o seguinte:  $T$  é uma transformação linear injetora se, e somente se,  $\text{Ker}(T) = \{0\}$ .

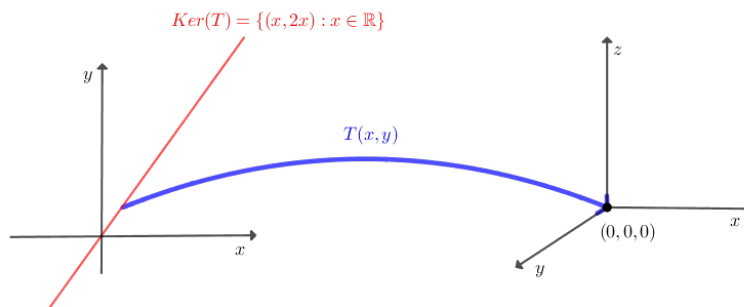


Figura 4.1: Núcleo (em vermelho) da transformação  $T(x, y) = (0, y - 2x, 0)$  (em azul).

A transformação linear dada no **Exemplo 4.2.1** não é injetora, uma vez que existem elementos  $(x, y) \neq (0, 0)$  que pertencem ao  $Ker(T)$ . Por exemplo,  $(1, 2)$  pertence ao  $Ker(T)$ , já que  $(1, 2)$  pode ser escrito na forma  $(x, 2x)$ , e assim  $T(1, 2) = (0, 2 - 2(1), 0) = (0, 0, 0)$ . Vejamos um exemplo de transformação linear injetora.

**Exemplo 4.2.2** Considere a transformação linear  $T(x, y) = (x, y, x + y)$ . Logo, o núcleo de  $T$  é dado por

$$(x, y) \in Ker(T) \iff T(x, y) = (0, 0, 0) \iff (x, y, x + y) = (0, 0, 0),$$

obtendo o seguinte sistema linear

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

Portanto, temos que o único elemento que pertence ao núcleo de  $T$  é o  $(0, 0)$ , isto é,

$$Ker(T) = \{(0, 0)\}.$$

A Figura 4.2 ilustra a transformação  $T$  que leva  $(0, 0)$  em  $(0, 0, 0)$ .

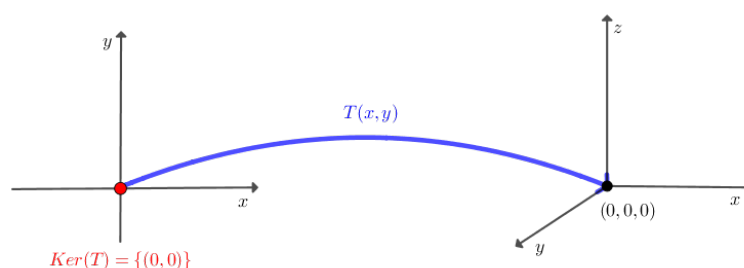


Figura 4.2: Núcleo (em vermelho) da transformação  $T(x, y) = (x, y, x + y)$  (em azul).

Sendo assim, como  $Ker(T) = \{(0, 0)\}$  então  $T$  é uma transformação linear injetora.

É importante frisar que é válida a recíproca do resultado, isto é, se  $T$  for uma transformação linear, então o núcleo só terá o elemento neutro do espaço  $U$ . Para ilustrar esse fato, considere o seguinte exemplo.

**Exemplo 4.2.3** Seja  $T(x, y) = (x, 2x + y)$  uma transformação linear. Vejamos se  $T$  é uma aplicação injetora. Para isso, considere  $(x, y)$  e  $(u, v)$  elementos quaisquer de  $U = \mathbb{R}^2$ , que satisfazem a propriedade  $T(x, y) =$

$T(u, v)$ . Se mostrarmos que  $(x, y) = (u, v)$  então concluiremos que  $T$  é injetora. Pois bem,

$$\begin{aligned} T(x, y) &= T(u, v) \\ (x, 2x + y) &= (u, 2u + v) \end{aligned}$$

obtendo o seguinte sistema linear

$$\begin{cases} x = u \\ 2x + y = 2u + v \end{cases}$$

Da primeira equação obtemos  $x = u$ , e portanto, substituindo  $x = u$  na segunda equação obtemos que  $y = v$ . Concluímos então que  $(x, y) = (u, v)$ , e assim,  $T$  é injetora.

Por fim, perceba que o núcleo de  $T$  é composto apenas pelo  $(0, 0)$ , pois

$$T(x, y) = (0, 0) \iff (x, 2x + y) = (0, 0)$$

concluindo que  $(x, y) = (0, 0)$ .

No exemplo anterior é possível perceber que estudar a injetividade de uma transformação linear através do núcleo é muito mais simples do que verificar pela definição de função injetora.

Estudamos até aqui um importante subconjunto do domínio de uma transformação linear. Agora nosso foco é estudar um subconjunto do contradomínio de uma transformação. Dada uma transformação linear  $T : U \rightarrow V$ , o conjunto *Imagem*, denotado por  $Im(T)$  é um subconjunto de  $V$  definido por todos os elementos de  $V$  que são obtidos através de algum elemento de  $U$ , por meio da aplicação  $T$ , isto é,

$$Im(T) = \{T(u) : u \in U\}.$$

**Exemplo 4.2.4** Seja  $T(x, y) = (x, 0)$  uma transformação linear. Então o conjunto imagem de  $T$  é dado por todos os elementos da forma  $(x, 0)$ , isto é, todos os elementos de  $\mathbb{R}^2$ , cuja segunda coordenada é 0. Em outras palavras,

$$Im(T) = \{T(x, y) : (x, y) \in \mathbb{R}^2\} = \{(x, 0) : x \in \mathbb{R}\}$$

Perceba que um gerador para o conjunto imagem dessa transformação linear é dado pelo vetor  $(1, 0)$ , já que  $(x, 0) = x(1, 0)$  para todo  $x \in \mathbb{R}$ . Sendo assim, podemos escrever  $Im(T) = \{(x, 0) : x \in \mathbb{R}\} = [(1, 0)]$ . A Figura 4.3 ilustra a imagem dessa aplicação.

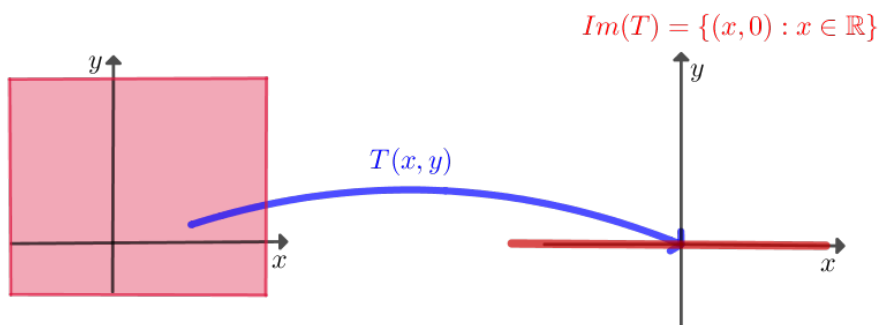


Figura 4.3: Imagem (reta em vermelho) da transformação  $T(x, y) = (x, 0)$  (em azul).

Os Exemplos 4.2.1 e 4.2.4 mostram que os subconjuntos  $Ker(T)$  e  $Im(T)$  podem ser gerados

por alguns elementos específicos. Na verdade, o núcleo e a imagem de uma transformação linear são subespaços de  $U$  e  $V$ , respectivamente. Sendo assim, o conjunto  $B = \{(1, 2)\}$  forma uma base para o núcleo da transformação linear dada no **Exemplo 4.2.1**, enquanto que  $B = \{(1, 0)\}$  forma uma base para a imagem da transformação linear dada no **Exemplo 4.2.4**.

**Exemplo 4.2.5** Considere a transformação linear  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  dada por  $T(x, y, z) = (x, 2y, 0)$ . Vamos determinar o núcleo e a imagem dessa transformação. Primeiro, vamos determinar a imagem. Para isso, note que

$$\begin{aligned} (x, y, z) \in \text{Im}(T) &\iff T(x, y, z) = (x, 2y, 0) \\ &= x(1, 0, 0) + y(0, 2, 0) \end{aligned}$$

logo o conjunto  $B_1 = \{(1, 0, 0), (0, 2, 0)\}$  gera todos os elementos da imagem. Além disso, como o conjunto  $B_1$  é linearmente independente, temos que  $B_1$  é uma base para  $\text{Im}(T)$ .

Agora vejamos o núcleo dessa transformação linear.

$$(x, y, z) \in \text{Ker}(T) \iff T(x, y, z) = (0, 0, 0) \iff (x, 2y, 0) = (0, 0, 0)$$

implicando que  $x = 0$  e  $y = 0$ . Note então que um elemento qualquer que pertence ao núcleo tem a forma  $(x, y, z) \in \text{Ker}(T) \iff (x, y, z) = (0, 0, z) = z(0, 0, 1)$ . Portanto,  $B_2 = \{(0, 0, 1)\}$  é um conjunto gerador do núcleo, e mais, como  $B_2$  é um conjunto linearmente independente, temos que  $B_2$  é uma base para o núcleo.

É interessante notar que no **Exemplo 4.2.5** temos dois elementos na base da imagem e um elemento na base do núcleo. Nesse caso, dizemos que a *dimensão* da imagem é igual a 2 e a *dimensão* do núcleo é igual a 1, isto é, a *dimensão* de um (sub)espaço vetorial  $V$  é definido pela quantidade de vetores na base, e denotamos por  $\dim(V)$ .

Além disso, note que a dimensão de  $V = \mathbb{R}^3$  é igual a 3, já que  $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  é uma base para  $\mathbb{R}^3$ . Com isso, é possível perceber que a soma das dimensões dos núcleo e da imagem é igual a dimensão do espaço  $V$ . Esse resultado ocorre em geral, isto é, dada uma transformação linear  $T : U \rightarrow V$ , temos

$$\dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V) \tag{4.1}$$

Essa relação nos traz um importante resultado. Se uma transformação linear  $T : U \rightarrow U$  for injetora, significa que  $\text{Ker}(T) = \{(0_U)\}$ , e portanto,  $\dim(\text{Ker}(T)) = 0$ . Com isso, a equação (4.1) nos informa que  $\dim(\text{Im}(T)) = \dim(U)$ . Isso significa que como  $\text{Im}(T) \subseteq U$  e possuem a mesma quantidade de geradores, temos que  $\text{Im}(T) = U$ . Portanto, se  $T$  for injetora, concluímos que  $T$  também é sobrejetora (e vice-versa).

Toda transformação linear cujo domínio é igual ao contradomínio, isto é,  $T : U \rightarrow U$ , é chamada de *operador linear*.

### 4.3 Operadores lineares

Na última discussão exploramos os conceitos de transformações lineares injetoras e sobrejetoras. Por exemplo, a transformação linear  $T : \mathbb{R}^2 \rightarrow \mathbb{C}$ , dada por  $T(x, y) = x + iy$  é uma transformação linear injetora e sobrejetora, já que:

- $T(x, y) = 0 + 0i$  implica que  $x + iy = 0 + 0i$ , e portanto,  $x = 0$  e  $y = 0$ , isto é,

$$\text{Ker}(T) = \{(x, y) \in \mathbb{R}^2 : T(x, y) = 0 + 0i\} = \{(0, 0)\}$$

concluindo que  $T$  é injetora.

- Dado um elemento  $a + bi$  qualquer de  $\mathbb{C}$ , existe  $(a, b) \in \mathbb{R}^2$ , tal que  $T(a, b) = a + bi$ . Isto é,  $\text{Im}(T) = \mathbb{C}$ . Portanto,  $T$  é sobrejetora.

Quando uma transformação linear  $T$  satisfaz a condição de injetividade e sobrejetividade (ou seja, é bijetora), dizemos que  $T$  é um *isomorfismo* entre os espaços  $\mathbb{R}^2$  e  $\mathbb{C}$ . Do ponto de vista matemático, isso significa que mesmo sendo diferentes um do outro, eles podem ser “visualizados” da mesma forma (veja Figura 4.4). Existem conexões interessantes entre espaços *isomorfos*, por exemplo, as dimensões de espaços isomorfos são sempre iguais. Note que  $B_1 = \{(1, 0), (0, 1)\}$  é uma base para  $\mathbb{R}^2$ , enquanto que  $B_2 = \{1, i\}$  é uma base para  $\mathbb{C}$ , e assim,  $\dim(\mathbb{C}) = \dim(\mathbb{R}^2) = 2$ .

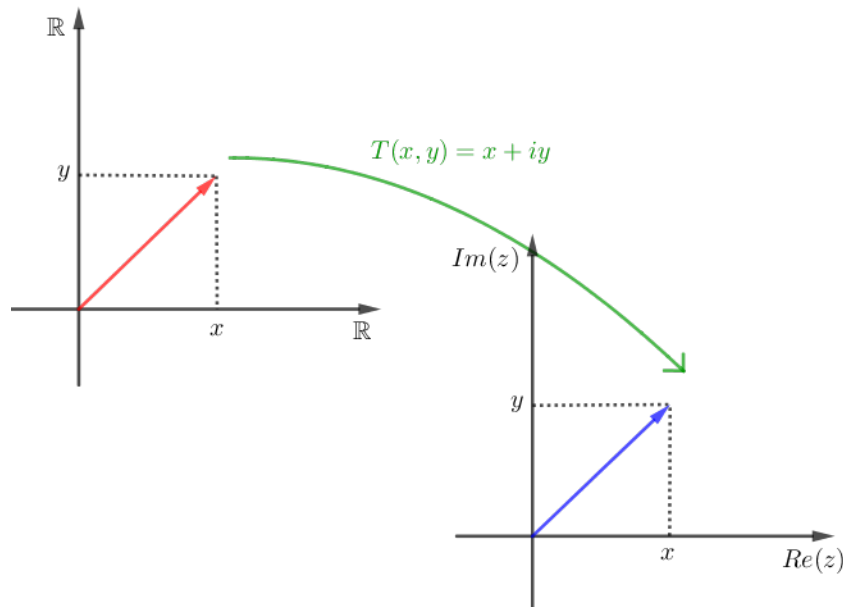


Figura 4.4: Transformação  $T(x, y) = x + iy$  (em verde), que leva um elemento qualquer  $(x, y)$  do  $\mathbb{R}^2$  (em vermelho) em um elemento  $z = x + iy$  dos números complexos  $\mathbb{C}$  (em azul). A notação  $\text{Im}(z)$  e  $\text{Re}(z)$  representa a parte imaginária e parte real do número complexo  $z$ , respectivamente.

Como mencionado anteriormente, uma aplicação  $T : U \rightarrow V$  é chamada de operador linear, se  $T$  for uma transformação linear, e mais,  $U = V$ . O exemplo acima se trata de uma transformação linear que não é um operador, já que  $\mathbb{R}^2 \neq \mathbb{C}$ . Por outro lado,  $T(x, y) = (x - y, 2x + y)$  é um operador linear já que, neste caso,  $U = V = \mathbb{R}^2$ . Do ponto de vista de visualização, podemos representar essa aplicação em um mesmo espaço. Por exemplo,  $T(1, 2) = (-1, 4)$ , sendo assim, podemos representar a associação entre  $(1, 2)$  e  $(-1, 4)$  em um “mesmo espaço”, assim como na Figura 4.5.

No próximo exemplo veremos um caso interessante de operador linear. Para isso, vamos novamente frisar que um vetor  $u = (x, y)$  também pode ser denotado em sua forma matricial como um “vetor coluna”, isto é:

$$u = \begin{bmatrix} x \\ y \end{bmatrix}.$$



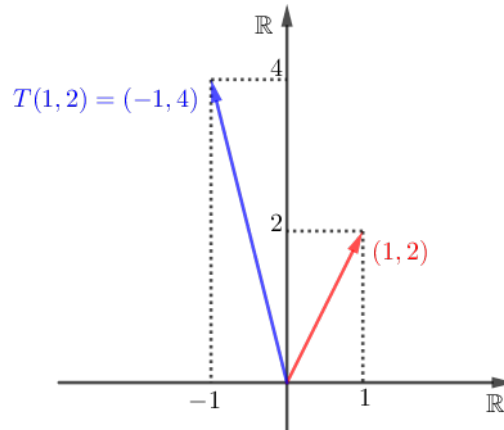


Figura 4.5: Transformação  $T(x, y) = (x - y, 2x + y)$ , que leva um elemento  $(1, 2)$  (em vermelho) em um elemento  $(-1, 4)$  (em azul).

**Exemplo 4.3.1** Considere uma matriz  $2 \times 2$  dada por:

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

Fixada essa matriz, definimos a seguinte aplicação  $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por:

$$T_A \left( \begin{bmatrix} x \\ y \end{bmatrix} \right) = A \begin{bmatrix} x \\ y \end{bmatrix}.$$

Vejamos como essa aplicação atua sobre  $\mathbb{R}^2$ .

$$T_A \left( \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \cdot 2 + 1 \cdot 3 \\ -1 \cdot 2 + 0 \cdot 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

ou ainda

$$T_A \left( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 1 \cdot 0 \\ -1 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

De um modo geral, temos que:

$$T_A \left( \begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \cdot x + 1 \cdot y \\ -1 \cdot x + 0 \cdot y \end{bmatrix} = \begin{bmatrix} y \\ -x \end{bmatrix}.$$

Note então que essa aplicação leva de fato elementos de  $\mathbb{R}^2$  em elementos de  $\mathbb{R}^2$ . Além disso, a aplicação  $T_A$  é uma transformação linear, e portanto, concluímos que  $T_A$  é um operador.

O operador linear dado no exemplo acima poderia também ser escrito na forma que estamos acostumados a trabalhar, ou seja,  $T(x, y) = (y, -x)$ . Isto é, se um operador é dado na forma matricial podemos recuperar a sua forma “vetorial”.

**Exercício 4.3.1** A recíproca da afirmação acima é verdadeira? Isto é, uma transformação linear na forma “vetorial” pode ser escrita na forma matricial?

## Capítulo 5

# Autovalores e autovetores

### 5.1 Autovalores e autovetores

Dizemos que o escalar  $\lambda$  é um *autovalor* de  $A \in \mathbb{R}^{n \times n}$  se a matriz  $A - \lambda I$  é singular, ou seja, se  $A - \lambda I$  não admitir inversa. Em outras palavras, existe  $v \neq 0$  tal que

$$Av = \lambda v. \quad (5.1)$$

O vetor  $v \neq 0$  que verifica (5.1) se chama *autovetor* associado ao autovalor  $\lambda$ .

Uma matriz não admite inversa quando seu determinante é igual a zero. Com isso, temos que  $\lambda$  será um autovalor de  $A$  se, e somente se,

$$\det(A - \lambda I) = 0. \quad (5.2)$$

A expressão  $\det(A - \lambda I)$ , quando desenvolvida, se revela um polinômio de grau  $n$ . Portanto, a equação (5.2), chamada *equação característica* e tem a forma:

$$c_n \lambda^n + c_{n-1} \lambda^{n-1} + \dots + c_1 \lambda + c_0 = 0.$$

Equações deste tipo podem ter raízes reais ou complexas. O *Teorema Fundamental da Álgebra* diz que o número total de raízes, reais ou complexas, contando suas multiplicidades, é igual a  $n$ .

Felizmente, quando a matriz  $A$  é simétrica, isto é quando  $A = A^t$ , a equação característica somente tem raízes reais. A outra propriedade fundamental de autovalores e autovetores de matrizes simétricas é que autovetores correspondentes a autovalores diferentes são ortogonais. Para a leitora que tiver mais curiosidade sobre o resultado, indicamos a consulta de [Martínez \(2004\)](#).

As matrizes de Pauli têm uma grande importância na área de mecânica quântica, descrevendo a interação do spin de uma partícula. Tais matrizes são dadas abaixo:

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Para ilustrar o processo de obtenção de tais autovalores, vamos determinar os autovalores e au-

tovetores associados a tais matrizes. Para obter os autovalores associados a uma determinada matriz  $A$ , é preciso resolver o seguinte problema  $Av = \lambda v$ , isto é, determinar valores numéricos  $\lambda$  de tal forma que o produto  $Av$  resulte em um múltiplo de  $v$ .

**Exemplo 5.1.1 (Autovalores e autovetores da matriz de Pauli  $\sigma_1$ )** Considere  $A = \sigma_1$ . Assim,

$$\begin{aligned} A - \lambda I &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix}. \end{aligned}$$

Para que a matriz acima seja singular, devemos calcular o determinante e igualar a 0. Assim,

$$(-\lambda)^2 - 1 = 0 \Rightarrow \lambda^2 = 1 \Rightarrow \lambda = \pm 1$$

Vamos determinar agora os autovetores associados a  $\lambda = 1$ . Para isso, vamos substituir esse valor na matriz  $(A - \lambda I)$ , multiplicar por um vetor  $v$  e igualar a matriz coluna composta por zeros. Assim,

$$(A - 1I)v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} -x + y = 0 \\ x - y = 0 \end{cases}$$

Portanto, concluímos que  $y = x$ , e assim, os autovetores associados a  $\lambda = 1$  têm a forma

$$\begin{pmatrix} x \\ x \end{pmatrix}, \quad x \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} x \\ x \end{pmatrix} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \dots \right\}$$

Em geral, escolhemos um "representante" para esses autovetores. Nesse caso específico, vamos considerar

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

mas poderia ser qualquer outro representante, dependendo da conveniência.

Façamos agora o mesmo para o autovalor  $\lambda = -1$ .

$$(A - (-1)I)v = 0 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} x + y = 0 \\ x + y = 0 \end{cases}$$

Portanto, concluímos que  $y = -x$ , e assim, os autovetores associados a  $\lambda = -1$  têm a forma

$$\begin{pmatrix} x \\ -x \end{pmatrix}, \quad x \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} x \\ -x \end{pmatrix} = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \dots \right\}$$

Como representante, poderíamos escolher

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Dizemos que  $(\lambda, v)$  é um autopar de  $A$  e o conjunto formado pelos autovetores  $W_\lambda = \{v \in \mathbb{R}^n : (A - \lambda I)v = 0\}$  é chamado de autoespaço.

**Exemplo 5.1.2 (Autovalores e autovetores da matriz de Pauli  $\sigma_2$ )** Considere  $A = \sigma_2$ . Assim,

$$\begin{aligned} A - \lambda I &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\lambda & -i \\ i & -\lambda \end{pmatrix}. \end{aligned}$$

Assim, o polinômio característico é dado por

$$(-\lambda)^2 - (-i^2) = 0 \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = \pm 1$$

Vamos determinar agora os autovetores associados a  $\lambda = 1$ . Assim,

$$(A - 1I)v = 0 \Rightarrow \begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} a + bi \\ c + di \end{pmatrix} = \begin{pmatrix} 0 + 0i \\ 0 + 0i \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} -a - bi - ci - di^2 = 0 + 0i \\ ai + bi^2 - c - di = 0 + 0i \end{cases} \sim \begin{cases} -a - bi - ci + d = 0 + 0i \\ ai - b - c - di = 0 + 0i \end{cases} \sim \begin{cases} -a + d + (-b - c)i = 0 + 0i \\ -b - c + (a - d)i = 0 + 0i \end{cases}$$

Perceba que se multiplicarmos a primeira equação por  $i$ , obtemos como resposta a segunda equação. Sendo assim, descartamos a segunda equação e ficamos apenas com a primeira, obtendo  $a = d$  e

$c = -b$ . Portanto, os autovetores associados a  $\lambda = 1$  têm a forma

$$\begin{pmatrix} a + bi \\ -b + ai \end{pmatrix}, \quad a, b \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} a + bi \\ -b + ai \end{pmatrix} = \left\{ \begin{pmatrix} 1 + i \\ -1 + i \end{pmatrix}, \begin{pmatrix} 1 + 2i \\ -2 + i \end{pmatrix}, \begin{pmatrix} -1 + 3i \\ -3 - i \end{pmatrix}, \dots \right\}$$

Vamos obter agora os autovetores associados a  $\lambda = -1$ .

$$(A - (-1)I)v = 0 \Rightarrow \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \begin{pmatrix} a + bi \\ c + di \end{pmatrix} = \begin{pmatrix} 0 + 0i \\ 0 + 0i \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} a + bi - ci - di^2 = 0 + 0i \\ ai + bi^2 + c + di = 0 + 0i \end{cases} \sim \begin{cases} a + bi - ci + d = 0 + 0i \\ ai - b + c + di = 0 + 0i \end{cases} \sim \begin{cases} a + d + (b - c)i = 0 + 0i \\ -b + c + (a + d)i = 0 + 0i \end{cases}$$

Obtemos então as relações  $d = -a$  e  $c = b$ . Portanto, os autovetores associados a  $\lambda = -1$  têm a forma

$$\begin{pmatrix} a + bi \\ b - ai \end{pmatrix}, \quad a, b \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} a + bi \\ b - ai \end{pmatrix} = \left\{ \begin{pmatrix} 1 + i \\ 1 - i \end{pmatrix}, \begin{pmatrix} 1 + 2i \\ 2 - i \end{pmatrix}, \begin{pmatrix} -1 + 3i \\ 3 + i \end{pmatrix}, \dots \right\}$$

**Exemplo 5.1.3 (Autovalores e autovetores da matriz de Pauli  $\sigma_3$ )** Considere  $A = \sigma_3$ . Assim,

$$\begin{aligned} A - \lambda I &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 - \lambda & 0 \\ 0 & -1 - \lambda \end{pmatrix}. \end{aligned}$$

O polinômio característico é dado por

$$(1 - \lambda)(-1 - \lambda) = 0 \Rightarrow \lambda = \pm 1$$

Vamos determinar agora os autovetores associados a  $\lambda = 1$ .

$$(A - 1I)v = 0 \Rightarrow \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} 0x + 0y = 0 \\ 0x - 2y = 0 \end{cases}$$

Portanto, concluímos que  $y = 0$  e  $x$  pode assumir qualquer valor real. Logo, os autovetores associados a  $\lambda = 1$  têm a forma

$$\begin{pmatrix} x \\ 0 \end{pmatrix}, \quad x \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} x \\ 0 \end{pmatrix} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \dots \right\}$$

Para o autovalor  $\lambda = -1$ , temos

$$(A - (-1)I)v = 0 \Rightarrow \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Multiplicando as matrizes, obtemos o seguinte sistema linear:

$$\begin{cases} 2x + 0y = 0 \\ 0x + 0y = 0 \end{cases}$$

Portanto, concluímos que  $x = 0$  e  $y$  pode assumir qualquer valor real. Logo, os autovetores associados a  $\lambda = -1$  têm a forma

$$\begin{pmatrix} 0 \\ y \end{pmatrix}, \quad x \in \mathbb{R}$$

Isso significa que os autovetores são:

$$\begin{pmatrix} 0 \\ y \end{pmatrix} = \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \dots \right\}$$

**Exemplo 5.1.4** Uma matriz de rotação, como o nome indica, rotaciona um vetor por meio de um ângulo  $\theta$ , preservando seu "tamanho". Tal matriz é dada por

$$M_\theta = \begin{pmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{pmatrix}$$

Independente do ângulo  $\theta$ , essa matriz possui 2 autovalores. Vamos seguir o mesmo processo do exemplo anterior para obter os autovalores da matriz de rotação. Isto é,

$$\begin{aligned} M_\theta - \lambda I &= \begin{pmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta) - \lambda & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) - \lambda \end{pmatrix}. \end{aligned}$$

O polinômio característico dessa matriz é dado por

$$(\cos(\theta) - \lambda)^2 - (-\operatorname{sen}^2(\theta)) = \lambda^2 - 2\lambda\cos(\theta) + \underbrace{\cos^2(\theta) + \operatorname{sen}^2(\theta)}_{=1}$$

Assim, calculando as raízes (em relação a  $\lambda$ ) do polinômio  $\lambda^2 - 2\lambda\cos(\theta) + 1$  a partir da fórmula quadrática, obtemos

$$\begin{aligned}\lambda &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{2\cos(\theta) \pm \sqrt{4\cos^2(\theta) - 4}}{2} \\ &= \frac{2\cos(\theta) \pm \sqrt{4(\cos^2(\theta) - 1)}}{2} \\ &= \frac{2\cos(\theta) \pm 2\sqrt{\cos^2(\theta) - 1}}{2} \\ &= \cos(\theta) \pm \sqrt{\cos^2(\theta) - 1}\end{aligned}$$

Como  $\cos^2(\theta) + \operatorname{sen}^2(\theta) = 1$ , então  $\cos^2(\theta) - 1 = -\operatorname{sen}^2(\theta) = i^2\operatorname{sen}^2(\theta)$ . Assim, substituindo na equação acima, obtemos os autovalores:

$$\begin{aligned}\lambda &= \cos(\theta) \pm \sqrt{i^2\operatorname{sen}^2(\theta)} \\ &= \cos(\theta) \pm i\operatorname{sen}(\theta)\end{aligned}$$

A representação  $e^{i\theta} = \cos(\theta) + i\operatorname{sen}(\theta)$  é chamada de fórmula de Euler e estabelece uma conexão entre funções trigonométricas e funções exponenciais complexas. Essa representação é muito utilizada na área de Física, Engenharia e na área de análise complexa dentro da matemática.

**Exercício 5.1.1 (Um modelo esquemático de reprodução de células vermelhas sanguíneas (CVS))** *Barros and Esmi (2019)* No sistema circulatório, as CVS estão constantemente sendo destruídas e repostas. Uma vez que estas células transportam oxigênio pelo organismo, a sua quantidade deve ser mantida em algum nível fixo. Assumindo que uma fração dessas células é destruída diariamente pelo baço e que a medula óssea produz um número de CVS proporcional ao número perdido no dia anterior, pode-se modelar matematicamente essa dinâmica através do sistema de equações de diferenças abaixo:

$$\begin{cases} R_{t+1} = (1 - f)R_t + M_t \\ M_{t+1} = \gamma f R_t \end{cases},$$

sendo que

- $R_t$  é o número de CVS na circulação sanguínea no dia  $t$
- $M_t$  é o número de CVS produzidas pela medula óssea no dia  $t$
- $f$  é a fração de CVS removidas pelo baço
- $\gamma$  é a taxa de produção de CVS.

Os autovalores  $\lambda$  da matriz associada ao sistema cumprem a seguinte regra:

a) se o módulo dos autovalores for menor que 1, então o número de CVS tende a zero à medida que  $t$  aumenta;

b) se algum dos autovalores possui módulo maior que 1, então o número de CVS aumenta indefinidamente à medida que  $t$  aumenta;

c) se o módulo de um autovalor for menor que 1 e o módulo do outro autovalor for igual a 1, então o número de CVS atinge um valor constante.

Considere um exemplo hipotético em que essa dinâmica é dada pelo seguinte sistema:

$$\begin{cases} R_{t+1} = 0,25R_t + M_t \\ M_{t+1} = 0,75R_t \end{cases} .$$

Determine qual seria o comportamento de CVS na circulação sanguínea. Argumente como chegou nesse resultado.



## Capítulo 6

# Operadores lineares auto-adjuntos e ortogonais

Anteriormente vimos que uma transformação linear  $T$  é chamada de operador linear quando seu domínio é igual ao contradomínio. Isso permite, por exemplo, estudar autovalores e autovetores.

Agora, vamos estudar tipos importantes de operadores lineares que são muito utilizados na mecânica quântica (geralmente associados ao estudo de energia de um sistema), como por exemplo os operadores auto-adjuntos. Além desses operadores vamos estudar também os operadores ortogonais, que aparecem em problemas de dinâmica de corpos rígidos.

Antes de começarmos nosso estudo, vamos relembrar algumas ferramentas matemáticas. Uma matriz  $A$  é chamada de simétrica quando  $A = A^t$ . Por outro lado, uma matriz  $A$  é chamada de *ortogonal* quando  $AA^t = A^tA = I$ , sendo  $I$  a matriz identidade. Em outras palavras,  $A$  é uma matriz ortogonal quando sua inversa é dada por  $A^t$ .

**Exemplo 6.0.1** A matriz de rotação  $M_\theta$  dada por

$$M_\theta = \begin{pmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{pmatrix}$$

é uma matriz ortogonal, uma vez que

$$\begin{aligned} (M_\theta)(M_\theta)^t &= \begin{pmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\text{sen}(\theta) \\ \text{sen}(\theta) & \cos(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\theta) + \text{sen}^2(\theta) & -\cos(\theta)\text{sen}(\theta) + \cos(\theta)\text{sen}(\theta) \\ -\cos(\theta)\text{sen}(\theta) + \cos(\theta)\text{sen}(\theta) & \text{sen}^2(\theta) + \cos^2(\theta) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

e de modo similar, temos que  $(M_\theta)^t(M_\theta) = I$ .

O motivo de estarmos falando sobre matrizes é que sempre é possível associar uma matriz a uma transformação linear (e vice-versa). Por exemplo, considere o operador linear  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dado por

$$T(x, y) = (2x, -y).$$

Seendo  $B = \{(1, 0), (0, 1)\}$  a base canônica de  $\mathbb{R}^2$ , podemos escrever:

$$\begin{aligned} T(1, 0) &= (2, 0) = 2(1, 0) + 0(0, 1) \\ T(0, 1) &= (0, -1) = 0(1, 0) + (-1)(0, 1) \end{aligned}$$

e assim, obtemos:

$$[T]_B^B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

A matriz  $[T]_B^B$  é chamada de *matriz de transformação* da base  $B$  para a base  $B$ . Perceba também que a partir de  $[T]_B^B$  é possível obter a forma original da transformação linear  $T$ , fazendo:

$$\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ -y \end{pmatrix}$$

Isto é,  $T(x, y) = (2x, -y)$ . Isso nos diz que as duas formas são equivalentes, ou seja, tanto faz trabalhar com matrizes ou com a forma “vetorial” da transformação linear.

**Daqui em diante, a menos que se diga o contrário, consideraremos que as bases dos espaços vetoriais são ortonormais.**

## 6.1 Operadores ortogonais

Dizemos que um operador  $T : U \rightarrow U$  é um *operador ortogonal* se  $[T]_B^B$  for uma matriz ortogonal.

**Exemplo 6.1.1** Como visto no **Exemplo 6.0.1**, a matriz de rotação  $M_\theta$  é ortogonal, e portanto, o operador  $T_\theta(x, y) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta))$  é ortogonal.

## 6.2 Operadores auto-adjuntos

Dizemos que um operador  $T : U \rightarrow U$  é um *operador auto-adjunto* se  $[T]_B^B$  for uma matriz simétrica.

**Exemplo 6.2.1** Considere o operador linear  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dado por  $T(x, y) = (2x - 2y, -2x + 5y)$ . Então a matriz de transformação  $[T]_B^B$ , em que  $B$  é a base canônica de  $\mathbb{R}^2$ , é dada por:

$$[T]_B^B = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

Como

$$([T]_B^B)^t = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix} = [T]_B^B,$$

segue que  $[T]_B^B$  é uma matriz simétrica, e portanto,  $T$  é um operador auto-adjunto.

**Observação 6.2.1** Ressaltamos que os operadores ortogonais e auto-adjuntos independem da base, isto é, se  $[T]_B^B$  for ortogonal (auto-adjunto), então  $[T]_C^C$  também é ortogonal (auto-adjunto), contanto que  $C$  também seja uma base ortonormal.

## Referências

Barros, L.C. and Esmi, E., 2019. *Notas de aula de biomatemática i*, vol. 1. IMECC–UNICAMP.

Boldrini, J.L., 1984. *Álgebra linear*, vol. 1. Harbra.

Callioli, C.A., Domingues, H.H. and Costa, R.C.F., 1983. *Álgebra linear e aplicações*, vol. 1. Atual.

Martínez, J.M., 2004. *Notas de geometria analítica*. Campinas: IMECC-UNICAMP.

# Índice Remissivo

- autoespaço, 42
- autopar, 42
- autovalor, 40
- autovetor, 40
  
- base, 26
  
- combinação linear, 22
- conjunto gerador, 24
- coordenadas de um vetor, 26
  
- dimensão, 37
- distância entre vetores, 19
  
- eliminação Gaussiana, 4
- equação característica, 40
  
- funções lineares, 28, 31
- fórmula de Euler, 45
  
- grau de liberdade, 3
  
- imagem, 36
- isomorfismo, 38
  
- linearmente dependente, 22
- linearmente independente, 22
  
- matriz de transformação, 48
- matriz ortogonal, 47
  
- norma, 18
- núcleo, 34
  
- operador auto-adjunto, 48
- operador linear, 37
- operador ortogonal, 48
- origem, 14
  
- produto escalar, 19
  
- regra do paralelogramo, 16
  
- sistema linear, 2
  
- sistema linear escalonado, 4
- subespaço vetorial, 18
  
- Teorema Fundamental da Álgebra, 40
- transformação linear, 28, 31
  
- vetor, 13
- vetores equipolentes, 14