UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

Diagonalização de matrizes e suas aplicações

Carla Andrucioli Carnesecca

Dissertação de Mestrado do Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT)

ICMC USP
SÃO CARLOS

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Carla Andrucioli Carnesecca

Diagonalização de matrizes e suas aplicações

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Mestrado Profissional em Matemática em Rede Nacional. *EXEMPLAR DE DEFESA*

Área de Concentração: Mestrado Profissional em Matemática em Rede Nacional

Orientadora: Profa. Dra. Katia Andreia Gonçalves de Azevedo

USP – São Carlos
Junho de 2019

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

A289d

Andrucioli Carnesecca, Carla
Diagonalização de matrizes e suas aplicações / Carla Andrucioli Carnesecca; orientadora Kátia Andreia Gonçalves de Azevedo. -- São Carlos, 2019.
113 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Mestrado Profissional em Matemática em Rede Nacional) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2019.

1. Diagonalização de matriz. 2. Reconhecimento de cônicas. 3. Resolução de sistemas de equações diferenciais ordinárias . 4. Mudança de base. 5. Autovalores e autovetores. I. Andreia Gonçalves de Azevedo, Kátia, orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:
Gláucia Maria Saia Cristianini - CRB - 8/4938
Juliana de Souza Moraes - CRB - 8/6176

Carla Andrucioli Carnesecca

Diagonalization of matrices and applications

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of Mathematics Professional Master's Program. *EXAMINATION BOARD PRESENTATION COPY*

Concentration Area: Professional Master Degree Program in Mathematics in National Network

Advisor: Profa. Dra. Katia Andreia Gonçalves de Azevedo

USP – São Carlos
June 2019

Dedico este trabalho ao meu filho Arthur Vicente,
luz onde é escuridão, alegria onde é tristeza,
força para a conclusão desse projeto.

AGRADECIMENTOS

A Deus por estar sempre presente em minha vida, iluminando meus caminhos tornando tudo possível.

A minha mãe Sonia por se empenhar para proporcionar melhores oportunidades na minha vida.

Ao meu irmão Flávio por sempre me apoiar e cobrar.

A todos meus colegas do curso, pela amizade e companherismo, em especial à Lívia, Daniele, Paula e Rosa pelos momentos de estudos, descontrações e risadas.

Ao Diego, pelas risadas, choros, desabafos e cobranças, porque sem ele não teria conseguido terminar.Obrigada!

Agradeço aos professores do PROFMAT, pelo conhecimento compartilhado.

A Profa Dra Kátia Andreia Gonçalves de Azevedo, pela dedicação, compreenção, paciência e quanta paciência que conduziu a orientação deste trabalho. Muito, muito obrigada!

Agradeço à CAPES pelo apoio financeiro.

“Eu tentei 99 vezes e falhei, mas na centésima tentativa eu consegui, nunca desista de seus objetivos mesmo que esses pareçam impossíveis, a próxima tentativa pode ser a vitoriosa.”

(Albert Einstein)

RESUMO

CARNESECCA, C. L. **Diagonalização de matrizes e suas aplicações**. 2019. 113 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

O principal objetivo desse trabalho é apresentar a teoria necessária para compreender o processo de diagonalização de operadores lineares e, consequentemente, de matrizes, como uma técnica para resolver sistemas de equações diferenciais ordinárias lineares homogêneos com coeficientes constantes e para reconhecer cônicas não degeneradas, as elipses, hipérboles e parábolas.

Palavras-chave: Diagonalização de Operadores Lineares, Cônicas, Sistemas Lineares de Equações Diferenciais Ordinárias.

ABSTRACT

CARNESECCA, C. L. **Diagonalization of matrices and applications**. 2019. 113 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

The main goal of this work is showing theories that are necessaries to understand the process of diagonalization of linear operators and, consequently, matrices, as a technique for solving homogeneous linear ordinary differential equations with constant coefficients and recognize non-degenerate cones, ellipses, hyperbolas, and parabolas.

Keywords: Diagonalization of Linear Operators, Conics, Linear Systems of Ordinary Differential Equations.

LISTA DE ILUSTRAÇÕES

Figura 1 – Reflexão em torno do eixo Ox. 47
Figura 2 – Projeção 48
Figura 3 – Rotação 49
Figura 4 – Rotação 50
Figura 5 – Dilatação e contração por α em $\mathbb{R}$ 51
Figura 6 – Projeção de v ao longo de w. 59
Figura 7 – Elipse 81
Figura 8 – Elipse 82
Figura 9 – Hipérbole 84
Figura 10 – Hipérbole 84
Figura 11 – Parábola 86
Figura 12 – $y^2 = 4px$ 87
Figura 13 – $x^2 = -4px$ 87
Figura 14 – Parabola $\mathscr{P}$ e círculo C 88
Figura 15 – A elipse $\frac{(x-1)^2}{9} + \frac{(x-2)^2}{4} = 1$ 90
Figura 16 – $\{(x,y) \in \mathbb{R}^2 : (y+3)^2 = 8(x+1)\}$ 91
Figura 17 – Elipse $(x'+2)^2 + \frac{(y'-1)^2}{3} = 1$ 93
Figura 18 – Hiperbole - $\mathscr{H} = 3x^2 - 4\sqrt{3}xy - y^2 + 20y = 25$. 94
Figura 19 – Par de retas: $x^2 - 6xy - 7y^2 + 10x + 2y + 9 = 0$. 96
Figura 20 – Parábola $\mathscr{P} : x^2 + 2xy + y^2 - 2x + 2y + 3 = 0$. 97

LISTA DE TABELAS

Tabela 1 – Reflexões mais comuns em $\mathbb{R}^2$ 48
Tabela 2 – Projeções mais comum em $\mathbb{R}^2$ 49
Tabela 3 – Rotações em $\mathbb{R}^3$ cujos eixos de rotação são os eixos coordenados 50

SUMÁRIO

1 INTRODUÇÃO 21

2 PRÉ REQUISITOS 23
2.1 Espaços Vetoriais 23
2.1.1 Definição e Exemplos 23
2.1.2 Subespaços 29
2.1.3 Dependência e Independência Linear 32

3 TRANSFORMAÇÕES LINEARES E MATRIZES 41
3.1 Transformações Lineares 41
3.2 Matriz de uma Transformação Linear 44
3.2.1 Operadores Lineares em $\mathbb{R}^2$ e em $\mathbb{R}^3$ 47
3.2.2 Mudança de Base e Matrizes semelhantes 51

4 ESPAÇO COM PRODUTO INTERNO 55

5 DIAGONALIZAÇÃO DE OPERADORES 67
5.1 Operadores Diagonalizáveis 67
5.2 Matrizes Diagonalizáveis 76
5.3 Teorema Espectral para operadores simétricos 78

6 CÔNICAS 81
6.1 Reconhecimento de Cônicas 89

7 SISTEMAS DE EQUAÇÕES DIFERENCIAIS ORDINÁRIAS 99

8 CONCLUSÃO 111

REFERÊNCIAS 113

CAPÍTULO

1

INTRODUÇÃO

A Matemática é uma ciência que sempre esteve presente em nosso cotidiano, dos primórdios históricos da humanidade até os dias atuais. E. mesmo sem notarmos, está presente em várias áreas do conhecimento. A Álgebra Linear é um ramo da Matemática que tem muita aplicabilidade em diversos campos de estudos como a criptografia, programação linear, circuitos elétricos, produção mecânica de peças, modelos econômicos lineares, entre outros. Por ser uma disciplina versátil nas aplicações, a Álgebra Linear tem uma grande relevância, pois contribui para os avanços tecnológicos e científicos. O objetivo deste trabalho é entender conceitos básicos e essenciais de Álgebra Linear como espaços vetoriais, autovalor e autovetor de transformações lineares, tendo como enfoque principal o processo de diagonalização de matrizes e suas aplicações, em particular a identificação de cônicas não degeneradas e resolução de sistemas lineares de equações diferenciais ordinárias com coeficientes constantes. A diagonalização de matrizes não faz parte do currículo do Ensino Médio. No entanto, são vários os conteúdos do currículo do Ensino Médio onde esse fundamento pode ser utilizado. Um deles está relacionado com a identificação de cônicas não degeneradas: elipse, hipérbole e parábola, como citado.

Desta forma, consideramos ser importante para os professores de Matemática conhecer técnicas para reconhecer uma cônica não degenerada ou resolver um sistema de equações diferenciais, ampliando sua visão sobre as possíveis aplicações desta teoria. Há vários livros que abordam estes problemas, mas compilamos os resultados essenciais para a compreensão da teoria.

Assim surgiu a ideia deste trabalho, que consiste em auxiliar professores da educação básica, em sua prática, no conteúdo de Álgebra Linear.

O trabalho será apresentado da seguinte maneira.

No capítulo 2 abordamos os pré-requisitos necessários para entender onde as transformações lineares estão definidas, conceitos como: espaços vetoriais, subespaços vetoriais, dependência e independência linear, base e dimensão. O capítulo 3 trata das transformações lineares, os operadores lineares em $\mathbb{R}^2$ e em $\mathbb{R}^3$, mudança de base e matrizes semelhantes, obtendo os resultados necessários para o processo de diagonalização. O capítulo 4 aborda os conceitos de espaço vetorial com produto interno e no capítulo 5, tratamos essencialmente do processo de diagonalização de operadores lineares e, consequentemente, diagonalização de matrizes. No capítulo 6, descrevemos as cônicas e aplicamos a teoria diagonalização de matrizes, para o processo de identificação de cônicas não degeneradas. Por fim, no capítulo 7, uma aplicação desta teoria sobre diagonaliização é feita para resolvermos sistemas lineares homogêneos de equações diferenciais ordinárias com coeficientes constantes.

CAPÍTULO

2

PRÉ REQUISITOS

Neste capítulo faremos um estudo sobre os pré-requisitos necessários para o entendimento do processo de diagonalização de operadores lineares. Os tópicos aqui estudados podem ser encontrados em (ZANI,), (HEFEZ A.; SOUZA FERNANDES, 2012), (LIPSCHUTZ, 1994) e (STEINBRUCH, 1987).

2.1 Espaços Vetoriais

2.1.1 Definição e Exemplos

Nesta seção serão apresentados a definição de espaço vetorial e alguns exemplos.

Definição 1. Um espaço vetorial V é um conjunto onde estão definidas duas operações, uma chamada adição e outra chamada multiplicação por um escalar. Os elementos do conjunto V são chamados de vetores. A adição faz corresponder a cada par de vetores $u, v \in V$, um novo vetor $u + v \in V$, chamado a soma de u e v. A multiplicação por escalar, a cada número $\alpha \in \mathbb{R}$ e a cada vetor $v \in V$, faz corresponder um vetor $\alpha \cdot v$ ou αv, chamado produto de α por v. Essas operações devem satisfazer, para quaisquer $\alpha, \beta \in \mathbb{R}$ e $u, v, w \in V$, as seguintes condições :

1. $u + v = v + u$;
2. $(u + v) + w = u + (v + w)$;
3. Existe um vetor $0 \in V$, chamado vetor nulo, denotado também por $\vec{0}$, tal que $v + 0 = 0 + v$ para todo $v \in V$;
4. Para cada vetor $v \in V$ existe um vetor $-v \in V$, chamado inverso aditivo ou simétrico de v, tal que: $-v + v = v + (-v) = 0$;
5. $(\alpha + \beta)v = \alpha v + \beta v$;

6. $(\alpha\beta)v = \alpha(\beta v)$;

7. $\alpha(u+v) = \alpha u + \alpha v$;

8. $1 \cdot v = v$.

Exemplo 1. $V = \mathbb{R}^2 = \{(x,y)/x,y \in \mathbb{R}\}$ é um espaço vetorial com as operações de adição e multiplicação por um número real assim definidas.

$$(x_1,y_1)+(x_2,y_2) = (x_1+x_2, y_1+y_2)$$

$$\alpha(x_1,y_1) = (\alpha x_1, \alpha x_2)$$

Vamos verificar as oito propriedades que um espaço vetorial deve satisfazer. Para isso considere $u = (x_1,y_1)$, $v = (x_2,y_2)$ e $w = (x_3,y_3)$ elementos de $\mathbb{R}^2$. Assim,

1. $u+v = (x_1,y_1)+(x_2,y_2) = (x_1+x_2,y_1+y_2) \stackrel{*}{=} (x_2+x_1,y_2+y_1) = (x_2,y_2)+(x_1,y_1) = v+u$

2. $u+(v+w) = (x_1,y_1)+((x_2,y_2)+(x_3,y_3)) = (x_1,y_1)+(x_2+x_3,y_2+y_3) =$

   $= (x_1+(x_2+x_3),y_1+(y_2+y_3)) \stackrel{*}{=} ((x_1+x_2)+x_3,(y_1+y_2)+y_3) =$

   $= (x_1+x_2,y_1+y_2)+(x_3,y_3) = (u+v)+w$

3. Existe $\vec{0} = (0,0) \in \mathbb{R}^2$ onde 0 é o elemento neutro da adição em $\mathbb{R}$ tal que para todo $u = (x_1,y_1) \in \mathbb{R}^2$ temos:

   $u+\vec{0} = (x_1,y_1)+(0,0) = (x_1+0,y_1+0) \stackrel{*}{=} (x_1,y_1) = u$

4. Para todo $u = (x_1,y_1) \in \mathbb{R}^2$ existe $-u = (-x_1,-y_1) \in \mathbb{R}^2$ tal que:

   $u+(-u) = (x_1,y_1)+(-x_1,-y_1) = (x_1-x_1,y_1-y_1) \stackrel{*}{=} (0,0)$, onde $-x_1$ e $-y_1$ serão os opostos de x_1 e y_1 respectivamente em $\mathbb{R}$

5. $\alpha(\beta u) = \alpha(\beta(x_1,y_1)) = \alpha(\beta x_1,\beta y_1) = (\alpha(\beta x_1),\alpha(\beta y_1)) \stackrel{*}{=} ((\alpha\beta)x_1),(\alpha\beta)y_1)) = (\alpha\beta)(x_1,y_1) = (\alpha\beta)u$.

6. $(\alpha+\beta)u = (\alpha+\beta)(x_1,y_1) = ((\alpha+\beta)x_1,(\alpha+\beta)y_1)) \stackrel{*}{=} (\alpha x_1+\beta x_1,\alpha y_1+\beta y_1) =$

   $= (\alpha x_1,\alpha y_1)+(\beta x_1,\beta y_1) = \alpha u+\beta u$.

7. $\alpha(u+v) = \alpha((x_1,y_1)+(x_2,y_2)) = \alpha(x_1+x_2,y_1+y_2) = (\alpha(x_1+x_2),\alpha(y_1+y_2)) \stackrel{*}{=} (\alpha x_1+\alpha x_2,\alpha y_1+\alpha y_2) = = (\alpha x_1,\alpha y_1)+(\alpha x_2,\alpha y_2) = \alpha(x_1,y_1)+\alpha(x_2,y_2) = \alpha u+\alpha v$.

8. $1u = 1(x_1,y_1) = (1x_1,1y_1) \stackrel{*}{=} (x_1,y_1) = u$.

Quando usamos o sinal $\overset{*}{=}$ significa que nesta passagem estamosassumindo o fato de que essas propriedades são válidas para os números reais. Logo, $V = \mathbb{R}^2$, com as operações de adição e multiplicação definidas acima, é um espaço vetorial.

Observe que se considerarmos $\mathbb{R}^n = \{(x_1, \cdots, x_n)/x_i \in \mathbb{R}, i = 1, \cdots, n\}$ e as operações de adição e multiplicação por escalar definidas por $(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$ e $\alpha(x_1, \cdots, x_n) = (\alpha x_1, \cdots, \alpha x_n)$, com $\alpha \in \mathbb{R}$, usando o mesmo raciocínio, provamos que $(\mathbb{R}^n, +, \cdot)$ é um espaço vetorial.

Exemplo 2. $V = \mathbb{R}_+^* = \{x \in \mathbb{R}/x > 0\}$ é um espaço vetorial com as operações de adição e multiplicação por um número real assim definidas:

$$x \oplus y = x \cdot y,\ x, y \in \mathbb{R}_+^*,$$

$$\lambda \odot x = x^\lambda,\ x \in \mathbb{R}_+^*,\ \lambda \in \mathbb{R}.$$

Vamos verificar as oito propriedades e provar que $(\mathbb{R}_+^*, \oplus, \odot)$ é um espaço vetorial.

1. $x \oplus y = x \cdot y = y \cdot x = y \oplus x$

2. $(x \oplus y) \oplus z = (x \cdot y) \oplus z = ((x \cdot y) \cdot z) = x \cdot (y \cdot z) = x \oplus (y \cdot z) = x \oplus (y \oplus z)$

3. Observe que $x \oplus 1 = x \cdot 1 = x, \forall x \in \mathbb{R}_+^*$. Logo, 1 é o elemento neutro da adição, ou seja, $\vec{0} = 1$

4. Dado $x \in \mathbb{R}_+^*$, considere o inverso aditivo $-x$ como sendo o número $\frac{1}{x}$. Assim, $x \oplus (-x) = x \cdot (\frac{1}{x}) = 1 = \vec{0}$.

5. $\alpha \odot (\beta \odot x) = \alpha(x^\beta) = (x^\beta)^\alpha = x^{\alpha\beta} = (\alpha\beta) \odot x$

6. $(\alpha + \beta) \odot x = x^{\alpha+\beta} = x^\alpha x^\beta = x^\alpha \oplus x^\beta = (\alpha \odot x) \oplus (\beta \odot y)$

7. $\alpha \odot (x \oplus y) = \alpha \odot (x \cdot y) = (x \cdot y)^\alpha = x^\alpha \cdot y^\alpha = x^\alpha \oplus y^\alpha = (\alpha \odot x) \oplus (\alpha \odot y)$

8. $1 \odot x = x^1 = x$

Exemplo 3. Seja $M_{m\times n}(\mathbb{R})$ o conjunto das matrizes reais de ordem $m \times n$.Se $A \in M_{m\times n}(\mathbb{R})$ então A é representada por:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}_{m\times n}$$

Dadas duas matrizes A e $B \in M_{m\times n}(\mathbb{R})$ então a adição e a multiplicação por um escalar são definidas por:

$$A+B=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} := \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

$$\text{e } \alpha A := \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}, \alpha \in \mathbb{R}, \text{ onde } A+B \text{ e } \alpha A \in M_{m\times n}(\mathbb{R}).$$

O conjunto das matrizes com a adição e multiplicação por escalar definidas acima é um espaço vetorial. De fato,

$$1.\ A+B=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} = \begin{bmatrix} b_{11}+a_{11} & b_{12}+a_{12} & \cdots & b_{1n}+a_{1n} \\ b_{21}+a_{21} & b_{22}+a_{22} & \cdots & b_{2n}+a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1}+a_{m1} & b_{m2}+a_{m2} & \cdots & b_{mn}+a_{mn} \end{bmatrix}$$

$$= B+A.$$

$$2.\ (A+B)+C=\left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}\right) + \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} (a_{11}+b_{11})+c_{11} & (a_{12}+b_{12})+c_{12} & \cdots & (a_{1n}+b_{1n})+c_{1n} \\ (a_{21}+b_{21})+c_{21} & (a_{22}+b_{22})+c_{22} & \cdots & (a_{2n}+b_{2n})+c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m1}+b_{m1})+c_{m1} & (a_{m2}+b_{m2})+c_{m2} & \cdots & (a_{mn}+b_{mn})+c_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11}+c_{11} & b_{12}+c_{12} & \cdots & b_{1n}+c_{1n} \\ b_{21}+c_{21} & b_{22}+c_{22} & \cdots & b_{2n}+c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1}+c_{m1} & b_{m2}+c_{m2} & \cdots & b_{mn}+c_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} +$$

$$+ \left(\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} + \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} \right) = A + (B + C).$$

3. Existe $\vec{0} \in M_{m\times n}$ dada por $\vec{0} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n}$, pois

$$A + \vec{0} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} a_{11}+0 & a_{12}+0 & \cdots & a_{1n}+0 \\ a_{21}+0 & a_{22}+0 & \cdots & a_{2n}+0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+0 & a_{m2}+0 & \cdots & a_{mn}+0 \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A.$$

4. Dado $A \in M_{m\times n}(\mathbb{R})$, existe $-A \in M_{m\times n}(\mathbb{R})$ de tal forma que

$$A + (-A) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}-a_{11} & a_{12}-a_{12} & \cdots & a_{1n}-a_{1n} \\ a_{21}-a_{21} & a_{22}-a_{22} & \cdots & a_{2n}-a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}-a_{m1} & a_{m2}-a_{m2} & \cdots & a_{mn}-a_{mn} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \vec{0}$$

5. $(\alpha+\beta)\cdot A = (\alpha+\beta) \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} (\alpha+\beta)a_{11} & \cdots & (\alpha+\beta)a_{1n} \\ \vdots & & \\ (\alpha+\beta)a_{m1} & \cdots & (\alpha+\beta)a_{mn} \end{bmatrix}$

$$= \begin{bmatrix} \alpha a_{11}+\beta a_{11} & \cdots & \alpha a_{1n}+\beta a_{1n} \\ \vdots & & \\ \alpha a_{m1}+\beta a_{m1} & \cdots & \alpha a_{mn}+\beta a_{mn} \end{bmatrix} = \alpha \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \beta \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

$$= \alpha A + \beta A$$

6. $(\alpha\beta)A = (\alpha\beta) \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \alpha \begin{bmatrix} \beta a_{11} & \cdots & \beta a_{1n} \\ \beta a_{21} & \cdots & \beta a_{2n} \\ \vdots & & \\ \beta a_{m1} & \cdots & \beta a_{mn} \end{bmatrix} = \alpha(\beta A)$

7. $\alpha(A+B) = \alpha \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \right)$

$$= \alpha \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} = \begin{bmatrix} \alpha(a_{11}+b_{11}) & \alpha(a_{12}+b_{12}) & \cdots & \alpha(a_{1n}+b_{1n}) \\ \alpha(a_{21}+b_{21}) & \alpha(a_{22}+b_{22}) & \cdots & \alpha(a_{2n}+b_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha(a_{m1}+b_{m1}) & \alpha(a_{m2}+b_{m2}) & \cdots & \alpha(a_{mn}+b_{mn}) \end{bmatrix} =$$

$$= \begin{bmatrix} \alpha a_{11}+\alpha b_{11} & \alpha a_{12}+\alpha b_{12} & \cdots & \alpha a_{1n}+\alpha b_{1n} \\ \alpha a_{21}+\alpha b_{21} & \alpha a_{22}+\alpha b_{22} & \cdots & \alpha a_{2n}+\alpha b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1}+\alpha b_{m1} & \alpha a_{m2}+\alpha b_{m2} & \cdots & \alpha a_{mn}+\alpha b_{mn} \end{bmatrix} =$$

$$= \alpha \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \alpha \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \alpha A + \alpha B.$$

8. $(1 \cdot A) = 1 \cdot \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} 1 \cdot a_{11} & 1 \cdot a_{12} & \cdots & 1 \cdot a_{1n} \\ 1 \cdot a_{21} & 1 \cdot a_{22} & \cdots & 1 \cdot a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 \cdot a_{m1} & 1 \cdot a_{m2} & \cdots & 1 \cdot a_{mn} \end{bmatrix} =$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A.$$

Exemplo 4. Seja $P_n = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \, a_i \in \mathbb{R}\}$ o conjunto dos polinômios de grau menor ou igual a n. P_n é um espaço vetorial em relação as operações usuais definidas da seguinte forma, se $P_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ e $Q_n(x) = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n$, então

$$P_n(x) + Q_n(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n, \text{ e}$$

$$\alpha P_n(x) = \alpha a_0 + \alpha a_1 x + +\alpha a_2 x^2 + \cdots + \alpha a_n x^n.$$

As oito propriedades são facilmente verificadas observando que o polinômio nulo $0 = 0 + 0\alpha + \cdots + 0\alpha^n$ é o elemento neutro da adição e dado $p_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, o oposto aditivo é dado por $-p_n(x) = -a_0 - a_1x - \cdots - a_nx^n$.

2.1.2 Subespaços

Definição 2. Um subconjunto W, não vazio, de um espaço vetorial V é um subespaço vetorial de V se estiverem satisfeitas as seguintes condições.

1. $0 \in W$;
2. se $u, v \in W$, então $u + v \in W$;
3. se $\alpha \in \mathbb{R}$ e $u \in W$, então $\alpha u \in W$.

Todo espaço vetorial V admite pelo menos dois subespaços vetoriais: o subespaço nulo $\{0\}$ e o próprio espaço vetorial V. Estes subespaços são chamados subespaços triviais. Os demais subespaços, se existirem, são chamados subespaços próprios.

Seja V um espaço vetorial e W um subconjunto não vazio de V. É facil ver que, W é um subespaço vetorial de V, se $u + \alpha v \in W$, para todo $\alpha \in \mathbb{R}$ e para todo u, $v \in W$.

Exemplo 5. Sejam $V = \mathbb{R}^2$ e $W = \{(x,y) \in \mathbb{R}^2; y = 2x\}$. W é um subespaço vetorial de V. De fato, se $u \in W$ então $u = (x_1, 2x_1)$ para algum $x_1 \in \mathbb{R}$ e se $v \in W$ então $v = (x_2, 2x_2)$, para algum $x_2 \in \mathbb{R}$. Logo $u + v = (x_1, 2x_1) + (x_2, 2x_2) = (x_1 + x_2, 2x_1 + 2x_2) = (x_1 + x_2, 2(x_1 + x_2))$ e podemos ver que $u + v \in W$, pois a segunda componente é o dobro da primeira. Dado $\alpha \in \mathbb{R}$, temos $\alpha u = \alpha(x_1, 2x_1) = (\alpha x_1, 2(\alpha x_1))$ que também pertence a W. Além disso, $\vec{0} \in W$ pois $(0,0) = (0, 2 \cdot 0)$. Portanto W é um subespaço vetorial.

Exemplo 6. Seja $V = M_{2\times 3}(\mathbb{R})$ o espaço vetorial dado por:

$$V = \{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}, a, b, c, d, e, f \in \mathbb{R}\}$$

e seja $W \subset V$, dada por:

$$W = \{ \begin{bmatrix} a & b & 0 \\ 0 & e & f \end{bmatrix}, a, b, e, f \in \mathbb{R}\}$$

W é um subespaço vetorial de V. De fato, W é não vazio, pois $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in W$.

Se as matrizes $A, B \in W$, então

$$A + B = \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & e_1 & f_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 & 0 \\ 0 & e_2 & f_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 & 0 \\ 0 & e_1 + e_2 & f_1 + f_2 \end{bmatrix} \in W$$

e

$$\alpha A = \alpha \begin{bmatrix} a_1 & b_1 & 0 \\ 0 & e_1 & f1 \end{bmatrix} = \begin{bmatrix} \alpha a_1 & \alpha b_1 & 0 \\ 0 & \alpha e_1 & \alpha f1 \end{bmatrix} \in W$$

Portanto W é subespaço vetorial do espaço V.

Exemplo 7. Considere o sistema linear homogêneo: $\begin{cases} a_{11}x + a_{12}y + a_{13}z = 0 \\ a_{21}x + a_{22}y + a_{23}z = 0 \\ a_{31}x + a_{32}y + a_{33}z = 0 \end{cases}$.

Denotando por

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, 0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o sistema linear homogêneo pode ser escrito como $AX = 0$.

Seja: $W = \{X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, x, y, z \in \mathbb{R} / AX = 0\}$ o conjunto de todas as soluções homogêneas do sistema dado. Vamos verificar que W é um subespaço vetorial do $\mathbb{R}^3$. Podemos ver que $W \neq \emptyset$, pois $A0 = 0$, e assim $0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in W$. Também se $X_1, X_2 \in W$ então $X_1 + X_2 \in W$ e $\forall \alpha \in \mathbb{R}$ temos $\alpha X_1 \in W$.De fato,

$$\begin{cases} X_1 \in W \Rightarrow AX_1 = 0 \\ X_2 \in W \Rightarrow AX_2 = 0 \end{cases} \longrightarrow AX_1 + AX_2 = A(X_1 + X_2) = 0 \text{ e } A(\alpha X_1) = \alpha AX_1 = \alpha \cdot 0 = 0.$$

Portanto W é subespaço vetorial de V.

Como os espaços vetoriais são conjuntos é natural perguntar se a união e a interseção de conjuntos preservam a propriedade de espaço vetorial.

1 A interseção de dois subespaços de um espaço vetorial V é um subespaço de V.

Demonstração. Sejam U e V subespaços de V. Vamos verificar se $U \cap W$ é também um subespaço de V. $U \cap W$ é um subconjunto não vazio de V, pois $0 \in U$ e $0 \in W$, já que ambos são subespaços de V. Agora, $\alpha \in \mathbb{R}$ e $u, v \in U \cap W$, como $u, v \in U$ e $u, v \in W$, segue-se que $u + \alpha v \in U$ e $u + \alpha v \in W$, ou seja $u + \alpha v \in U \cap W$. Assim $U \cap W$ é um subespaço de V. □

2 A união de dois subespaços de um espaço vetorial V não é necessariamente um subespaço de V. Como exemplo, podemos considerar $U = \{(x,y) \in \mathbb{R}^2; x + y = 0\}$ e $W = \{(x,y) \in \mathbb{R}^2; x - y = 0\}$, subespaços de $\mathbb{R}^2$, o conjunto $U \cup W$ não é subespaço de $\mathbb{R}^2$, pois $u = (1,1) \in U \cup W$ e $w = (1,-1) \in U \cup W$, mas $u + w = (2,0) \notin U \cup W$.

Definição 3. Sejam S_1 e S_2 subespaços de V. Definimos $S = S_1 + S_2$ como sendo o conjunto de todos os vetores $u + v$ tal que $u \in S_1$ e $v \in S_2$.

3 A soma S de dois subespaços vetoriais S_1 e S_2 de V é um subespaço vetorial.

Demonstração. Sejam u e $w \in S$. Então existem $w_1, v_1 \in S_1$ e $u_2, v_2 \in S_2$ tais que $u = u_1 + u_2$ e $v = v_1 + v_2$. Logo, $u + \alpha v = u_1 + u_2 + \alpha(v_1 + \alpha_2) = (u_1 + \alpha v_1) + (u_2 + \alpha v_2)$ e como S_1 e S_2 são subespaços de $V, u_1 + \alpha v_1 \in S_1$ e $u_2 + \alpha v_2 \in S_2$. Assim, $u + \alpha v \in S$. □

Definição 4. Sejam S_1 e S_2 dois subespaços vetorial V. W é a soma direta de S_1 e S_2 e representa-se por $W = S_1 \oplus S_2$ se, $V = S_1 + S_2$ e $S_1 \cap S_2 = \{0\}$.

Exemplo 8. Sejam $S_1 = \{(a,0,c,0); a, c \in \mathbb{R}\}$ e $S_2 = \{(0,b,0,d); b, d \in \mathbb{R}\}$ subespaços vetoriais de $\mathbb{R}^4$. Então:

$S_1 + S_2 = \{(a,b,c,d); a,b,c,d \in \mathbb{R}\} = \mathbb{R}^4$ e

$S_1 \cap S_2 = \{(0,0,0,0)\}$

Logo, $S_1 \oplus S_2 = \mathbb{R}^4$ e dizemos que $\mathbb{R}^4$ pode ser escrito como soma direta dos subespaços S_1 e S_2.

Definição 5. Seja V um espaço vetorial e sejam $v_1, v_2, \cdots, v_r$ vetores de V. Diremos que um vetor v de V é uma combinação linear de $v_1, v_2, \cdots, v_r$ se existirem números reais $a_1, a_2, \cdots, a_r$ tais que:

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_r v_r. \tag{2.1}$$

Exemplo 9. Considere os seguintes vetores do $\mathbb{R}^3, v_1 = (1,2,1), v_2 = (1,0,2)$ e $v_3 = (1,1,0)$. Podemos escrever $v = (1,2,4)$ como uma combinação linear de v_1, v_2 e v_3. De fato,

$$(1,2,4) = a_1(1,2,1) + a_2(1,0,2) + a_3(1,1,0),$$

que é equivalente ao sistema $\begin{cases} a_1 + a_2 + a_3 = 1 \\ 2a_1 + a_3 = 2 \\ a_1 + 2a_2 = 4 \end{cases}$. Concluímos que $a_1 = 2, a_2 = 1$ e $a_3 = 2$.

Assim, v é combinação linear de v_1, v_2 e v_3.

Sejam $v_1, v_2, \cdots, v_r$ vetores de um espaço vetorial V. Denotaremos por $G(v_1, v_2, \cdots, v_r)$, o conjunto de todas as combinações lineares de $v_1, v_2, \cdots, v_r$ em V.O seguinte resultado é de facil verificação.

Proposição 1. Seja $W = G(v_1, v_2, \cdots, v_r)$ onde $v_1, v_2, \cdots, v_r$ são vetores de um espaço vetorial V. Valem as seguintes afirmações:

1. W é um subespaço de V.
2. W é o menor subespaço de V contendo $v_1, v_2, \cdots, v_r$ ou seja qualquer subespaço de V que contém $v_1, v_2, \cdots, v_r$ também contém W.

Exemplo 10. Seja $V = \mathbb{R}^3$. Vamos determinar o subespaço gerado pelos vetores $v_1 = (1,1,1); v_2 = (1,1,0)$ e $v_3 = (1,0,0)$. Considere $v = (x,y,z)$ e suponha que $v = a_1v_1 + a_2v_2 + a_3v_3$ para algum $a_1, a_2, a_3 \in \mathbb{R}$. Assim, $(x,y,z) = a_1(1,1,1) + a_2(1,1,0) + a_3(1,0,0)$ que equivale ao sistema

$$\begin{cases} a_1 + a_2 + a_3 = x \\ a_1 + a_2 = y \\ a_1 = z \end{cases}$$

onde concluímos que $a_1 = z, a_2 = y - z$ e $a_3 = x - y$. Logo, $(x,y,z) = z(1,1,1) + (y-z)(1,1,0) + (x-y)(1,0,0)$ e concluímos que $\mathbb{R}^3 \subset G(v_1, v_2, v_3)$. Como, $G(v_1, v_2, v_3)$ é subespaço vetorial de $\mathbb{R}^3$, temos $G(v_1, v_2, v_3) = \mathbb{R}^3$.

Proposição 2. Sejam $\alpha = \{v_1, v_2, \cdots, v_r\}$ e $\beta = \{w_1, w_2, \cdots, w_m\}$ dois conjuntos de vetores em um espaço vetorial V. As seguintes afirmações são equivalentes.

(a) $G(v_1, v_2, \cdots, v_r) = G(w_1, w_2, \cdots, w_m)$;

(b) Cada vetor de α é uma combinação linear de vetores de β e cada vetor de β é uma combinação linear de vetores de α.

2.1.3 Dependência e Independência Linear

Definição 6. Sejam $v_1, v_2, \cdots, v_r$ vetores em um espaço vetorial V. Dizemos que os vetores $v_1, v_2, \cdots, v_r$ são linearmente independentes ou simplesmente l.i. se a equação

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r = 0 \tag{2.2}$$

é satisfeita somente quando $\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$. Caso exista algum $\alpha_i \neq 0$ dizemos que os vetores $v_1, v_2, \cdots, v_r$ são linearmente dependentes ou simplesmente l. d..
O conjunto $\{v_1, v_2, \cdots, v_r\}$ é dito ser independente ou dependente se os vetores $v_1, v_2, \cdots, v_r$ são linearmente independentes ou dependentes, respectivamente. Dizemos que 2.2 é uma combinação linear nula de $v_1, \cdots, v_r$.

Exemplo 11. Sejam V um espaço vetorial e u, v e $w \in V$. Suponha que w é combinação linear de u e v. Então u, v e w são linearmente dependentes. De fato, como w é combinação linear de u, v existem α e $\beta \in \mathbb{R}$ tais que $\alpha u + \beta v = w$. Assim $\alpha u + \beta v - w = 0$.

Exemplo 12. Considere o $\mathbb{R}^4$ com as operações usuais. Temos que $(0,1,0,1),(4,6,2,6)$ e $(2,0,1,0)$ são linearmente dependentes, pois $3(0,1,0,1) - \frac{1}{2}(4,6,2,6) + 1(2,0,1,0) = 0$

Exemplo 13. Conside M_2 com as operações usuais. Então $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 2 & 2 \end{bmatrix}$ são linearmente independentes. De fato, sejam $\alpha, \beta, \gamma \in \mathbb{R}$ tais que:

$$\alpha A + \beta B + \gamma C = 0.$$

Então, $\begin{cases} \alpha + \beta = 0 \\ \alpha + \beta = 0 \\ 2\gamma = 0 \\ \alpha + 2\gamma = 0 \end{cases}$, de onde concluímos que $\alpha = \beta = \gamma = 0$.

Proposição 3. Sejam $v_1, v_2, \cdots, v_n$ vetores em $\mathbb{R}^n$, onde, para cada i, com $1 \leq i \leq r$, temos $v_i = (a_{i1}, a_{i2}, \cdots, a_{in})$ e considere $A = [a_{ij}]$. Temos que $\{v_1, v_2, \cdots, v_n\}$ é linearmente independente se, e somente se, A é invertível.

Demonstração. Basta observar que o sistema linear $k_1 v_1 + \cdots + k_n v_n = 0$ terá apenas a solução nula. □

Proposição 4. Sejam $v_1, v_2, \cdots, v_r$ vetores em $\mathbb{R}^n$. Se $r > n$, então os vetores $v_1, v_2, \cdots, v_r$ são linearmente dependentes.

Demonstração. Suponha que para cada $1 \leq i \leq r$, $v_i = (a_{i1}, \cdots, a_{in})$. Consideremos a equação:

$$k_1 v_1 + k_2 v_2 + \cdots + k_r v_r = 0 \tag{2.3}$$

que é equivalente ao sistema:

$$\begin{cases} a_{11}k_1 + a_{21}k_2 + \cdots + a_{r1}k_r = 0 \\ a_{12}k_1 + a_{22}k_2 + \cdots + a_{r2}k_r = 0 \\ \cdots \\ a_{1n}k_1 + a_{2n}k_2 + \cdots + a_{rn}k_r = 0 \end{cases}$$

O sistema dado é linearmente homogênio de n equações e r incógnitas $k_1, k_2, \cdots, k_r$. Como $r > n$ segue que o sistema tem soluções não triviais. Isto mostra que $v_1, v_2, \cdots, v_r$ são linearmente dependentes. □

Proposição 5. Um conjunto finito $\alpha = \{v_1, \cdots, v_r\}$, com $r \geq 2$, de um espaço vetorial V é linearmente dependente se, e somente se, pelo menos um dos vetores de α pode ser escrito como combinação linear dos outros vetores.

Definição 7. Seja $\alpha = \{v_1, v_2, \cdots, v_r\}$ um conjunto **ordenado** de vetores em um espaço vetorial não nulo V. Dizemos que α é uma base de V se as seguintes condições são verificadas:

i) α é linearmente independente;

ii) $V = G(\alpha)$.

Exemplo 14. Considere $\mathbb{R}^4$ com as operações usuais. Temos que

$$B = \{(1,0,1,0),(0,1,0,1),(1,0,0,1),(0,0,1,1)\}$$

é uma base do $\mathbb{R}^4$. De fato, seja $(a,b,c,d) \in \mathbb{R}^4$ e considere $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ tais que:

$$\alpha(1,0,1,0) + \beta(0,1,0,1) + \gamma(1,0,0,1) + \delta(0,0,1,1) = (a,b,c,d).$$

Temos que

$$\begin{cases} \alpha + \gamma = a \\ \beta = b \\ \alpha + \delta = c \\ \beta + \gamma + \delta = d \end{cases}$$

e assim, $\alpha = a - d + b + \dfrac{c-a+d-b}{2}$, $\beta = b$, $\gamma = d - b - \dfrac{c-a+d-b}{2}$, $\delta = \dfrac{c-a+d-b}{2}$. Assim temos que $G(B) = \mathbb{R}^4$. Para verificar que B é l.i., basta considerar $a = b = c = d = 0$ e verificar que $\alpha = \beta = \gamma = \delta = 0$.

Exemplo 15. Considere o espaço vetorial M_{23} de todas as matrizes 2×3 sobre o corpo dos números reais. Então o conjunto B forma uma base de M_{23}, onde

$$B = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}.$$

A demonstrção deste fato é similar ao exempolo anterior.

De modo mais geral, no espaço vetorial M_{rs} das matrizes $r \times s$, seja E_{ij} a matriz com 1 como elemento de ordem ij, e o 0 nos demais. Então, o conjunto de todas as matrizes E_{ij} formam uma base de M_{rs} chamada base usual de M_{rs}. Em particular $e_1 = (1,0,0,\cdots,0), e_2 = (0,1,0,\cdots,0), e_n = (0,0,0,\cdots,1)$ formam a base usual de $\mathbb{R}^n$.

Uma base de um espaço vetorial V é um conjunto gerador no qual cada vetor de V pode ser escrito de modo único como combinação linear desses vetores. Esse é o resultado da próxima proposição.

Proposição 6. Sejam $\alpha = \{v_1, v_2, \cdots, v_n\}$ um conjunto ordenado de vetores de um espaço vetorial não nulo. As seguintes afirmações são equivalentes.

i) α é uma base de V;

ii) cada vetor de v em V pode ser escrito de modo único na forma $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$.

Demonstração. Suponhamos que α é uma base de V. Tomemos $v \in V$. Como α gera V, existem números reais $a_1, a_2, \cdots, a_n$ tais que:

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n. \tag{2.4}$$

Para mostrar que a combinação linear em 2.4 é única, suponhamos que existem $b_1, b_2, \cdots, b_n$ tais que:

$$v = b_1v_1 + b_2v_2 + \cdots + b_nv_n \tag{2.5}$$

De 2.4 e 2.5 segue que:

$$(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n = 0 \tag{2.6}$$

Como α é linearmente independente, a equação 2.6 é satisfeita somente se $(a_1 - b_1) = (a_2 - b_2) = \cdots = (a_n - b_n) = 0$. Como $v \in V$ foi tomado de modo arbitrário, segue o resultado.
Suponhamos, agora, que α tem a propriedade de que cada vetor v em V pode ser escrito de modo único como combinação linear dos espaço gerado, claramente α gera V e para mostrar que α é linearmente independente, considere a equação $\mathrm{k}_1v_1 + k_2v_2 + \cdots + k_nv_n = 0$. Como $0 = 0v_1 + 0v_2 + \cdots + 0v_n$ e esta forma de escrever o vetor nulo é única segue que $k_1 = k_2 = \cdots = k_n = 0$. □

Definição 8. Seja V um espaço vetorial gerado por um conjunto finito de vetores não nulos $\alpha = \{v_1, v_2, \cdots, v_n\}$, isto é, $G(\alpha) = V$. Diremos que V é finitamente gerado.

Proposição 7. Seja V um espaço vetorial finitamente gerado e seja $B = \{v_1, v_2, \cdots, v_n\}$ uma base de $V \neq \{0\}$. Então, qualquer conjunto com mais de n vetores de V é linearmente dependente.

Demonstração. De fato, seja $B' = \{w_1, w_2, \cdots, w_m\}$ um conjunto de m vetores de V, com $m > n$. Queremos mostrar que B' é LD. Sejam $x_1, x_2, \cdots, x_m$ tais que $x_1w_1 + x_2w_2 + \cdots + x_mw_m = 0$.

Como B é uma base de V e $w_i \in B' \subset V$, $i = 1, \ldots, m$, então, existem escalares $\alpha_j, \beta_j, \cdots, \eta_j$, com $j = 1, \ldots, n$, tais que:

$$\begin{cases} w_1 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n \\ w_2 = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n \\ \vdots \\ w_m = \eta_1 v_1 + \eta_2 v_2 + \cdots + \eta_n v_n. \end{cases} \tag{2.7}$$

Desta forma,

$(\alpha_1 x_1 + \beta_1 x_2 + \cdots + \eta_1 x_m) v_1 +$

$(\alpha_2 x_1 + \beta_2 x_2 + \cdots + \eta_2 x_m) v_2 + \cdots (\alpha_n x_1 + \beta_n x_2 + \cdots + \eta_n x_m) v_n = 0.$

Como os vetores $v_1, v_2, \cdots, v_n$ são uma base para V então eles são l.i. Logo,

$$\begin{cases} \alpha_1 x_1 + \beta_1 x_2 + \cdots + \eta_1 x_m = 0 \\ \alpha_2 x_1 + \beta_2 x_2 + \cdots + \eta_2 x_m = 0 \\ \vdots \\ \alpha_n x_1 + \beta_n x_2 + \cdots + \eta_n x_m = 0 \end{cases}.$$

Sendo $m > n$, o sistema admite mais de uma solução, além da trivial, logo $B' = \{w_1, w_2, \cdots, w_m\}$ é l.d. □

Proposição 8. Duas bases quaisquer de um espaço vetorial $V \neq \{0\}$, V finitamente gerado, têm o mesmo número de vetores.

Demonstração. De fato, sejam $A = \{v_1, v_2, \cdots, v_n\}$ e $B = \{w_1, w_2, \cdots, w_m\}$ duas bases para V. Como A é base e B é l.i., então $m \leq n$. Por outro lado, como B é base e A é l.i., então $n \leq m$. Logo, $m = n$. □

Definição 9. Seja V um espaço vetorial finitamente gerado. Se $V = \{0\}$ definimos a dimensão de V como sendo 0. Se $V \neq \{0\}$ e finitamente gerado, definimos a dimensão de V como sendo o número de elementos de uma base qualquer de V. Usaremos o simbolo $\dim V$ para designar a dimensão de V. Se um espaço vetorial V não é finitamente gerado dizemos que V possui dimensão infinita.

Como exemplos, podemos citar, $\dim \mathbb{R} = 1$, $\dim \mathbb{R}^n = n$, $\dim M_{2\times 2} = 4$, $\dim M_{m\times n} = m \times n$ e $\dim P_n = n + 1$.

Teorema 1. Seja V um espaço vetorial de dimensão n. Qualquer conjunto de vetores L.I. em V é parte de uma base, isto é, pode ser completado até formar uma base de V.

Demonstração. Suponha que esse conjunto contenha r vetores linearmente independentes. Como $r < n$ existe $u_{r+1} \in V$ tal que $u_1, \cdots, u_r, u_{r+1}$ são l.i., pois caso contrário os vetores $u_1, \cdots, u_r$ formariam uma base de V; o que é impossível pois $\dim V = n > r$. Se $r+1 = n$ então $u_1, \cdots, u_r, u_{r+1}$ formam uma base de V. Se $r+1 < n$ então é possível encontrar $u_{r+2} \in V$ tal que $u_1, \cdots, u_r, u_{r+1}, u_{r+2}$ são l.i., pois caso contrário a sequência $u_1, \cdots, u_r, u_{r+1}$ seria uma base de V; o que é impossível pois $\dim V = n > r+1$. Repetindo os argumentos acima, encontramos vetores $u_{r+1}, u_{r+2}, \cdots, u_{r+k}$, onde $r+k = n$, de forma que $u_1, \cdots, u_r, u_{r+1}, \cdots, u_{r+k}$ são l.i. e, como $\dim V = n = r+k$ segue que esta sequência de vetores é uma base de V que contém os vetores $u_1, \cdots, u_r$. □

Exemplo 16. Sejam $v_1 = (1,0,2)$ e $v_2 = (0,-1,3)$, podemos completar o conjunto $\{v_1, v_2\}$ de modo a formar uma base do $\mathbb{R}^3$. De fato, sabemos que $\dim \mathbb{R}^3 = 3$, logo devemos acrescentar um vetor $v_3 = (a,b,c) \neq \alpha_1 v_1 + \alpha_2 v_2 = (\alpha_1, -\alpha_2, 2\alpha_1 + 3\alpha_2)$, pois v_3 nao pode ser uma combinação linear de v_1 e v_2. Existem infinitos vetores possíveis, por exemplo, podemos escolher $v_3 = (2,-1,0)$. Assim, $\{v_1, v_2, v_3\}$ é uma base do $\mathbb{R}^3$.

Proposição 9. Seja V um espaço vetorial de dimensão finita. Se W é um subespaço de V, então W tem também dimensão finita e $\dim W \leq \dim V$. Além disso, se $\dim W = \dim V$, então $W = V$.

Como exemplo, suponha $V = \mathbb{R}^3$. Então, se $S \subset \mathbb{R}^3$, temos $\dim S = 0, 1, 2$ ou 3 e

(1) $\dim S = 0 \Rightarrow S = \{0\}$

(2) $\dim S = 1 \Rightarrow S$ é uma reta, passando pela origem.

(3) $\dim S = 2 \Rightarrow S$ é um plano, passando pela origem.

(4) $\dim S = 3 \Rightarrow S = \mathbb{R}^3$.

Definição 10. Sejam V um espaço vetorial finitamente gerado e B uma base de V formada pelos vetores $u_1, \cdots, u_n$. Como B é uma base de V, todo elemento de $u \in V$ se escreve como $\alpha_1 u_1 + \cdots + \alpha_n u_n$, com os coeficientes $\alpha_1, \cdots, a_n \in \mathbb{R}$. Pela proposição 6, os coeficientes $\alpha_1, \cdots, a_n$ são unicamente determinados pelo vetor u. Estes coeficientes são denominados coordenas de u com relação à base B. Representaremos as coordenadas de u com relação à base B como

$$[u]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}_B$$

ou simplesmente por

$$[u]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

quando B estiver subentendida.

Exemplo 17. Vamos mostrar que os vetores $(1,1,1),(0,1,1)$ e $(0,0,1)$ formam uma base de $\mathbb{R}^3$ e encontrar as coordenadas de $u=(1,2,0)\in\mathbb{R}^3$ com relação à essa base. Já sabemos que $dim\mathbb{R}^3=3$. Para verificar se os vetores acima formam uma base de V, basta verificar se eles são l.i.. Vemos que estes vetores são de fato l.i. pois a matriz

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

possui determinante igual a $1\neq 0$. Agora,

$$(1,2,0)=\alpha(1,1,1)+\beta(0,1,1)+\gamma(0,0,1)=(\alpha,\alpha+\beta,\alpha+\beta+\gamma)$$

que é equivalente ao sistema

$$\begin{cases} \alpha=1 \\ \alpha+\beta=2 \\ \alpha+\beta+\gamma=0 \end{cases}$$

cuja única solução é $\alpha=1,\beta=1$ e $\gamma=-2$. Desse modo, as coordenadas de $u=(1,2,0)$ com relação à base B são dadas por

$$[u]_B=\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}_B$$

Definição 11. Seja V um espaço vetorial e U um subespaço vetorial de V. O complemento ortogonal de U é o conjunto

$$U^{\perp}=\{v\in V; <u,v>=0,\ \forall u\in U\}.$$

Proposição 10. $U^{\perp}$ é um subespaço vetorial de V.

Demonstração. Temos $0\in U^{\perp}$ pois $<0,u>=0\ \forall u\in U$. Se $v,w\in U^{\perp}$ e $\alpha\in\mathbb{R}$ então para todo $u\in U$, temos:

$$<v+\alpha w,u>=<v,u>+\alpha<w,u>=0$$

Portanto, $v+\alpha w\in U^{\perp}$. □

Observação 1. Se V têm dimensão finita então $u\in U^{\perp}$ se, e somente se, u é ortogonal a todos os vetores de uma base qualquer de U.

Exemplo 18. Encontre $U^{\perp}$ se $U=\left\{(x,y,z)\in\mathbb{R}^3/x-y-z=0\right\}$

Temos $(x,y,z)\in U$ se $(x,y,z)=(y+z,y,z)=y(1,1,0)+z(1,0,1)$.

Logo $(1,1,0)$ e $(1,0,1)$ formam uma base para U.

Assim, $(x,y,z) \in U^{\perp}$ se $<(x,y,z),(1,1,0)>= 0$ e $<(x,y,z),(1,0,1)>= 0$

$$\begin{cases} x+y=0 \\ x+z=0 \end{cases} \iff (x,y,z) = x(1,-1,-1)$$

Assim:

$$U^{\perp} = [(1,-1,-1)].$$

Temos o seguinte resultado para complemento ortogonal:

Teorema 2. Se W é um subespaço de V,então

$$V = W \oplus W^{\perp}.$$

CAPÍTULO

3

TRANSFORMAÇÕES LINEARES E MATRIZES

3.1 Transformações Lineares

No capítulo anterior, estudamos espaços vetoriais e as propriedades dos vetores, elementos destes espaços, como, por exemplo, os conceitos de dependência linear e independência linear. Neste capítulo, vamos estudar as transformações entre espaços vetoriais, mas não qualquer uma, vamos estudar as transformações que possuem a propriedade de associar à soma de dois vetores, a soma dos resultados obtidos pela aplicação da transformação em cada um desses vetores e, associar ao produto de um escalar por um vetor, o produto deste escalar pelo resultado obtido pela transformação aplicada a esse vetor. Descrevemos de forma mais precisa essas condições na seguinte definição:

Definição 12. Sejam U e V espaçoes vetoriais. Dizemos que uma função $T : U \to V$é uma transformação linear se forem erificadas as seguintes condições:
a) $T(u+v) = T(u) + T(v)$, $\forall u, v \in U$;
b) $T(\lambda u) = \lambda T(u)$, $\forall \lambda \in \mathbb{R}$ e $\forall u \in U$.

Quando a transformação linear for de um espaço vetorial V nele mesmo, ela é chamada de operador linear em V.

Os resultados deste capítulo podem ser encontrados em (ZANI,), (LADEIRA, 2004), (HEFEZ A.; SOUZA FERNANDES, 2012) e (FRENSEL K.; DELGADO, 2011).

Observação 2. Toda transformação linear $T : U \to V$ leva o vetor nulo do espaço U no vetor nulo do espaço V. De fato, $T(0_U) = T(0 \cdot 0_U) = 0 \cdot T(0_U) = 0_V$.

Podemos citar como exemplos de transformações lineares:

Exemplo 19. A transformação nula $T : U \to V$ dada por $T(u) = 0$, $\forall u \in U$ e a transformação identidade $T : U \to U$ dada por $T(u) = u$, $\forall u \in U$.

Exemplo 20. $T : \mathscr{C}^1([a,b];\mathbb{R}) \to \mathscr{C}([a,b];\mathbb{R})$, dada por T(f)=f'. Aqui $\mathscr{C}([a,b];\mathbb{R})$ significa o espaço das funções contínuas em $[a,b]$ com valores reais e $\mathscr{C}^1([a,b];\mathbb{R})$ significa o espaço das funções com derivadas contínuas.

Exemplo 21. Seja $A \in M_{m\times n}$ uma matriz fixada e considere $T : M_{n\times 1} \to M_{m\times 1}$ dada por $T(X) = AX.T$ assim definida é uma transformação linear.

Podemos observar que uma transformação linear fica completamente determinada se conhecermos seus valores nos elementos da base do espaço de saída, isto é:

Proposição 11. Sejam U e V espaços vetoriais e seja $\alpha = \{u_1, u_2, \ldots, u_n\}$ uma base para o espaço U. Então, toda transformação linear $T : U \to V$ fica determinada conhecendo-se os valores de $T(u_1)$, $T(u_2)$, $\cdots$, $T(u_n)$.

Demonstração. Seja $u \in U$, então existem escalares $\alpha_1, \alpha_2, \cdots, \alpha_n$ tais que $u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots \alpha_n u_n$. Assim, $T(u) = T(\alpha_1 u_1 + \alpha_2 u_2 + \cdots \alpha_n u_n) = \alpha_1 T(u_1) + \alpha_2 T(u_2) + \cdots \alpha_n T(u_n)$. □

Definição 13. Sejam U e V espaços vetoriais. O conjunto de todas as transformações lineares de U em V é denotado por $\mathscr{L}(U,V)$. Quando $U = V$ denotamos $\mathscr{L}(U,V)$ por $\mathscr{L}(U)$.

Nosso objetivo é identificar o conjunto das transformações lineares $\mathscr{L}(U,V)$, com $dim(U) = n$ e $dim(V) = m$ com o conjunto das matrizes de ordem $m \times n$. Primeiramente, observamos que se definirmos em $\mathscr{L}(U,V)$ as operações de adição e multiplicação por escalar dadas por :

$$T + S : UV,\ (T+S)(u) = T(u) + S(u),\ \forall u \in U,$$

$$\lambda \cdot T : U \to V,\ (\lambda \cdot T)(u) = \lambda T(u),\ \forall \lambda \in \mathbb{R}, \forall u \in U,$$

é um espaço vetorial. A dimensão do espaço vetorial $\mathscr{L}(U,V)$ é $m \cdot n$, se $dim(U) = n$ e $dim(V) = m$.

Definição 14. Uma transformação linear $T : U \to V$ é :
1) injetora se $T(u) = T(v)$ implicar que $u = v$;
2) sobrejetora se para cada $v \in V$ existir um vetor $u \in U$ tal que $T(u) = v$;
3) bijetora se for injetora e sobrejetora.

Definição 15. Dizemos que $T \in \mathscr{L}(U,V)$ possui inversa se existir $S : V \to U$ tal que $S \circ T(u) = u$, $\forall u \in U$ e $T \circ S(v) = v$, $\forall v \in V$.Denotaremos S por T^{-1}.

A fim de que $T \in \mathscr{L}(U,V)$ possua uma função inversa é necessário e suficiente que T seja uma transformação bijetora. Neste caso, a inversa de T, T^{-1} é também uma transformação linear e $T^{-1} \in \mathscr{L}(V,U)$.

Definição 16. Diremos que uma transformação linear $T : U \to V$ é um isomorfismo se ela for uma transformação bijetora. Dizemos, neste caso, que os espaços U e V são isomorfos.

A seguir, listamos alguns resultados para ajudar-nos a provar quando uma transformação linear é uma bijeção.

Quando T é uma transformação linear, temos um resultado bastante prático para verificar se ela é ou não uma transformação injetora, dado pela proposição a seguir:

Proposição 12. Uma transformação linera $T : U \to V$ é injetora se, e somente se, $T(u) = 0$ implicar em $u = 0$.

Demonstração. Se supormos que T é uma transformação linear injetora, então, se $T(u) = 0$, como sabemos que $T(0) = 0$ e T é injetora, segue que $u = 0$. Agora, se considerarmos $T(u) = T(v)$, então $T(u - v) = 0$ e, se por hipótese, isto implicar em $u - v = 0$, então $u = v$ e T é injetora. $\square$

Proposição 13. Seja $T : U \to V$ uma transformação linear. Temos que:
1) Se W é um subespaço vetorial de U, então $T(W) = \{T(w) \in V : w \in W\}$;
2) Se W é um subespaço vetorial de V então $T^{-1}(W)$ é um espaço vetorial de U.

Definição 17. O núcleo de uma transformação linear $T : UV$ é o subespaço vetorial de U dado por $T^{-1}(\{0\})$, ou seja, é o conjunto $\{u \in U : T(u) = 0\}$. O núcleo de uma transformação linear é indicado por $\mathscr{N}(T)$ ou por $Ker(T)$.

O seguinte teorema relaciona a dimensão do núcleo de uma transformação linear com a dimensão de sua imagem.

Teorema 3. Sejam U e V espaços vetoriais, com dimensão de U finita, e $T : U \to V$ uma transformação linear. Então:

$$dim(U) = dim(N(T)) + dim(T(U))$$

Proposição 14. Seja $T : U \to V$ uma transformação linear. T é injetora se, e somente se, $\mathscr{N}(T) = \{0\}$.

Corolário 1. Se U e V são espaços vetoriais de dimensão finita tais que $dim(U) = dim(V)$ e se $T : U \to V$ é uma transformação linear, então as seguintes condições são equivalentes:
1) T é sobrejetora;
2) T é injetora;
3) T é bijetora;

4) T leva bases de U em bases de V.

Proposição 15. Se $T : U \to V$ é um isomorfismo e U tem dimensão finita, então V também tem dimensão finita e $dim(U) = dim(V)$. Por outro lado, se V tem dimensão finita , U também terá dimensão finita e também teremos $dim(U) = dim(V)$.

Proposição 16. Se U e V são espaços vetoriais de dimensão finita n, existe um isomorfismo $T : U \to V$, isto é, os espaços são isomorfos.

Corolário 2. Dois espaços vetoriais de dimensão finita são isomorfos se, e somente se, têm a mesma dimensão.

Corolário 3. Se U é um espaço vetorial de dimensão finita n e V é um espaço vetorial de dimensão finita m, então $\mathscr{L}(U,V)$ é isomorfo ao espaço vetorial $M_{m\times n}$.

3.2 Matriz de uma Transformação Linear

Se V e W são espaço vetoriais de dimensão finita, com bases fixadas, então uma transformação linear $T : V \to W$ pode ser representada por uma matriz. A vantagem de tal representação é que muitos problemas associados às tranformações lineares entre espaços de dimensão finita podem ser resolvidos com a teoria das matrizes.

Seja $T : V \to W$ uma transformação linear, em que $\dim V = n$ e $\dim W = m$.

Sejam $\alpha = \{v_1, v_2, \cdots, v_n\}$ e $\beta = \{w_1, w_2, \cdots, w_m\}$ bases de V e W respectivamente. Como β é uma base de W, podemos determinar de modo único números reais a_{ij}, com $1 \leq i \leq n, 1 \leq j \leq m$, tais que:

$$T(v_i) = a_{1i}w_1 + \cdots + a_{ji}wj + \cdots + a_{mi}w_m.$$

Tomemos agora v em V. Temos que $v = k_1v_1 + \cdots + k_nv_n$, em que $k_i \in \mathbb{R}$ para $1 \leq i \leq n$. Então,

$$T(v) = k_1T(v_1) + \cdots + k_nT(v_n)$$

$$= k_1(a_{11}w_1 + \cdots + a_{m1}w_m) + \cdots + k_n(a_{1n}w_1 + \cdots + a_{mn}w_m) =$$

$$= (a_{11}k_1 + \cdots + a_{1n}k_n)w_1 + \cdots + (a_{m1}k_1 + \cdots + a_{mn}k_n)w_m.$$

Logo,

$$[T(v)]_\beta = \begin{bmatrix} a_{11}k_1 + \cdots + a_{1n}k_n \\ \vdots \\ a_{m1}k_1 + \cdots + a_{mn}k_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ \vdots \\ k_n \end{bmatrix} = [T]^\alpha_\beta \cdot [v]_\alpha,$$ definindo por

$[T]^{\alpha}_{\beta} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$ como sendo a matriz que representa T em relação às bases α e β. Assim, temos,

$$[T(v)]_{\beta} = [T]^{\alpha}_{\beta} \cdot [v]_{\alpha}. \tag{3.1}$$

Exemplo 22. Sejam $\alpha = \{(1,1),(0,2)\}$ e $\beta = \{(1,0,1),(0,1,0),(1,2,0)\}$, bases de $\mathbb{R}^2$ e $\mathbb{R}^3$, respectivamente. Calculemos $[T]^{\alpha}_{\beta}$ onde $T: \mathbb{R}^2 \to \mathbb{R}^3$ é dada por:

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

Como T é uma transformação linear de $\mathbb{R}^2$ em $\mathbb{R}^3$, $[T]^{\alpha}_{\beta}$ é uma matriz 3×2, digamos

$[T]^{\alpha}_{\beta} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$, onde a_{11}, a_{21}, a_{31} são as coordenadas de $T(1,1)$ na base β e a_{12}, a_{22} e a_{32} são as coordenadas de $T(0,2)$ na base β.

$T(1,1) = a_{11}(1,0,1) + a_{21}(0,1,0) + a_{31}(1,2,0) = (2,0,2)$ e

$T(0,2) = a_{12}(1,0,1) + a_{22}(0,1,0) + a_{32}(1,2,0) = (0,-2,4)$.

Assim, $\begin{cases} a_{11} + a_{31} = 2 \\ a_{21} + 2a_{31} = 0 \\ a_{11} = 2 \end{cases}$ e, resolvendo o sistema temos:

$$[T]^{\alpha}_{\beta} = \begin{bmatrix} 2 & 4 \\ 0 & 6 \\ 0 & -4 \end{bmatrix}.$$

Exemplo 23. Vamos agora fazer o inverso, isto é, dada a matriz $[T]^{\alpha}_{\beta}$ vamos determinar qual é a transformação T. Sejam α e β as bases dadas no exemplo anterior $\alpha = \{(1,1),(0,2)\}$ e $\beta = \{(1,0,1),(0,1,0),(1,2,0)\}$. Queremos determinar a transformação linear $T: \mathbb{R}^2 \to \mathbb{R}^3$ tal que $[T]^{\alpha}_{\beta} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}$.

Para determinar T, usaremos a expressão 3.1. Acharemos inicialmente $[v]_{\alpha}$. Ora, se $(x,y) \in \mathbb{R}^2$, então $(x,y) = x(1,1) + \left(\dfrac{y-x}{2}\right)(0,2)$

assim,

$$[(x,y)]_{\alpha} = \begin{bmatrix} x \\ \dfrac{y-x}{2} \end{bmatrix}.$$

Portanto,

$$[T(x,y)]_\beta = \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ \dfrac{y-x}{2} \end{bmatrix} = \begin{bmatrix} x \\ y \\ \dfrac{y-x}{2} \end{bmatrix}.$$

Consequentemente,

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

Proposição 17. Temos as seguintes operações de transformações lineares representadas por matrizes:

1. Sejam T e T' transformações lineares de V em W, onde V e W são espaços vetoriais de dimensão finita. Se α e β são bases de V e W, respectivamente, então:

   $[T+T']^\alpha_\beta = [T]^\alpha_\beta + [T']^\alpha_\beta$,

   $[kT]^\alpha_\beta = k[T]^\alpha_\beta$, onde k é um número real arbitrário.

2. Sejam $T: V \to W$ e $S: W \to U$ transformações lineares, em que V, W e U são espaços vetoriais de dimensão finita. Se α, β e γ são bases de V, W e U, respectivamente, então:

   $[S \circ T]^\alpha_\gamma = [S]^\beta_\gamma \cdot [T]^\alpha_\beta$.

Teorema 4. Seja $T: V \to W$ um isomorfismo, onde V e W são espaços vetoriais de dimensão finita. Se α é uma base de V e β uma base de W, então:

$$[T^{-1}]^\beta_\alpha = ([T]^\alpha_\beta)^{-1} \tag{3.2}$$

Corolário 4. Seja $T: V \to W$ uma transformação linear, onde V e W são espaços vetoriais de mesma dimensão finita. Sejam α e β bases de V e W, respectivamente. Temos que T é invertível se, e somente se, a matriz $[T]^\alpha_\beta$ é invertível.

Exemplo 24. Seja $T: \mathbb{R}^2 \to \mathbb{R}^2$ a transformação linear dada por $T(x,y) = (4x-3y, -2x+2y)$. Vamos verificar que T é invertível e vamos encontrar T^{-1}. Para verificarmos que T é invertível, podemos calcular $[T]^\alpha_\alpha$ onde α é uma base qualquer de $\mathbb{R}^2$, e usar o corolário. Se α é a base canônica de $\mathbb{R}^2$, então, $[T]^\alpha_\alpha = \begin{pmatrix} 4 & -3 \\ -2 & 2 \end{pmatrix}$. Portanto,

$$[T^{-1}]^\alpha_\alpha = ([T]^\alpha_\alpha)^{-1} = \begin{pmatrix} 1 & \frac{3}{2} \\ 1 & 2 \end{pmatrix}$$

Assim,

$$[T^{-1}(x,y)]_\alpha = [T^{-1}]^\alpha_\alpha [(x,y)]_\alpha = \begin{pmatrix} 1 & \frac{3}{2} \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \frac{3}{2}y \\ x + 2y \end{pmatrix}.$$

Logo,

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

3.2.1 Operadores Lineares em $\mathbb{R}^2$ e em $\mathbb{R}^3$

Dentre os operadores lineares mais importantes em $\mathbb{R}^2$ e em $\mathbb{R}^3$ estão os que produzem reflexão, projeções, rotações e homotetias.

• Reflexões:

Podemos considerar o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$, chamado de reflexão em torno do eixo Ox, que transforma cada vetor $v = (x,y) \in \mathbb{R}^2$ em sua imagem simétrica em relação ao eixo Ox.

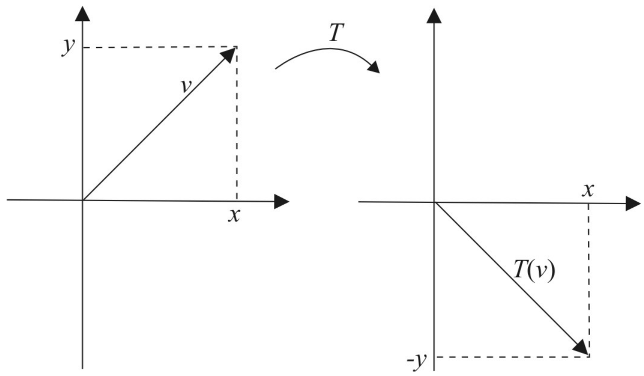

Figura 1 – Reflexão em torno do eixo Ox.

Escrevendo $w = T(v) = (w_1, w_2)$, obteremos as equações $w_1 = x = 1x + 0y$ e $w_2 = -y = 0x - 1y$. Assim se α denota a base canônica de $\mathbb{R}^2$, segue que:

$$[T(v)]_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} [v]_\alpha$$

Tabela 1 – Reflexões mais comuns em $\mathbb{R}^2$

Operador	Equações	Matriz $[T]_\alpha^\alpha$
Reflexão em torno do eixo Oy	$\begin{cases} w_1 = -x \\ w_2 = y \end{cases}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Reflexão em torno da reta $y = x$	$\begin{cases} w_1 = y \\ w_2 = x \end{cases}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Reflexão em torno do plano xOy	$\begin{cases} w_1 = x \\ w_2 = y \\ w_3 = -z \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
Reflexão em torno do plano yOz	$\begin{cases} w_1 = -x \\ w_2 = y \\ w_3 = z \end{cases}$	$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Reflexão em torno do plano xOz	$\begin{cases} w_1 = x \\ w_2 = -y \\ w_3 = z \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

• **Projeções:**

Podemos considerar o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada vetor $v = (x, y) \in \mathbb{R}^2$ em sua projeção ortogonal sobre o eixo Ox.

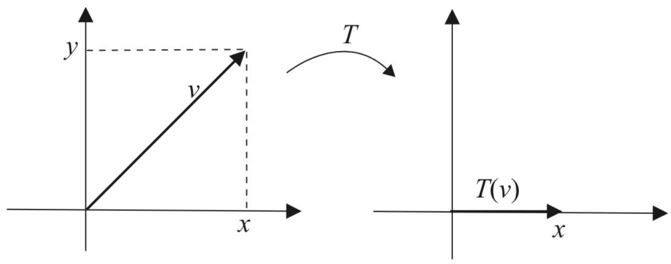

Figura 2 – Projeção

Escrevendo $w = T(v) = (w_1, w_2)$, obteremos as equações $w_1 = x = 1x + 0y$ e $w_2 = 0 = 0x + 0y$.

Assim se α denota a base canônica de $\mathbb{R}^2$, temos:

$$[T(v)]_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} [v]_\alpha$$

Em geral uma projeção ou projeção ortogonal de $\mathbb{R}^2$ ou $\mathbb{R}^3$ é um operador linear que transforma cada vetor em sua projeção ortogonal sobre alguma reta ou algum plano que passa pela origem.

Tabela 2 – Projeções mais comum em $\mathbb{R}^2$

Operador	Equações	Matriz $[T]_\alpha^\alpha$
Projeção sobre o eixo Oy	$\begin{cases} w_1 = 0 \\ w_2 = y \end{cases}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Projeção sobre o plano xOy	$\begin{cases} w_1 = x \\ w_2 = y \\ w_3 = 0 \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
Projeção sobre o plano yOz	$\begin{cases} w_1 = 0 \\ w_2 = y \\ w_3 = z \end{cases}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Projeção sobre o plano xOz	$\begin{cases} w_1 = x \\ w_2 = 0 \\ w_3 = z \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

• **Rotação:**

Considere o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ que rotaciona cada vetor $v = (x,y) \in \mathbb{R}^2$ de um ângulo fixado θ.

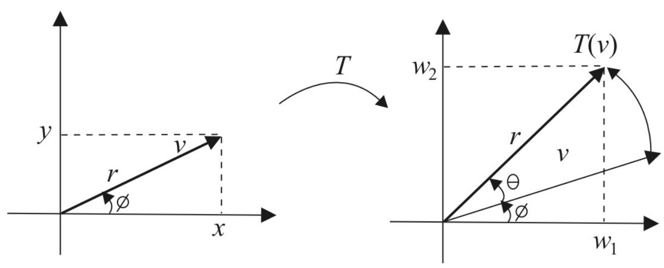

Figura 3 – Rotação

T é chamado de rotação por θ em $\mathbb{R}^2$. Escrevendo $w = T(v) = (w_1, w_2)$ segue da trigonometria que:

$$x = r\cos\varnothing \quad y = r\sin\varnothing \tag{3.3}$$

e

$$w_1 = r\cos(\theta + \varnothing), w_2 = r\sin(\theta + \varnothing) \tag{3.4}$$

onde r é o comprimento de v e $\varnothing$ é o ângulo entre v e o eixo Ox positivo no sentido anti - horário. Aplicando as identidades trigonométricas em 3.4 temos:

$$\begin{cases} w_1 = r\cos\theta\cos\varnothing - r\sin\theta\sin\varnothing \\ w_2 = r\sin\theta\cos\varnothing + r\cos\theta\sin\varnothing \end{cases}$$

Substituindo 3.3 nas expressões acima, temos:

$$\begin{cases} w_1 = x\cos\theta - y\sin\theta \\ w_2 = x\sin\theta + y\cos\theta \end{cases}$$

Assim se α denota a base canônica de $\mathbb{R}^2$, obteremos:

$$[T(v)]_\alpha = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} [v]_\alpha$$

Em geral a rotação de vetores em $\mathbb{R}^3$ é feita em relação a uma reta partindo da origem, chamada eixo de rotação. À medida que um vetor gira em torno do eixo de rotação, ele varre uma porção de um cone.

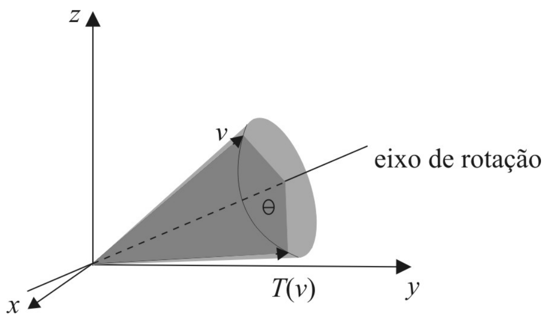

Figura 4 – Rotação

O ângulo de rotação que é medido na base do cone é descrito no sentido horário ou anti - horário, em relação a um ponto de vista ao longo do eixo de rotação olhando para a origem.

Assim como em $\mathbb{R}^2$, os ângulos são positivos se gerados por rotações no sentido anti - horário e negativos se gerados por rotações no sentido horário.

Tabela 3 – Rotações em $\mathbb{R}^3$ cujos eixos de rotação são os eixos coordenados

Operador	Equações	Matriz $[T]^\alpha_\alpha$
Rotação anti - horária em torno do eixo Ox por um ângulo θ	$\begin{cases} w_1 = x \\ w_2 = y\cos\theta - z\sin\theta \\ w_3 = y\sin\theta + z\cos\theta \end{cases}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$
Rotação anti - horária em torno do eixo Oy por um ângulo θ	$\begin{cases} w_1 = x\cos\theta + z\sin\theta \\ w_2 = y \\ w_3 = -x\sin\theta + z\cos\theta \end{cases}$	$\begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$
Rotação anti - horária em torno do eixo Oz por um ângulo θ	$\begin{cases} w_1 = x\cos\theta - y\sin\theta \\ w_2 = x\sin\theta + y\cos\theta \\ w_3 = z \end{cases}$	$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$

• **Homotetias:**

Sabemos que a multiplicação por um escalar de um vetor $\mathbb{R}^2$ e em $\mathbb{R}^3$, dependendo do valor, produz no vetor uma dilatação, contração ou inversão.

Podemos representar estes efeitos geométricos por meio de operadores lineares.

Sendo o operador linear $T_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$, dado por $T_\alpha(v) = \alpha v$, em que $\alpha \in \mathbb{R}$ e $v \in \mathbb{R}^2$, a oparação dilata v, se $\alpha \geq 1$, contrai v, se $0 \leq \alpha < 1$; inverte o sentido de v, se $\alpha < 0$. No caso particular de $\alpha = -1$ o operador T_α é chamado reflexão em torno da origem. Isso vale também para o $\mathbb{R}^3$.

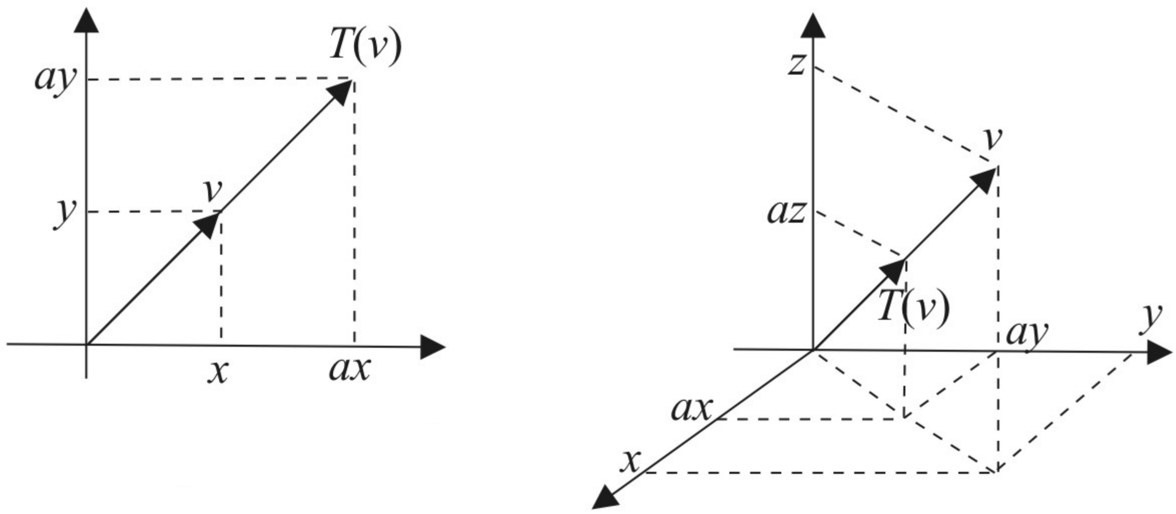

Figura 5 – Dilatação e contração por α em $\mathbb{R}$

3.2.2 Mudança de Base e Matrizes semelhantes

Definição 18. Dado um espaço vetorial V arbitrário de dimensão finita e duas bases α e β de V, podemos obter uma relação entre as matrizes $[v]_\alpha$ e $[v]_\beta$ de um vetor v em V, usando para isto, o operador identidade em V, usando a expressão:

$$[v]_\beta = [Iv]_\beta^\alpha \cdot [v]_\alpha \quad \forall\, v \in V.$$

A matriz $[Iv]_\beta^\alpha$ é chamada matriz mudança de base de α para β, pois pela igualdade acima, ela nos permite obter as coordenadas de um vetor v em V em relação à base β uma vez conhecidas suas coordenadas na base α.

Exemplo 25. Considerando as bases $\alpha = \{(1,0,1),(1,1,1),(1,1,2)\}$ e $\beta = \{(1,0,0),(0,1,0),(0,0,1)\}$

e encontraremos $[Iv]^{\alpha}_{\beta}$. Precisamos resolver

$$\begin{aligned}(1,0,0) &= a_{11}(1,0,1)+a_{21}(1,1,1)+a_{31}(1,1,2)\\(0,1,0) &= a_{12}(1,0,1)+a_{22}(1,1,1)+a_{32}(1,1,2)\\(0,0,1) &= a_{13}(1,0,1)+a_{23}(1,1,1)+a_{33}(1,1,2)\end{aligned}$$

$\Longleftrightarrow$

$$\begin{aligned}(a_{11}+a_{21}+a_{31},a_{21}+a_{31},a_{11}+a_{21}+2a_{31}) &= (1,0,0)\\(a_{12}+a_{22}+a_{32},a_{22}+a_{32},a_{12}+a_{22}+2a_{32}) &= (0,1,0)\\(a_{13}+a_{23}+a_{33},a_{23}+a_{33},a_{13}+a_{23}+2a_{33}) &= (0,0,1)\end{aligned}$$

Cada linha acima representa um sistema de três equações com três incógnitas e a matriz associada a cada um desses sistemas é a mesma, o que muda são os nomes das variáveis e o segundo membro. Utilizando como variáveis x, y e z, basta resolvermos o seguinte sistema

$\begin{pmatrix}1&1&1\\0&1&1\\1&1&2\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}a\\b\\c\end{pmatrix}$, onde a, b e $c\in\mathbb{R}$.

O sistema acima é equivalente a $\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}a\\b\\c-a\end{pmatrix}$

cuja solução única é dada por $x=a-b,\ \ y=a+b-c$ e $z=c-a$.

Tomando $(a,b,c)=(1,0,0)$ obtemos $(a_{11},a_{21},a_{31})=(1,1,-1)$,
$(a,b,c)=(0,1,0)$ obtemos $(a_{12},a_{22},a_{32})=(-1,1,0)$,
$(a,b,c)=(0,0,1)$ obtemos $(a_{13},a_{23},a_{33})=(0,-1,1)$.

Desta forma, $[Iv]^{\alpha}_{\beta}=\begin{pmatrix}1&-1&0\\1&1&-1\\-1&0&1\end{pmatrix}$.

Teorema 5. Sejam α e β duas bases de um espaço de dimensão finita V. Temos que a matriz $[Iv]^{\alpha}_{\beta}$ é invertível e sua inversa é a matriz $[Iv]^{\beta}_{\alpha}$, ou seja,

$$\left([Iv]^{\alpha}_{\beta}\right)^{-1}=[Iv]^{\beta}_{\alpha}.$$

Teorema 6. Sejam α e β duas bases de um espaço vetorial de dimensão finita V. Se T é um operador linear em V, então

$$[T]^{\alpha}_{\alpha}=P^{-1}\cdot[T]^{\beta}_{\beta}\cdot P$$

onde $P = [Iv]_{\beta}^{\alpha}$

Se A e B são matrizes quadradas de mesma ordem, dizemos que B é semelhante a A, quando existir uma matriz invertível P tal que $B = P^{-1}AP$. É facil verificar que se uma matriz B é semelhante a A, então A também é semelhante a B, desta forma, dizemos que A e B são semelhantes. Pelo teorema acima temos que $[T]_{\alpha}^{\alpha}$ e $[T]_{\beta}^{\beta}$ são semelhantes.

Exemplo 26. Vamos verificar se as matrizes $A = \begin{pmatrix} 5 & 2 \\ -8 & -3 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ são semelhantes. Para isso devemos encontrar uma matriz invertível P tal que $PA = BP$. Se tal matriz P existir ela é uma matriz quadrada de ordem 2 que pode ser denotada por $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$. Assim:

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix} \begin{pmatrix} 5 & 2 \\ -8 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

Esta igualdade equivale ao sistema $\begin{cases} 4x - 8y - 2z = 0 \\ 2x - 4y - 2t = 0 \\ 4z - 8t = 0 \end{cases}$, que admite a solução não trivial $(3,1,2,1)$. Portanto, $P = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$ e A e B são semelhantes.

CAPÍTULO

4

ESPAÇO COM PRODUTO INTERNO

Neste capítulo vamos apresentar a noção de produto interno em espaços vetoriais.Esta noção generaliza a noção de produto escalar em $\mathbb{R}^2$ e em $\mathbb{R}^3$ enriquecendo a estrutura de um espaço vetorial. Os resultados deste capítulo também podem ser encontrados em (HEFEZ A.; SOUZA FERNANDES, 2012) e (ZANI,).

Definição 19. Seja V um espaço vetorial. Um produto interno em V é uma função que a cada par de vetores u e v em V associa um número real, denotado por $< u,v >$ que satisfaz as seguintes condições:

Para quaisquer vetores u,v e w em V e $\forall$ número real α,

P1. $< u,v >\geq 0$

P2. $< u,v >= 0$ se, e somente se, $v = 0$

P3. $< u,v >=< v,u >$

P4. $< u+v,w >=< u,w > + < v,w >$

P5. $< \alpha u,v >= \alpha < u,v >$

Exemplo 27. Sejam $u = (x_1,x_2,\cdots,x_n)$ e $v = (y_1,y_2,\cdots,y_n)$ em $\mathbb{R}^n$. Definimos

$$< u,v >= x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{4.1}$$

Note que $< u,u >= x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0$ e que $< u,v >= x_1y_1 + x_2y_2 + \cdots + x_ny_n == y_1x_1 + y_2x_2 + \cdots + y_nx_n =< v,u >$, mostrando que as condições 1 e 3 da definição são satisfeitas. A condição 2 também é satisfeita já que

$$< u,u >= x_1^2 + x_2^2 + \cdots + x_n^2 = 0 \iff x_1 = x_2 = \cdots = x_n = 0.$$

Se $w = (z_1, z_2, \cdots, z_n)$ então: $< u+v, w >= (x_1+y_1)z_1 + (x_2+y_2)z_2 + \cdots + (x_n+y_n)z_n$ $= (x_1z_1 + x_2z_2 + \cdots + x_nz_n) + (y_1z_1 + y_2z_2 + \cdots + y_nz_n) =< u,w > + < v,w >$, mostrando que a condição 4 também é satisfeita e a condição 5 também segue, pois se $k \in \mathbb{R}$, então:

$$< ku, v >= (kx_1)y_1 + (kx_2)y_2 + \cdots + (kx_n)y_n$$

$$= k(x_1y_1 + x_2y_2 + \cdots + x_ny_n)$$

$$= k < u,v >$$

Assim, 4.1 define um produto interno em $\mathbb{R}^n$, chamado produto interno usual de $\mathbb{R}^n$ ou produto escalar de $\mathbb{R}^n$, generalizando a noção de produto escalar de $\mathbb{R}^2$ e de $\mathbb{R}^3$.

Definição 20. Seja V um espaço com produto interno. Definimos norma do vetor v de V, ou comprimento de v, denotado por $\|v\|$, como o número real

$$\|v\| = \sqrt{< v,v >}.$$

Se $\|v\| = 1$, dizemos que v é um vetor unitário.

A distância de (u,v) entre dois vetores u e v de V é definida como:

$$d(u,v) = \|u-v\| = \sqrt{< u-v, u-v >}.$$

Exemplo 28. Se $u = (x_1, x_2, \cdots, x_n)$ e $v = (y_1, y_2, \cdots, y_n)$ são vetores em $\mathbb{R}^n$, com o produto interno usual, então

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

e

$$d(u,v) = \|u-v\| = \sqrt{< u-v, u-v >}$$

$$= \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + \cdots + (x_n-y_n)^2}.$$

Proposição 18. Seja V um espaço vetorial com um produto interno. Temos:

1. $\|\alpha u\| = |\alpha|\ \|u\|$, para todo $u \in V$ e todo $\alpha \in \mathbb{R}$.

2. $\|u\| \geq 0$, para todo $u \in V$.

3. $\|u\| = 0$ se, e somente se, $u = 0$.

4. $| < u,v > | \leq \|u\|\,\|v\|$ para todo $u,v \in V$ (desigualdade de Cauchy - Schwarz).

5. $\|u+v\| \leq \|u\| + \|v\|$, para todo $u,v \in V$ (desigualdade triângular).

Proposição 19. (Identidade do Paralelogramo) Sejam u e v vetores de um espaço vetorial V com um produto interno. Então:

$$\|u+v\|^2+\|u-v\|^2=2(|u\|^2+\|v\|^2).$$

Demonstração.

$$\|u+v\|^2+\|u-v\|^2=<u+v,u+v>+<u-v,u-v>$$

$$=<u,u>+<v,v>+2<u,v>+<u,u>+<v,v>-2<u,v>$$

$$=2<u,u>+2<v,v>=2(\|u\|^2+\|v\|^2).$$

□

Proposição 20. Se u,v e w são vetores em um espaço com produto interno V, então:

1. $d(u,v)\geq 0$.
2. $d(u,v)=$0 se, e somente se, $u=v$.
3. $d(u,v)=d(v,u)$.
4. $d(u,v)\leq d(u,w)+d(w,v)$ (desigualdade triângular).

Seja V um espço vetorial com produto interno e $u,v\in V$ ambos não nulos. Pela desigualdade de Cauchy - Schwarz, ver proposição (18), temos:

$$-\|u\|\,\|v\|\leq<u,v>\leq\|u\|\,\|v\|$$

ou ainda

$$-1\leq\frac{<u,v>}{\|u\|\,\|v\|}\leq 1$$

Desta forma, existe um único número real $\theta\in[0,\pi]$ tal que:

$$\cos\theta=\frac{<u,v>}{\|u\|\,\|v\|}$$

Este número θ é chamado de ângulo entre os vetores u e v.

Exemplo 29. Sabe-se que $\|u\|=\|v\|=1$ e $\|u-v\|=2$. Calcule o ângulo entre u e v.

Como $\|u-v\|=2$, então:

$$4=\|u-v\|^2=<u-v,u-v>=\|u\|+\|v\|-2<u,v>=2-2<u,v>$$

Assim, $<u,v>=1$ e $\cos\theta=\dfrac{<u,v>}{\|u\|\,\|v\|}=-1$, ou seja, $\theta=\pi$.

Definição 21. Seja V um espaço com produto interno. Dizemos que $u, v \in V$ são ortogonais se $< u, v >= 0$ e, neste caso, denotamos por $u \perp s$.

Dizemos que um conjunto $S = \{u_1, \cdots, u_n\} \subset V$ é ortogonal se $u_i \perp u_j$ quando $i \neq j$.

Dizemos que um conjunto $S = \{u_1, \cdots, u_n\} \subset V$ é ortonormal se for ortogonal e também $\|u_j\| = 1, j = 1, 2, \cdots, n$.

Dizemos que $u \in V$ é ortogonal a um subconjunto não vazio S de V se u for ortogonal a todos os elementos de S. Neste caso usaremos a definição de $u \perp S$.

Exemplo 30. Seja um vetor arbitrário $u = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$. Então um vetor $v = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$ é ortogonal a u se $< u, v >= x_1y_1 + x_2y_2, \cdots + x_ny_n = 0$.

Observação 3. É fácil verificar que

1. Se $u = 0$ ou $v = 0$, então $u \perp v$. Se $u \neq 0$ e $v \neq 0$ então $u \perp v$ se, e somente se o ângulo entre u e $v = \dfrac{\pi}{2}$.
2. Se $S = \{u_1, \cdots, u_n\} \subset V$ é um conjunto ortogonal com $u_j \neq 0, j = 1, 2, \cdots, n$, então $S' = \left\{ \dfrac{u_1}{\|u_1\|}, \cdots, \dfrac{u_n}{\|u_n\|} \right\}$ é um conjunto ortonormal.

Proposição 21. Todo conjunto ortogonal de vetores não nulos de V é linearmente independente.

Demonstração. Se

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \tag{4.2}$$

então tomando o produto interno do vetor acima com u_1 e lembrando que $< u_1, u_1 >= \|u_1\|^2 = 1$ e $< u_j, u_1 >= 0$ se $j = 2, 3, \cdots, n$, obteremos:

$$\alpha_1 = \alpha_1 < u_1, u_1 > + \cdots + \alpha < u_n, u_1 >=< 0, u_1 >= 0$$

isto é, $\alpha_1 = 0$ e 4.2 fica $\alpha_2 u_2 + \cdots + \alpha_n u_n = 0$.

Tomando o produto interno do vetor acima com u_2, obtemos como acima, que $\alpha_2 = 0$.

Repetindo o processo chegamos à conclusão que a única possibilidade para 4.2 é $\alpha_1 = \cdots = \alpha_n = 0$. □

Observação 4. A proposição acima continua válida se S for apenas um conjunto ortogonal com elementos não nulos.

Definição 22. Se V é um espaço vetorial com produto interno de dimensão n e se $\{u_1, \cdots, u_n\}$ é conjunto ortonormal, então diremos que estes vetores formam uma base ortonormal de V.

Proposição 22. Sejam V um espaço euclidiano que possui uma base ortonormal dada por $\{u_1, \cdots, u_n\}$. Então, se $u \in V$ temos

$$u =< u, u_1 > u_1 + \cdots + < u, u_n > u_n.$$

Demonstração. Como $u_1, \cdots, u_n$ formam uma base de V, existem $\alpha_1, \cdots, \alpha_n \in \mathbb{R}$ tais que $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Tomando o produto interno de V com v_1, temos que,

$< v, v_1 >= \alpha_1 < v_1, v_1 > + \cdots + \alpha_n < v_n, v_1 >= \alpha_1$, pois a base é ortogonal. O resultado segue tomando o produto interno de v por v_2, v_3, etc. $\square$

Se $\beta = \{v_1, v_2, \cdots, v_n\}$ é uma base ortogonal de V normalizando cada um dos vetores de β, obteremos a base ortonormal α de V, onde $\alpha = \left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \cdots, \frac{v_n}{\|v_n\|} \right\}$.

Proposição 23. Seja w um vetor não nulo de V. Se $v \in V$, então:

$$k = \frac{< v, w >}{< w, w >} = \frac{< v, w >}{\|w\|^2}$$

é o único número real tal que $v' = v - kw$ é ortogonal a w.

A projeção de v ao longo de w, é denotado por $proj_w(v)$ e é definida por:

$$proj_w(v) = kw = \frac{< v, w >}{< w, w >} w.$$

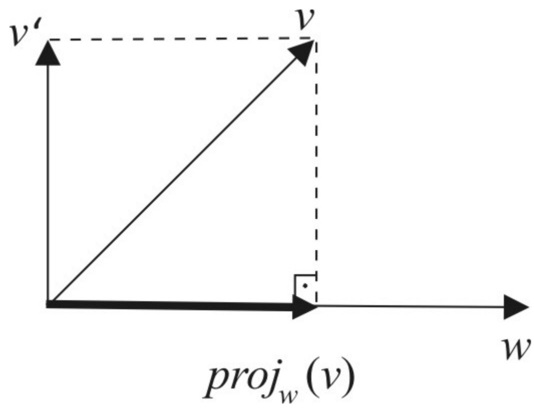

Figura 6 – Projeção de v ao longo de w.

Proposição 24. Suponhamos que $\{w_1, w_2, \cdots, w_r\}$ seja um conjunto ortogonal de vetores não nulos de V. Se $v \in V$, então $k_i = \frac{< v, w_i >}{\|w_i\|^2}, 1 \leq i \leq r$, são os únicos números reais tais que o vetor

$$v' = v - k_1 w_1 - k_2 w_2 - \cdots - k_r w_r$$

é ortogonal aos vetores $w_1, w_2, \cdots, w_r$.

Demonstração. Para $i = 1, 2, \cdots, r$ e usando $< w_i, w_j >= 0$ para $i \neq j$, temos:

$$< v - k_1 w_1 - k_2 w_2 - \cdots - k_r w_r, w_i >=< v, w_i > - k_i < w_i, w_i >=$$
$$=< v, w_i > - \frac{< v, w_i >}{< w_i, w_i >} < w_i, w_i >= 0$$

□

Com estes resultados, podemos afirmar o seguinte teorema:

Teorema 7. Um espaço vetorial V com produto interno possui uma base ortogonal.

Demonstração. Seja $\{v_1, v_2, \cdots, v_n\}$ uma base para um espaço V com produto interno. Podemos obter uma uma base ortogonal $\{w_1, w_2, \cdots, w_n\}$ para V da maneira descrita a seguir. Considere

$$w_1 = v_1$$
$$w_2 = v_2 - \frac{< v_2, w_1 >}{\|w_1\|^2} w_1$$
$$w_3 = v_3 - \frac{< v_3, w_1 >}{\|w_1\|^2} w_1 - \frac{< v_3, w_2 >}{\|w_2\|^2} w_2$$
$$\vdots$$
$$w_n = v_n - \frac{< v_n, w_1 >}{\|w_1\|^2} w_1 - \frac{< v_n, w_2 >}{\|w_2\|^2} w_2 - \cdots - \frac{< v_n, w_n >}{\|w_n\|^2} w_n.$$

Pela proposição 24, o conjunto $\{w_1, w_2, \cdots, w_n\}$ é um conjunto ortogonal. Além disso, como o conjunto $\{v_1, v_2, \cdots, v_n\}$ é linearmente independente, cada vetor w_i é não nulo. Assim, o conjunto $\{w_1, w_2, \cdots, w_n\}$ é um conjunto ortogonal de vetores não nulos de V. Como, por definição, $n = \dim V$, $\{w_1, w_2, \cdots, w_n\}$ é base ortogonal de V. A normalização de cada w_k garante uma base ortonormal para V. □

Exemplo 31. Para encontrar uma base ortonormal para $W = \{(x, y, z) \in \mathbb{R}^3, x - 2y = 0\}$, notamos, primeiramente, que $(x, y, z) \in W$ se

$$(x, y, z) = (2y, y, z) = y(2, 1, 0) + z(0, 0, 1).$$

Assim, $(2, 1, 0)$ e $(0, 0, 1)$ formam uma base de W. A base ortonormal para W será dada pelos vetores u_1 e u_2, considerando $u_1 = (0, 0, 1)$ pois é um vetor unitário e pelo processo anterior, u_2 é a projeção ortogonal unitária de $(2, 1, 0)$ sobre u_1, isto é:

$$u_2 = \frac{(2,1,0) - < (2,1,0), (0,0,1) > (0,0,1)}{\|(2,1,0) - < (2,1,0), (0,0,1) > (0,0,1)\|} = \frac{(2,1,0)}{\|2,1,0\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0 \right).$$

Operadores em Espaços com Produto Interno

Agora, vamos mostrar a existência do operador adjunto de um operador linear e apartir deste, introduzir as noções de operadores simétricos e operadores ortogonais. Suponhamos que V é um espaço com produto interno de dimensão finita $n > 0$.

Mostraremos que existe um isomorfismo entre V e $\mathscr{L}(V,\mathbb{R})$. Dado um vetor $v \in V$, a ele associamos de modo natural um funcional linear em V, como segue:

$$\begin{aligned} \phi_v : V \to \ & \mathbb{R} \\ u \to \ & < u, v > \end{aligned}$$

De fato ϕ_v é um funcional linear, pois, para cada $u_1, u_2 \in V$ e todo $\alpha \in \mathbb{R}$ temos:

$\phi_v(u_1 + \alpha u_2) =< u_1 + \alpha u_2, v >=< u_1, v > +\alpha < u_2, v >= \phi_v(u_1) + \alpha \phi_v(u_2)$.

Logo, cada $v \in V$ define um funcional linear ϕ_v em V, ou seja, um elemento de $\mathscr{L}(V,\mathbb{R})$.

Suponhamos, agora, que $\phi_v = \phi_{v'}$, isto é, $< u, v >=< u, v' >$ para todo $u \in V$. Logo $< u, v - v' >= 0$, para todo $u \in V$. Portanto, $v - v'$ é ortogonal a todos os vetores de V o que acarreta que $v = v'$. Desta forma, a função $v \mapsto \phi_v$, onde $\phi_v(u) =< u, v >, (u \in V)$, é um isomorfismo entre V e $\mathscr{L}(V,\mathbb{R})$.

Teorema 8. Dado o operador linear T em V, existe um único operador linear T^* em V tal que:

$$< T(v), w >=< v, T^*(w) >, \forall v, w \in V.$$

Demonstração. Tome $w \in V$. Como a função definida por $v \mapsto < T(v), w >$ é um funcional linear em V, segue do resultado anterior que existe um único vetor $w' \in V$ tal que:
$< T(v), w >=< v, w' >$, para todo $v \in V$. Basta definir $T^*(w) = w'$.

Se $v_1, \cdots$,v_n é uma base ortonormal de V, então,

$w' =< w', v_1 > v_1 + \ldots + < w', v_n > v_n =< T(v_1), w > v_1 + \ldots < t(v_n), w > v_n$.

Como $T^*(w) = w'$, podemos ver pela igualdade anterior que T^* é linear. □

Definição 23. O operador T^* acima definido é chamado de operador adjunto de T.

Assim, o teorema anterior afirma que todo operador linear T, em um espaço com o produto interno de dimensão finita, possui um operador adjunto T^*.

Proposição 25. Para toda base ortonormal α de V e para todo operador linear T em V, temos que

$$[T^*]_\alpha^\alpha = \left[[T]_\alpha^\alpha\right]^t.$$

Para a prova desta proposição precisamos do lema a seguir.

Lema 1. Seja $\alpha = \{v_1, \cdots, v_n\}$ uma base ortonormal de V. Se $A = [a_{ij}]_{n\times n}$ é a matriz que representa um operador T em V, com relação à base α (ou seja, $A = [T]_\alpha^\alpha$), então:

$$a_{ij} =< T(v_j), v_i > \text{para todos } i, j,\ 1 \leq i, j \leq n.$$

Demonstração do Teorema: Considere as matrizes $[T]_\alpha^\alpha = [a_{ij}]_{n\times n}$ e $[T^*]_\alpha^\alpha = [b_{ij}]_{n\times n}$.

Pelo lema 1,

$a_{ij} =< T(v_j), v_i >$ e $b_{ij} =< T^*(v_j), v_i >$, para todos i, j, $1 \leq i, j \leq n$.

Logo,

$$b_{ij} =< T^*(v_j), v_i >=< v_i, T^*(v_j) >=< T(v_i), v_j >= a_{ji}, \text{para todos } i, j, \text{com} 1 \leq i, j \leq n.$$

Definição 24. Um operador linear $T : V \to V$ é dito operador simétrico quando $T^* = T$.

Pela proposição 25, observamos que se T é um operador simétrico em V, então para toda base ortonormal α de V temos:

$$[T]_\alpha^\alpha = ([T]_\alpha^\alpha)^t$$

Assim $T : V \to V$ é simétrico se, e somente se, $[T]_\alpha^\alpha$ é uma matriz simétrica.

Exemplo 32. Seja $T : \mathbb{R}^3 \to \mathbb{R}^3$ o operador linear definido por $T(x,y,z) = (2x-y+z, -x+y+3z, x+3y)$ Se α é a base canônica de $\mathbb{R}^3$, então:

$$[T]_\alpha^\alpha = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 3 \\ 1 & 3 & 0 \end{pmatrix}$$

é uma matriz simétrica, e portanto, T é um operador simétrico.

Definição 25. Um operador linear $T : V \to V$ é dito ser um operador ortogonal quando

$T^*T = TT^* = I_v$.

Dizemos que um operador T em V preserva norma, preserva distância, ou preserva produto interno, quando para todos $u, v \in V$, se tenha $\|T(v)\| = \|v\|, d(T(u), T(v)) = d(u, v)$, ou $< T(u), T(v) >=< u, v >$, respectivamente.

Teorema 9. Seja $T : V \to V$ um operador linear. As seguintes afirmações são equivalentes:

i. T é ortogonal.

ii. T preserva a norma.

iii. T preserva a distância.

iv. T preserva o produto interno.

v. T transforma toda base ortonormal numa base ortonormal.

vi. T transforma alguma base ortonormal numa base ortonormal.

Demonstração. $(i) \Rightarrow (ii)$

Se $v \in V$, então:

$\|T(v)\|^2 =< T(v), T(v) >=< v, T^*(T(v)) >=< v, I_v(v) >=< v, v >= \|v\|^2$.

$(ii) \Rightarrow (iii)$

Se $v, u \in V$, então $d(T(v), T(u)) = \|T(v) - T(u)\| = \|T(v-u)\| = \|v-u\| = d(v,u)$.

$(iii) \Rightarrow (iv)$

Se $v, u \in V$, então $d(T(v+u), 0) = d(v+u, 0)$ ou seja, $\|T(v+u)\|^2 = \|v+u\|^2$

Note que:

$\|T(v+u)\| =< T(v), T(v) > +2 < T(v), T(u) > + < T(u), T(u) >$ e

$\|v+u\|^2 =< v, v > +2 < v, u > + < u, u >$

Como:

$< v, v >= (d(v,0))^2 = (d(T(v),0))^2 =< T(v), T(v) >$

o mesmo valendo para u, temos que $< T(v), T(u) >=< v, u >$, como desejado.

$(iv) \Rightarrow (i)$

Se $v, u \in V$, então:

$$< v, u >=< T(v), T(u) >=< v, T^*(T(u)) >,$$

mostrando que para todo $u, v \in V$,

$$< v, (T^*T - I_v)(u) >= 0$$

Temos que $(T^*T - I_v)(u) = 0$, para todo $u \in V$, o que acarreta que $T^*T = I_v$, logo T é ortogonal.

$(i) \Rightarrow (v)$

Seja $v_1, v_2, \cdots, v_n$ uma base ortonormal de V. Então

$$< T(v_i), T(v_j) >=< v_i, v_j >= \begin{pmatrix} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j \end{pmatrix}$$

Logo, o conjunto $T(v_1), T(v_2), \cdots, T(v_n)$ é ortonormal e consequentemente, linearmente independente. Como $\dim V = n$, concluímos que esse conjunto é uma base de V.

$(v) \Rightarrow (vi)$

Esta implicação é óbvia.

$(vi) \Rightarrow (iv)$

Seja $\{v_1, v_2, \cdots, v_n\}$ uma base ortonormal de V tal que $\{T(v_1), T(v_2), \cdots, T(v_n)\}$ também é uma base ortonormal de V. Sejam v e u em V com

$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ e $u = b_1v_1 + b_2v_2 + \cdots + b_nv_n$, então

$$< v, u >= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j < v_i, v_j >= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j. \tag{4.3}$$

Por outro lado, temos

$$T(v) = a_1T(v_1) + \cdots + a_nT(v_n)$$

e

$$T(u) = b_1T(v_1) + \cdots + b_nT(v_n),$$

donde

$$< T(v), T(u) >= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j < T(v_i), T(u_j) >= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j. \tag{4.4}$$

Assim de 4.3 e 4.4 concluímos que

$$< T(v), T(u) >=< v, u > .$$

□

Definição 26. Uma matriz quadrada A tal que $A^tA = I$ é chamada de matriz ortogonal.

Proposição 26. Seja $A = [a_{ij}]_{n\times n}$ uma matriz. São equivalentes

i. A é ortogonal.

ii. As colunas de A formam um conjunto ortonormal em $\mathbb{R}^n$.

iii. As linhas de A formam um conjunto ortonormal em $\mathbb{R}^n$, *Demonstração.* $(i) \Leftrightarrow (ii)$ Chamaremos $A^tA = [b_{ij}]_{n\times n}$.Pela definição de produto de matrizes, o elemento b_{ij} é dado por:

$$b_{ij} = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} =< (a_{1i}, a_{2i}, \cdots, a_{ni}), (a_{1j}, a_{2j}, \cdots, a_{nj}) >$$

Assim,

$A^tA = I_n$ se, e somente se,

$$< (a_{1i}, a_{2i}, \cdots, a_{ni}), (a_{1j}, a_{2j}, \cdots, a_{nj}) >= \begin{cases} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j. \end{cases}$$

$(i) \Leftrightarrow (iii)$ Basta utilizar o fato que A é ortogonal se, e somente se, A^t é ortogonal, que as linhas de A^t são as colunas de A e aplicar o que foi provado acima. □

Teorema 10. Se α e β são bases ortonormais de V, então a matriz de mudança de base $[I_v]^{\alpha}_{\beta}$ é uma matriz ortogonal.

Demonstração. Sejam $\alpha = \{v_1, v_2, \cdots, v_n\}$ e $\beta = \{w_1, w_2, \cdots, w_n\}$. Suponhamos $[I_v]^{\alpha}_{\beta} = [a_{ij}]$. Para $1 \leq i \leq n$ temos que:

$$v_i = a_{1i}w_1 + a_{2i}w_2 + \cdots + a_{ni}w_n$$

Como v_i e v_j são ortogonais, quando $i \neq j$, então:

$$\begin{aligned} 0 &=< v_i, v_j >= a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} \\ &=< (a_{1i}, a_{2i}, \cdots, a_{ni}), (a_{1j}, a_{2j}, \cdots, a_{nj}) > \end{aligned}$$

pois β é ortonormal e concluímos que as colunas de $[I_v]^{\alpha}_{\beta}$ formam vetores ortogonais em $\mathbb{R}^n$. Vejamos agora que cada coluna de $[I_v]^{\alpha}_{\beta}$ formam um vetor unitário em $\mathbb{R}^n$. De fato, se $1 \leq i \leq n$, então:

$$1 =< v_i, v_i >= a_{1i}^2 + a_{2i}^2 + \cdots + a_{ni}^2,$$

já que β é ortonormal. Assim, as colunas de $[I_v]^{\alpha}_{\beta}$ formam vetores unitários em $\mathbb{R}^n$ e $[I_v]^{\alpha}_{\beta}$ é uma matriz ortogonal.

Agora vamos mostrar a relação entre os operadores ortogonais e as matrizes ortogonais.

Sejam dados um espaço vetorial V, com uma base $\alpha = \{v_1, v_2, \cdots, v_n\}$ e uma matriz quadrada $A = [a_ij]$ de ordem n. Podemos, como já vimos, associar à matriz A um operador linear T_A, definido como se segue

$$T_A(v) = (a_{11}x_1 + \cdots + \cdots + a_{1n}x_n, \cdots, a_{n1}x_1 + \cdots + a_{nn}x_n),$$

onde $x_1, \cdots, x_n$ são coordenadas de v relativamente à base α, ou seja

$$v = x_1 v_1 + \cdots + x_n v_n$$

$\square$

Proposição 27. Sejam α uma base ortonormal de V, T um operador linear em V e $A \in M_{n\times m}$

i. T é ortogonal se, e somente se, $[T]_{\alpha}^{\alpha}$ é ortogonal.

ii. A é ortogonal se, e somente se, T_A é ortogonal.

CAPÍTULO

5

DIAGONALIZAÇÃO DE OPERADORES

5.1 Operadores Diagonalizáveis

Vimos anteriormente que uma operador linear $T : V \to V$, onde V tem dimensão finita, pode ser representada por uma matriz. Sendo as matrizes diagonais as mais simples do ponto de vista das operações matriciais, queremos saber se para todo operador T existe uma base α de V tal que $[T]_\alpha^\alpha$ seja uma matriz diagonal. A resposta é que nem sempre existe tal base. Por exemplo, o operador T em $\mathbb{R}^2$, cuja matriz na base canônica é dada por:

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

não admite uma tal representação. De fato, se fosse possível achar uma base α tal que a matriz de T nesta base fosse diagonal, teríamos $PAP^{-1} = C$ onde P é uma matriz 2×2 invertível e C uma matriz diagonal. Como $A^2 = 0$, isto acarretaria que:

$$C^2 = (PAP^{-1})^2 = PA^2P^{-1} = 0$$

Logo $C = 0$, o que implica que $A = 0$; uma contradição.

Definição 27. Dizemos que um operador definido sobre um espaço vetorial V de dimensão finita é diagonalizável, quando for possível representá-lo por uma matriz diagonal em alguma base de V.

Nosso objetivo, então, é obter resultados para garantir se um operador é diagonalizável ou não. Para isso, iniciamos com a definição de autovalor e autovetor associados a um operador linear T. Para descrever o conteúdo deste capítulo utilizamos a bibliografia (HEFEZ A.; SOUZA FERNANDES, 2012) e (ZANI,).

Definição 28. Seja $T : V \to V$ um operador linear. Um número real α será dito um autovalor de T se existir um vetor não nulo v, chamado de autovetor de T associado a α, tal que $T(v) = \alpha v$.

Exemplo 33. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ o operador linear dado por $T(x,y) = (4x-y, 2x+y)$. Vamos determinar $\alpha \in \mathbb{R}$ e $v = (x,y) \in \mathbb{R}^2$, não nulo, tais que $T(x,y) = \alpha(x,y)$, ou seja $(4x-y, 2x+y) = \alpha(x,y)$.

Resolvendo o sistema $\begin{cases} 4x - y = \alpha x \\ 2x + y = \alpha y \end{cases}$, temos que $\alpha_1 = 3$ e $\alpha_2 = 2$ são os autovalores de T. Vamos agora calcular os autovetores de T. Primeiramente, para $\alpha = 2$, fazemos:

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

Assim, $y = 2x$ e o conjunto solução da equação $T(x,y) = 2(x,y)$ é dado por $\{(x,2x); x \in \mathbb{R}\}$. Logo, os autovetores de T associados a $\alpha = 2$ são os vetores da forma $(x,2x)$ em que $x \in \mathbb{R}$ e $x \neq 0$.

Agora, vamos calcular os autovetores de T associados ao autovalor $\alpha = 3$, fazemos:

$$\begin{cases} 4x - y = 3x \\ 2x + y = 3y \end{cases}.$$

Assim $x = y$, e o conjunto solução dessa equação é dado por $\{(x,x); x \in \mathbb{R}\}$. Logo, os autovalores de T associados a $\alpha = 3$ são os vetores da forma (x,x) em que $x \in \mathbb{R}$ e $x \neq 0$.

O exemplo a seguir nos mostra que nem todo operador linear possui autovalores e autovetores.

Exemplo 34. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ o operador linear dado por $T(x,y) = (-y,x)$. Se $\alpha \in \mathbb{R}$ e $v = (x,y) \in \mathbb{R}^2$, $v \neq 0$, são tais que $T(x,y) = \alpha(x,y)$ então:

$$(-y,x) = \alpha(x,y).$$

Equivalentemente,

$$\begin{cases} \alpha x = -y \\ \alpha y = x \end{cases}$$

Assim $y(\alpha^2 + 1) = 0$ e como $\alpha \in \mathbb{R}$ a equação $y(\alpha^2 + 1) = 0$ é verificada apenas se $y = 0$, o que implicaria $x = 0$. Como v não é o vetor nulo, isso não pode ocorrer. Concluímos que T não tem autovalor e nem autovetores.

Proposição 28. Seja $T : V \to V$ um operador linear e sejam $c_1, c_2, \cdots, c_r$ autovalores distintos de T. Se $v_1, v_2, \cdots, v_r$ são autovetores associados aos autovalores $c_1, c_2, \cdots, c_r$, respectivamente, então $v_1, v_2, \cdots, v_r$ é linearmente independente.

Demonstração. A prova é feita por indução sobre r. O resultado é valido para $r = 1$ pois se $T : V \to V$ é um operador linear com autovalor c_1 e se v_1 é um autovetor de T associado a c_1, então v_1 é linearmente independente, pois $v_1 \neq 0$. Suponhamos agora o resultado válido para $r-1$, vamos provar para r, $r \geq 2$.

Considere a equação

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = 0 \tag{5.1}$$

onde $a_1, a_2, \cdots, a_r$ são números reais. Aplicando T em 5.1, e observando que $T(v_j) = c_jv_j$, $1 \leq j \leq r$, obtemos

$$a_1(c_1v_1) + a_2(c_2v_2) + \cdots + a_r(c_rv_r) = 0 \tag{5.2}$$

Por outro lado, T possui pelo menos um autovalor não nulo, que podemos supor $c_r \neq 0$. Multiplicando 5.1 por c_r, obtemos:

$$a_1(c_rv_1) + a_2(c_rv_2) + \cdots + a_r(c_rv_r) = 0 \tag{5.3}$$

De 5.2 e 5.3

$$a_1(c_1 - c_r)v_1 + a_2(c_2 - c_r)v_2 + \cdots + a_{r-1}(c_{r-1} - c_r)v_{r-1} = 0 \tag{5.4}$$

Pela hipótese de indução $\{v_1, v_2, \cdots, v_{r-1}\}$ é linearmente independente. Portanto, de 5.4 segue:

$$a_j(c_j - c_r) = 0,\ \ 1 \leq j \leq r-1. \tag{5.5}$$

Como os autovalores $c_1, c_2, \cdots, c_r$ são todos distintos, de 5.5 obtemos que $a_j = 0$ para todo $1 \leq j \leq r-1$. Substituindo os valores em 5.1, concluímos também que $a_r = 0$, já que $v_r \neq 0$. Portanto $\{v_1, v_2, \cdots, v_r\}$ é independente. □

Corolário 5. Seja $T : V \to V$ um operador linear. Se $\dim V = n$ e T possui n autovalores distintos, então V possui uma base formada por autovetores de T.

Demonstração. Pela proposição acima, n autovalores distintos implicam na existência de um conjunto de autovetores $v_1, v_2, \cdots, v_n$ linearmente independente. Como $G(v_1, v_2, \cdots, v_n) \subset V$ e $\dim(v_1, v_2, \cdots, v_n) = n = \dim V$, temos que $G(v_1, v_2, \cdots, v_n) = V$. Logo $\{v_1, v_2, \cdots, v_n\}$ é uma base de V. □

Definição 29. Seja A uma matriz quadrada de ordem n. A matriz $tI_n - A$, onde I_n é a matriz identidade de ordem n e t uma indeterminada é chamada de matriz característica de A. O determinante dessa matriz que é um polinômio em t, é o polinômio característico de A, denotado por $P_A(t)$.

Exemplo 35. Seja $\begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$, o polinômio característico de A será dado pelo determinante da matriz característica de A, $tI_2 - A = \begin{pmatrix} t-4 & 1 \\ -2 & t-1 \end{pmatrix}$, dado por

$$P_A(t) = \det \begin{pmatrix} t-4 & 1 \\ -2 & t-1 \end{pmatrix} = (t-4)(t-1)+2 = t^2 - 5t + 6.$$

Observamos que as raízes do polinômio deste exemplo, ou seja, os números reais t_0 tais que $P_A(t_0) = 0$, são os autovalores dados no exemplo 33.

Existe uma relação entre autovalores de um operador e as raízes do polinômio característico de alguma matriz associada a ele.

Teorema 11. Seja $T : V \to V$ um operador linear e seja $\alpha = \{v_1, v_2, \cdots, v_n\}$ uma base de V. Então:

(i) v é um autovetor de T associado a t_0 se, e somente se, v é uma solução não trivial do sistema linear $AX = 0$, onde $A = t_0 I_n - [T]_\alpha^\alpha$

(ii) $t_0 \in \mathbb{R}$ é um autovalor de T se, e somente se, t_0 é uma solução do polinômio característico da matriz $[T]_\alpha^\alpha$, ou seja, $P_{[T]_\alpha^\alpha}(t_0) = 0$

Demonstração. (i) Seja t_0 um autovalor de T e v um autovetor de T associado a t_0. Como $[T(v)]_\alpha = [T]_\alpha^\alpha [v]_\alpha$ e $T(v) = t_0 v$, temos:

$$[t_0 v]_\alpha = [T]_\alpha^\alpha [v]_\alpha$$

$t_0 I_n [v]_\alpha = [T]_\alpha^\alpha [v]_\alpha$.

Equivalentemente,

$$(t_0 I_n - [T]_\alpha^\alpha)[v]_\alpha = 0 \tag{5.6}$$

(ii) Considere o sistema linear $AX = 0$, onde $A = t_0 I_n - [T]_\alpha^\alpha$. De (i) segue que $AX = 0$ tem uma solução não trivial, a saber $[v]_\alpha$, já que v não é um vetor nulo. o que implica que A não é invertível. Assim, $P_{[T]_\alpha^\alpha}(t_0) = 0$, provando que t_0 é uma raíz de $P_{[T]_\alpha^\alpha}$. Reciprocamente, se $t_0 \in \mathbb{R}$ é uma raiz de $P_{[T]_\alpha^\alpha}$, então $P_{[T]_\alpha^\alpha}(t_0) = 0$. Portanto, o sistema linear $AX = 0$, onde $A = t_0 I_n - [T]_\alpha^\alpha$, tem uma solução $X_1 = \left([x_1 \quad x_2 \quad \cdots \quad x_n]^t \right)$ não nula, pois $\det A = 0$. Como $AX_1 = 0$,

$$(t_0 I_n - [T]_\alpha^\alpha) X_1 = t_0 X_1 - [T]_\alpha^\alpha X_1 = 0 \tag{5.7}$$

$$[t_0 v]_\alpha = t_0 [v]_\alpha = [T]_\alpha^\alpha [v]_\alpha = [T(v)]_\alpha \tag{5.8}$$

pois pela construção de v, $X_1 = [v]_\alpha$. Obtemos que $[T(v)]_\alpha = [t_0 v]_\alpha$,isto é, as coordenadas dos vetores $T(v)$ e $t_0 v$ na base α são iguais. Consequentemente, estes vetores são iguais, ou seja, $T(v) = t_0 v$. Como, por construção $v \neq 0$, segue-se que t_0 é um autovalor de T e v um autovetor de T associado a t_0. □

Exemplo 36. Vamos refazer o exemplo 33, utilizando o teorema acima. Reconsidere o operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por $T(x,y) = (4x - y, 2x + y)$ e seja α a base canônica de $\mathbb{R}^2$.

$$P_{[T]_\alpha^\alpha}(t) = \det \begin{pmatrix} t-4 & 1 \\ -2 & t-1 \end{pmatrix} = t^2 - 5t + 6t^2 - 5t + 6 = 0 \Longleftrightarrow t_1 = 2;\ t_2 = 3$$

O teorema acima nos mostra que t_1 e t_2 são os únicos autovalores de T. Para determinarmos os autovetores de T associados a t_1, devemos resolver o sistema

$$\begin{pmatrix} t_1 - 4 & 1 \\ -2 & t_1 - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ou seja,

$$\begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que equivale a equação $-2x_1 + x_2 = 0$ assim, o autoespaço de T associado a t_1 é $\{(x, 2x); x \in \mathbb{R} \neq 0\}$. Agora vamos determinar, da mesma forma, os autovetores associados ao autovalor $t_2 = 3$ Assim:

$$\begin{pmatrix} t_2 - 4 & 1 \\ -2 & t_2 - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ou seja,

$$\begin{pmatrix} -2 & 1 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

equivale à equação linear $-x_1 + x_2 = 0$ assim, o autoespaço de T associado a t_2 é $\{(x, x); x \in \mathbb{R} \neq 0\}$.

O teorema a seguir é um dos mais importantes teoremas da Álgebra Linear, o chamado Teorema de Cayley - Hamilton.

Teorema 12. Seja $A \in M_n$ e seja $P_A(t)$ o polinômio característico de A. Então, $P_A = 0$, onde 0 é a matriz nula de $M(n)$.

Uma consequência imediata do Teorema de Cayley - Hamilton é que a potência A^n, de uma matriz $A \in M_n$, pode ser escrita como combinação linear das potências de A com expoentes menores do que n, pois se $P_A(t) = t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0$, então $P_A(A) = 0$, o que equivale a

$$A^n = -b_{n-1}A^{n-1} - \cdots - b_1 A + b_0 I_n$$

Exemplo 37. Considere a matriz $A = \begin{pmatrix} 1 & 3 \\ -1 & 0 \end{pmatrix}$ seu polinômio característico será:

$$P_A(t) = \det(tI_n - A) = \det \begin{bmatrix} t-1 & -3 \\ 1 & t-0 \end{bmatrix} = t^2 - t + 3.$$

Pelo Teorema de Cayley - Hamilton, $P_A(A) = 0$. Vamos calcular A^3.

$A^2 - A + 3I_2 = 0$, ou seja, $A^2 = A - 3I_2$

$A^3 = AA^2 = A(A - 3I_2) = A^2 - 3A = -2A - 3I_2$.

Para A^4, temos $A^4 = AA^3 = A(-2A - 3I_2) = -2A^2 - 3A = -2(A - 3I_2) - 3A = -5A + 6I_2$.

Este procedimento mostra que, em geral, se $A \in M_2$, então para todo $m \in N$, a matriz A^m se escreve como combinação linear de I_2 e A.

Retornando ao nosso objetivo inicial, dado um operador linear $T : V \to V$, queremos obter, se possível, uma base α de V na qual a matriz $[T]_\alpha^\alpha$ seja uma matriz diagonal.

Teorema 13. Um operador linear $T : V \to V$ admite uma base β em relação à qual a matriz $[T]_\beta^\beta$ é diagonal se, e somente se, essa base β for formada por autovetores de T.

Demonstração. Suponhamos que $\beta = \{v_1, v_2, \cdots, v_n\}$ é uma base de V tal que $[T]_\beta^\beta$ é diagonal, digamos

$$[T]_\beta^\beta = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n. \end{pmatrix} \tag{5.9}$$

para cada $1 \leq j \leq n$; $T(v_j) = 0v_1 + \cdots + 0v_{j-1} + \alpha_j v_j + 0v_{j+1} + \cdots + 0v_n = \alpha_j v_j$, segue que α_j é um autovalor de T e v_j é autovetor de T associado a α_j. Portanto, β é uma base formada de autovetores de T.

Suponhamos agora que $\beta = \{u_1, u_2, \cdots, u_n\}$ é uma base de V formada por autovetores de T. Existem, então, números reais $b_1, b_2, \cdots, b_n$ tais que, para cada $1 \leq j \leq n$, $T(u_j) = b_j u_j$. Observamos que os b_{js} não são necessariamente todos distintos. Pela definição de $[T]_\beta^\beta$ e pelo fato da base ser autovetores,

$$[T]_\beta^\beta = \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix} \tag{5.10}$$

ou seja, $[T]_\beta^\beta$ é uma matriz diagonal. □

Na demonstrção do teorema acima fica claro que, se o operador linear T tem uma representação por uma matriz diagonal $[T]_\beta^\beta$, então as entradas da diagonal principal de $[T]_\beta^\beta$ são dadas pelos autovalores de T. Mais ainda, a ordem em que os autovalores aparecem na diagonal principal da matriz é a mesma em que seus respectivos autovetores são dados na base β. O polinõmio caracteristico de T tem a forma $P_T(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$, onde os números reais $\lambda_1, \cdots, \lambda_n$ são todos os autovalores de T.

Exemplo 38. O operador linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ dado por $T(x,y) = (4x - y, 2x + y)$ é diagonalizável. De fato, vimos anteriormente que os autovalores de T são 2 e 3 e os autovetores associados aos autovalores são $\{(x,2x); x \in \mathbb{R}, x \neq 0\}$ e $\{(x,x); x \in \mathbb{R}, x \neq 0\}$ respectivamente, então uma representação diagonal para T é dada por:

$[T]_\beta^\beta = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, onde $\beta = \{(1,2),(1,1)\}$.

Uma outra representação diagonal de T é dada por:

$[T]_\beta^\beta = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$, sendo $\beta = \{(1,1),(1,2)\}$.

Definimos a multiplicidade algébrica de um autovalor λ de T, como sendo sua multiplicidade como raiz do polinômio característico de T. A multiplicidade geométrica de um autovalor λ de T é, por definição, a dimensão do espaço gerado (ou autoespaço) pelos autovetores associados a λ, o qual denotaremos por $V(\lambda)$.

Teorema 14. Sejam U um espaço vetorial de dimensão finita e $T \in \mathscr{L}(U)$: Então, T é diagonalizável se, e somente se, os seus autovalores $\lambda_1, \cdots, \lambda_n$ forem tais que

$$U = V(\lambda_1) \oplus \cdots \oplus V(\lambda_n)$$

Demonstração. Se $U = V(\lambda_1) \oplus \cdots \oplus V(\lambda_n)$, então podemos formar uma base B de U formada por bases B_j de $V(\lambda_j)$; $j = 1, \cdots, n$. Como cada elemento de B_j é um autovetor de T, segue por definição que T é diagonalizável.Reciprocamente, se T for diagonalizável existe uma base B de U formada por autovetores de T. Como cada autovetor está associado a algum autovalor de T, vemos que cada elemento de B está contido em algum $V(\lambda_j)$. Desta forma, a soma de todos os subespaços próprios de T contém B e, portanto, é o próprio U. Logo a soma é direta, ou seja, $U = V(\lambda_1) \oplus \cdots \oplus V(\lambda_n)$. □

O seguinte teorema nos fornece uma forma de verificar se T é diagonalizável, conhecendo as multiplicidades algébricas e geométricas de seus autovalores.

Teorema 15. Sejam U um espaço vetorial de dimensão finita e $T \in \mathscr{L}(U)$. Então T é diagonalizável se e somente se ambas condições forem verificadas

1. para cada autovalor de T as suas multiplicidades algébrica e geométrica são iguais,
2. a soma das multiplicidades geométricas de todos os autovalores de T coincide com a dimensão de U.

Corolário 6. Sejam U um espaço vetorial de dimensão n e $T \in \mathscr{L}(U)$: Se

$$p_T(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$$

onde $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$ são distintos entre si então T é diagonalizável.

Demonstração. Como os autovalores de T são dois a dois distintos, vê-se que as raízes de $p_T(\lambda)$, são todas simples, isto é, têm multiplicidade um. Desta forma, se λ é um autovalor de T então a sua multiplicidade algébrica é um. Como $dimV(\lambda) \geq 1$, segue-se que a multiplicidade geométrica de λ é um, ou seja, igual à sua multiplicidade algébrica. □

Exemplo 39. Vamos verificar se $T : \mathbb{R}^4 \to \mathbb{R}^4$ dada por $T(x,y,z,t) = (x+y, y, 2z+t, 2z+t)$ é diagonalizável. A matriz T em relação a base canônica é dada por:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

e seu polinômio característico é:

$$p_T(\lambda) = det \begin{pmatrix} \lambda-1 & 1 & 0 & 0 \\ 0 & \lambda-1 & 0 & 0 \\ 0 & 0 & \lambda-2 & 1 \\ 0 & 0 & 2 & \lambda-1 \end{pmatrix} = \lambda(\lambda-3)(\lambda-1)^2 = 0$$

onde, $\lambda = 0$ ou $\lambda = 3$ ou $\lambda = 1$ são as raízes. Vamos procurar os autovetores associados aos autovalores encontrados.

Para $\lambda = 0$, temos que $v = (x,y,z,t)$ é autovetor associado a $\lambda = 0$ se

$$T(v) = \lambda v \Longleftrightarrow (x+y,y,2z+t,2z+t) = (0,0,0,0)$$

ou seja, $(x,y,z,t) = z(0,0,1,-2)$, com $z \in \mathbb{R}^*$. Logo, $[(0,0,1,-2)]$ é o espaço gerado pelo autovetor associado a $\lambda = 0$.

Para $\lambda = 3$, $T(v) = \lambda v \Longleftrightarrow (x+y,y,2z+t,2z+t) = (3x,3y,3z,3t)$, ou seja, $v = (x,y,z,t) = t(0,0,1,1)$, com $t \in \mathbb{R}^*$. Logo, $[(0,0,1,1)]$ é o espaço gerado pelo autovetor associado a $\lambda = 3$.

Para $\lambda = 1$, $T(v) = \lambda v \Longleftrightarrow (x+y,y,2z+t,2z+t) = (x,y,z,t)$ e $v = (x,y,z,t) = x(1,0,0,0)$, com $x \in \mathbb{R}^*$. Logo, $[(1,0,0,0)]$ é o espaço gerado pelo autovetor associado a $\lambda = 1$. Como a multiplicidade algébrica do autovalor 1 é dois e a sua multiplicidade geométrica é um, vimos que T não é diagonalizável.

Exemplo 40. Vamos verificar se $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $T(x,y,z) = (x+z, y+z, x+y+2z)$ é diagonalizável. Com relação à base canônica, a matriz de T é dada por $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Assim,

$$p_T(\lambda) = det \begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix} = \lambda(1-\lambda)(\lambda-3).$$

Desta forma, vemos que $p_T(\lambda)$ apresenta todas as raízes reais e simples e, pelo corolário 6, segue-se que T é diagonalizável.

Vamos encontrar uma base de autovetores para este operador. Para o autovalor $\lambda = 0$, temos que resolver $T(x,y,z) = (0,0,0)$, o que nos fornece o autovetor $u = (1,1,-1)$. Para o autovalor $\lambda = 1$, temos o autovetor $v = (1,-1,0)$ e para o autovalor $\lambda = 3$ temos o autovetor $w = (1,1,2)$.

A matriz de T com relação à base formada por u, v e w é dada por

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

O próximo exemplo será utilizado nas aplicações envolvendo cônicas.

Exemplo 41. Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ cuja matriz com relação a alguma base é dada por

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

Vamos mostrar que T é diagonalizável. O polinômio característico de T é dado por

$$p_T(\lambda) = \lambda^2 - (a+c)\lambda + ac - b^2$$

Vemos que $p_T(\lambda)$ apresenta duas raízes reais simples, isto é, com multiplicidade um, se, e somente se, o discriminante $(a+c)^2-4(ac-b^2)$ for positivo.
Temos que $(a+c)^2-4(ac-b^2)=a^2+c^2-2ac+4b^2=(a-c)^2+4b^2>0$ se, e somente se, $a\neq c$ ou $b\neq 0$. Portanto, se $a\neq c$ ou $b\neq 0$ as multiplicidades algébrica e geométrica de cada um dos autovalores de T (as raízes de $p_T(\lambda)$) coincidem e, portanto, T é diagonalizável.

Se $a=c$ e $b=0$ então vê-se claramente que T é diagonalizável pois, neste caso, A é diagonal.

5.2 Matrizes Diagonalizáveis

Toda matriz $A=(a_{ij})\in M_{m\times n}$ define uma transformação linear $T_A:\mathbb{R}^n\to\mathbb{R}^m$ definida por $T_A(v)=(a_{11}x_1+\cdots+\cdots+a_{1n}x_n,\cdots,a_{n1}x_1+\cdots+a_{nn}x_n)$, onde $x_1,\cdots,x_n$ são coordenadas de v relativamente à base $\alpha=\{v_1,v_2,\ldots,v_n\}$. Em particular, se A é uma matriz quadrada de ordem n, então A define um operador linear T_A em $\mathbb{R}^n$. Dizemos que a matriz A é diagonalizável, se o operador linear T_A for diagonalizável. Desta forma, existe uma representação diagonal D, onde $D=[T_A]_\beta^\beta$, para o operador T_A com relação a alguma base β de V. Como $[T]_\alpha^\alpha=A$ onde α denota a base canônica de $\mathbb{R}^n$, assim, se $P=[I_{\mathbb{R}^n}]_\alpha^\beta$, temos:

$$D=P^{-1}AP.$$

Teorema 16. Uma matriz $A\in M_{m\times n}$ é diagonalizável se, e somente se, existe uma matriz P invertível de ordem n tal que $P^{-1}AP$ é uma matriz diagonal.

Demonstração. Consideremos $\beta=\{v_1,v_2,\cdots,v_n\}$, onde v_j é o vetor formado pela j-ésima coluna de P. Seja α a base canônica de $\mathbb{R}^n$.

Assim,

$$[T_A]_\beta^\beta=[I_{\mathbb{R}^n}]_\beta^\alpha[T_A]_\alpha^\alpha[I_{\mathbb{R}^n}]_\alpha^\beta, \tag{5.11}$$

ou, equivalentemente,

$$[T_A]_\beta^\beta=P^{-1}AP \tag{5.12}$$

já que $[I_{\mathbb{R}^n}]_\alpha^\beta=P$ pela maneira como β foi tomada. Como $P^{-1}AP$ é uma matriz diagonal, segue-se que $[T_A]_\beta^\beta$ é uma matriz diagonal. Portanto, T_A é diagonalizável e. assim, A também é. $\square$

Exemplo 42. Vamos verificar se a matriz $A=\begin{pmatrix}1&0&2\\0&1&3\\0&0&-1\end{pmatrix}$ é diagonalizável.

Seja α a base canônica no $\mathbb{R}^3$. Então:

$$p_{[T_A]_\alpha^\alpha}=p_A(t)=\det\begin{pmatrix}t-1&0&-2\\0&t-1&-3\\0&0&t+1\end{pmatrix}=0$$

para $t = 1$ ou $t = -1$. O autoespaço associado ao autovalor $t = 1$ é o conjunto solução do sistema linear:

$$\begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & -3 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ou seja, é o conjunto $\{(x, y, 0); x, y \in \mathbb{R}\}$. O autoespaço associado ao autovalor $t = -1$ é o conjunto solução do sistema linear:

$$\begin{pmatrix} -2 & 0 & -2 \\ 0 & -2 & -3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ou seja, é o conjunto $\{(-z, \frac{-3}{2}z, z); z \in \mathbb{R}\}$. Se considerarmos $\beta = \{(1,1,0), (1,0,0), (1, \frac{3}{2}, -1)\}$ temos que β é uma base de $\mathbb{R}^3$ formada por autovetores de T_A. Assim T_A é diagonalizável. A matriz

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & \frac{3}{2} \\ 0 & 0 & -1 \end{pmatrix}$$

é uma matriz que diagonaliza A, no sentido que $D = P^{-1}AP$. Assim,

$$D = \begin{pmatrix} 0 & 1 & \dfrac{3}{2} \\ 1 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & \dfrac{3}{2} \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Observamo que o cálculo da potência m da matriz A, isto é, A^m, fica bastante simplificado quando A é diagonalizável. De fato, se $A \in M_{n)}$ e se $P \in M_n$ é invertível, então é fácil verificar que:

$$(P^{-1}AP)^m = P^{-1}A^mP$$

Logo, se A é diagonalizável e se $P^{-1}AP = D$ é uma matriz diagonal, temos que:

$$D^m = P^{-1}A^mP,$$

ou, equivalente,

$$A^m = PD^mP^{-1}.$$

Vejamos um exemplo deste cálculo.

Exemplo 43. Vamos calcular a matriz A^{50} onde $A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$.

Vamos verificar que A é diagonalizável e encontrar uma matriz P que diagonaliza A. Para isso, façamos

$$\det(tI - A) = \det \begin{pmatrix} t-1 & -2 \\ 0 & t+1 \end{pmatrix} = (t-1)(t+1)$$

que se anula para $t = 1$ ou $t = -1$. Como as raízes são simples, temos que A será diagonalizável e a matriz diagonal correspondente será $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Temos que $v_1 = (1,0)$ é um autovetor para $t = 1$ e $v_2 = (1,-1)$ é um autovetor para $t = -1$. Assim, $P = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$ e, como $D^{50} = I_2$ segue-se que

$$A^{50} = P^{-1}D^{50}P = P^{-1}I_2P = I_2.$$

5.3 Teorema Espectral para operadores simétricos

Vimos que $T : V \to V$ é um operador diagonalizável se, e somente se, existe uma base de V formada por autovetores de T. Agora, veremos que se V é um espaço com produto interno e se $T : V \to V$ é um operador simétrico, então existe uma base ortonormal de V formada por autovetores de T. Em particular, todo operador simétrico é diagonalizável. Este resultado é conhecido como Teorema Espectral e é um dos resultados mais importantes da Algebra linear.

Todos os resultados que foram provados até o momento que envolvem sistemas lineares e determinantes são válidos sobre um corpo arbitrário K. Utilizamos $K = \mathbb{R}$. Neste ponto precisaremos considerar $K = \mathbb{C}$.

Dado um operador linear $T : \mathbb{R}^n \to \mathbb{R}^n$ podemos estendê-lo a um operador $T_C : \mathbb{C}^n \to \mathbb{C}^n$ do seguinte modo: se $z = x + iy \in \mathbb{C}^n$, onde $x, y \in \mathbb{R}^n$, define-se $T_C(z) = T(x) + iT(y)$. Os polinômios característicos de T e de T_C coincidem, mas T_C pode possuir mais autovalores e autovetores do que T.

Proposição 29. Seja V um espaço vetorial de dimensão finita sobre $\mathbb{R}$. Se $T : V \to V$ é um operador simétrico e α uma base de V, então, todas as raízes do polinômio característico $P_{[T]_\alpha^\alpha}$ em $\mathbb{C}$ são números reais.

Teorema 17. (Teorema Espectral) Seja V um espaço vetorial de dimensão finita sobre $\mathbb{R}$. Se $T : V \to V$ é um operador simétrico, então existe uma base ortonormal β de V tal que $[T]_\beta^\beta$ é diagonal.

Demonstração. A prova será feita por indução sobre a dimensão de V. Denotaremos a matriz $[T]_{\alpha}^{\alpha}$ por A. Se $\dim V = 1$, o resultado é óbvio. Suponhamos que $n \geq 1$ e que o resultado é valido para espaços de dimensão n. Seja V um espaço vetorial tal que $\dim V = n+1$. Seja α uma base de V e seja c uma raiz complexa do polinômio P_A.Pela proposição acima $c \in \mathbb{R}$. Portanto é um autovalor de T. Seja v um autovetor unitário de T associado a c. Consideremos os subespaços.

$$W = \{w \in V; < w, v >= 0\}$$

Note que $W = G(v)^{\perp}$. Afirmamos que $T(W) \subset W$. De fato, seja $w \in W$. Como T é um operador simétrico, temos que: $< T(w), v >=< w, T(v) >=< w, cv >= c < w, v >= c0 = 0$, donde $T(w) \in W$. Assim podemos considerar o operador restrição.

$$S = T|W \in \mathscr{L}(W,W)$$

que é também um operador simétrico. Além disso, como $\dim G(v) = 1$, segue que $\dim W = n$. Assim podemos aplicar a hipótese de indução ao operador S para garantir a existência de uma base ortonormal $\{v_1, v_2, \cdots, v_n\}$ de W formada por autovetores de S. Consequentemente, $\beta = \{v, v_1, \cdots, v_n\}$ é uma base ortonormal de V formada por autovetores de T. Daí, $[T]_{\beta}^{\beta}$ é diagonal.

□

Teorema 18. Se $A \in M_n$ é simétrica, então existe uma matriz ortogonal $P \in M_n$ tal que $P^{-1}AP = (P^tAP)$ é diagonal.

Demonstração. Seja $A \in M_n$ uma matriz simétrica. Então o operador $T_A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)$ também é simétrico. Pelo Teorema Espectral, existe uma base ortonormal β de $\mathbb{R}^n$ tal que $[T_A]_{\beta}^{\beta} = D$ é diagonal. Se α é base canônica de $\mathbb{R}^n$, então:

$$D = [T_A]_{\beta}^{\beta} = [I_{\mathbb{R}^n}]_{\beta}^{\alpha} [T_A]_{\alpha}^{\alpha} [I_{\mathbb{R}^n}]_{\alpha}^{\beta} = P^{-1}AP$$

sendo $P = [I_{\mathbb{R}^n}]_{\alpha}^{\beta}$. Como α e β são bases ortonormais, segue que P é uma matriz ortonormal, ou seja, $P^{-1} = P^t$ □

CAPÍTULO

6

CÔNICAS

De maneira geral, uma cônica é o conjunto de pontos $P = (x, y)$ no plano tais que

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

onde $A, B, C, D, E, F \in \mathbb{R}$, com A, B, C não simultaneamente nulos. Esse capítulo se propõe ao estudo das cônicas não degeneradas, os elementos principais e características gerais. Também haverá uma breve discussão sobre cônicas degeneradas. A bibliografia usada para este capítulo pode ser encontrada em (FRENSEL K.; DELGADO, 2011), (HEFEZ A.; SOUZA FERNANDES, 2012) e (GOMEZ KATIA R. FRENSEL, 2017).

Elipse

Definição 30. Uma elipse, $\mathscr{E}$, de focos F_1 e F_2, é o conjunto do plano que consiste de todos os pontos P cuja soma das distâncias a F_1 e F_2 é igual a uma constante $2a > 0$, maior do que a distância entre os focos $2c \geqslant 0$. Ou seja.

$$\mathscr{E} = \{Pd(P, F_1) + d(P, F_2) = 2a\}$$

$$0 \leq c < a; d(F_1, F_2) = 2c$$

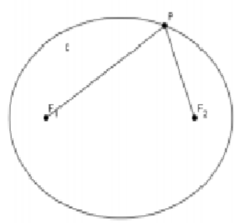

Figura 7 – Elipse

Os pontos F_1 e F_2 chamam-se Focos e a medida do segmento F_1F_2 é chamada de distância focal que será representada por $2c$. A reta a qual os focos pertencem é um eixo de simetria da curva que intercepta a elipse nos pontos A_1 e A_2 (que chamam-se de vértices da elipse). O segmento A_1A_2 é chamado eixo maior da elipse. No ponto médio dos focos se encontra o centro da elipse e passando uma reta perpendicular por ele, tem-se outro eixo de simetria da curva. Esse eixo intercepta a elipse nos pontos B_1 e B_2. O segmento determinado por esses pontos é chamado eixo menor e sua medida será representado por $2b$. Do triângulo retângulo formado decorre a relação $a^2 = b^2 + c^2$ e, portanto, sempre se tem $a > b$.

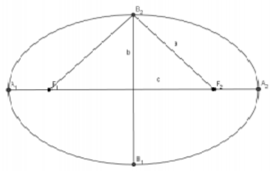

Figura 8 – Elipse

Chama-se de excentricidade ao quociente entre as distâncias entre focos e a distância entre vértices, isto é

$$e = \frac{d(F_1,F_2)}{d(A_1,A_2}.$$

Observa-se que, da definição, para a elipse $e < 1$. Considere A_1 e A_2 pontos onde a elipse intercepta o eixo maior. Observa-se que $d(A_1,A_2) = 2a$. De fato, sejam

$$d(A_1,F_1) = x, \;\; d(F_1,F_2) = 2c \;\text{ e }\; d(F_2,A_2) = y.$$

Como o A_1 está na elipse tem-se que, da equação,

$$d(A_1,F_1) + d(A_1,F_2q) = 2a.$$

Utilizando $d(A_1,F_2) = d(A_1,F_1) + d(F_1,F_2)$ tira-se que $2x + 2c = 2a$. Analogamente, tem-se $2y + 2c = 2a$. Subtraindo estas duas equações, vê-se que $x = y$, ou seja, $d(A_1,F_1) = d(A_2,F_2)$. Logo,

$$d(A_1,A_2) = d(A_1,F_1) + d(F_1,F_2) + d(F_2,A_2) = x + 2c + y = 2a.$$

Exemplo 44. Os vértices de uma elipse são os pontos $(4,0)$ e $(-4,0)$, e seus focos são os pontos $(3,0)$ e $(-3,0)$. Vamos determinar a equação da elipse. Como $F_1 = (-3,0)$ e $F_2 = (3,0)$, a reta focal é o eixo $-OX$, $A_1 = (-4,0)$ e $A_2 = (4,0)$ são os vértices sobre a reta focal ℓ. Então, $C = \dfrac{f_1 + f_2}{2} = \dfrac{A_1 + A_2}{2} = (0,0)$ é o centro da elipse, $a = d(C,A_1) = d(C,A_2) = 4$, $c = d(C,F_1) = d(C,F_2) = 3$ e $b = \sqrt{a^2 - c^2} = \sqrt{4^2 - 3^2} = \sqrt{7}$ Logo, a equação da elipse é $\mathscr{E} : \dfrac{x^2}{16} + \dfrac{y^2}{7} = 1$

Exemplo 45. A equação de uma elipse é

$$\mathscr{E} : x^2 + 4y^2 + 2x - 12y + 6 = 0.$$

Vamos determinar a equação da elipse $\mathscr{E}$ na forma canônica, o seu centro, os seus vértices, os seus focos e a sua excentricidade.

Completando os quadrados na equação de E, temos:

$$\mathscr{E} : (x^2 + 2x) + 4(y^2 - 3y) = -6$$

$$\mathscr{E} : (x^2 + 2x + 1)4(y^2 - 3y + \frac{9}{4}) = -6 + 1 + 4 \cdot \frac{9}{4} = 4$$

$$\mathscr{E} : (x+1)^2 + 4(y - \frac{3}{2})^2 = 4$$

$$\mathscr{E} : \frac{(x+1)^2}{4} + (y - \frac{3}{2})^2 = 1.$$

sendo esta última equação a forma canônica de $\mathscr{E}$. Dessa equação obtemos que o centro da elipse é $C = (-1, \frac{3}{2})$, $a = 2$, $b = 1$ e, portanto, $c^2 = a^2 - b^2 = 2^2 - 1^2 = 3$, ou seja $c = \sqrt{3}$. A reta focal de $\mathscr{E}$ é $\ell : y = \frac{3}{2}$, paralela ao eixo OX, e a reta não-focal é a reta vertical $\ell' : x = -1$, paralela ao eixo $-OY$. Os focos da elipse são $F_1 = -1 - \sqrt{3}, \frac{3}{2})$ e $F_2 = (-1 + \sqrt{3}, \frac{3}{2})$; os vértices sobre a reta focal são $A_1 = (-3, \frac{3}{2})$ e $A_2 = (1, \frac{3}{2})$ e os vértices sobre a reta não-focal são $B_1 = (-1, \frac{1}{2})$ e $B_2 = (-1, \frac{5}{2})4$ e a excentricidade de $\mathscr{E}$ é $e = \frac{\sqrt{3}}{2}$.

Exemplo 46. Vamos determinar se a equação $25x^2 + 9y^2 - 225 = 0$ representa uma elipse ou uma elipse degenerada. Caso seja uma elipse, vamos determinar seus principais elementos.

Como $25x^2 + 9y^2 = 225$, dividindo por 225, obtemos a equação $\frac{x^2}{9} + \frac{y^2}{25} = 1$representa uma elipse com:

- $a = 5, b = 3$ e $c == 4$.
- centro: $C = (0, 0)$.
- reta focal: $\ell = eixo - OY : x = 0$.
- reta não-focal: $\ell' = eixo - OX : y = 0$.
- vértices sobre a reta focal: $A_1 = (0, -5)$ e $A_2 = (0, 5)$.
- vértices sobre a reta não-focal: $B_1 = (-3, 0)$ e $B_2 = (3, 0)$.
- focos: $F_1 = (0, -4)$ e $F_2 = (0, 4)$.

Hipérbole

Definição 31. Uma hipérbole, $\mathscr{H}$, de focos F_1 e F_2, é o conjunto do plano que consiste de todos os pontos P tais que o módulo da diferença das distâncias de F_1 a F_2 é igual a uma constante $2a > 0$, menor do que a distância entre os focos $2c \geqslant 0$.

$$\mathscr{H} = \{P/d(P,F_1) - d(P,F_2)| = 2a\}$$

$$0 \leqslant a < c; d(F-1,F_2) = 2c$$

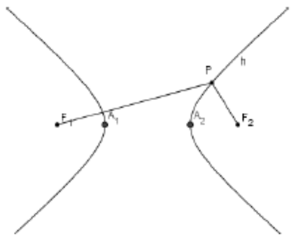

Figura 9 – Hipérbole

Os pontos F_1 e F_2 chamam-se Focos e a medida do segmento F_1F_2 é chamada de distância focal, que será representada por $2c$. A reta na qual os focos estão localizados é um eixo de simetria da curva que intercepta a hipérbole nos pontos A_1 e A_2, chamados de vértices. A distância entre os vértices será representada por $2a$. O segmento A_1A_2 é chamado eixo real da hipérbole. No ponto médio dos focos encontra-se o centro da hipérbole, e passando uma reta perpendicular por ele, tem-se o outro eixo de simetria da curva. Esse eixo intercepta a hipérbole nos pontos B_1 e B_2. O segmento determinado por esses pontos é chamado eixo conjugado (ou imaginário) e será representado com medida $2b$.

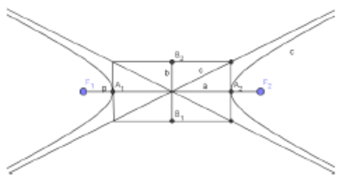

Figura 10 – Hipérbole

Do triângulo retângulo formado obtém-se a relação $c^2 = a^2 + b^2$. Chama-se de excentricidade o quociente entre as distâncias focais e as distâncias entres os vértices, isto é

$$e = \frac{d(F_1,F_2)}{d(A_1,A_2)}.$$

Observa-se que, da definição, para a hipérbole $e > 1$. Considere A_1 e A_2 pontos onde a hipérbole intercepta o eixo real. Observa-se que $d(A_1,A_2) = 2a$. De fato, sejam

$$d(A_1,F_1) = x,\ \ d(F_1,F_2) = 2c \ \text{ e } \ d(F_2,A_2) = y.$$

Como o A_1 está na hipérbole, tem-se da equação,

$$d(A_1,F_1) - d(A_2,F_2) = 2a.$$

Utilizando que $d(F_1,F_2) = d(F_1,A_1) + d(A_1,A_2)$ tem-se que

$$2c = 2x + 2a.$$

Analogamente prova-se que $2c = 2y - 2a$. Subtraindo estas duas equações se vê que $x = y$, ou seja, $d(F_1,A_1) = d(A_2,F_2)$. Logo,

$$d(A_1,A_2) = d(F_1,F_2) - [d(F_1,A_1) + d(A_2,F_2)] = 2c - [x - y] = 2c - 2x = 2a.$$

Exemplo 47. Vamos determinar a equação da hipérbole equilátera com focos nos pontos $(-\sqrt{8},0)$ e $(\sqrt{8},0)$. Como $F_1 = (-\sqrt{8},0)$ e $F_2 = (\sqrt{8},0)$, temos que o centro da hipérbole é $C = \dfrac{F_1 + f_2}{2} = (0,0)$ e a reta focal é o eixo $-OX$. Sendo a hipérbole equilátera, temos $a = b$. Como $c = \sqrt{8}$ e $c^2 + a^2 = b^2$, obtemos que $a^2 = 4$. Logo, $a = b = 2$ e $\mathscr{H} : \dfrac{x^2}{4} - \dfrac{y^2}{4} = 1$, é a equação da hipérbole. Além disso, $A_1 = (-2,0)$ e $A_2 = (2,0)$ são os vértices, $B_1 = (0,-2)$ e $B_2 = (0,2)$ são os vértices imaginários e $x = \pm y$ são as assíntotas da hipérbole $\mathscr{H}$.

Exemplo 48. Vamos mostrar que a excentricidade de qualquer hipérbole equilátera é $\sqrt{2}$.

Como $a = b$ e $c^2 = a^2 + b^2$, temos que $c^2 = 2a^2$, ou seja, $c = \sqrt{2a}$. Logo, $e = \dfrac{c}{a} = \dfrac{\sqrt{2a}}{a} = \sqrt{2}$.

Exemplo 49. Vamos determinar se a equação $9x^2 - 25y^2 - 225 = 0$ representa uma hipérbole ou uma hipérbole degenerada. Caso seja uma hipérbole, vamos determinar seus principais elementos.

Como $9x^2 - 25y^2 = 225$, dividindo por 225, encontramos a equação $\dfrac{x^2}{25} - \dfrac{y^2}{9} = 1$, que representa uma hipérbole com

- $a = 5, b = 3$ e $c = \sqrt{34}$.
- centro: $C = (0,0)$.
- reta focal: $\ell = eixo - OX : y = 0$.
- reta não-focal: $\ell = eixo - OY : x = 0$.

- vértices: $A_1 = (-5,0)$ e $A_2 = (5,0)$.
- vértices imaginários (na reta não-focal): $B_1 = (0,-3)$ e $B_2 = (0,3)$.
- focos: $F_1 = (-\sqrt{34},0)$ e $F_2 = (\sqrt{34},0)$.
- assíntotas: $y = \pm\frac{3}{5}x$, ou seja $3x \pm 5y = 0$.

Parábola

Definição 32. Sejam $\mathscr{L}$ uma reta no plano e F um ponto no plano não pertencente a $\mathscr{L}$. A parábola $\mathscr{P}$ de diretriz $\mathscr{L}$ e foco F é o conjunto que consiste de todos os pontos P do plano que são equidistantes do ponto F e da reta $\mathscr{L}$.

$$\mathscr{P} = \{P/d(P,F) = d(P,\mathscr{L})\}$$

Toma-se também um número positivo 4p, chamado de parâmetro que será a distância entre o foco e a diretriz. Na parábola se encontra o vértice no ponto médio do foco e da interseção da diretriz com o eixo da parábola.

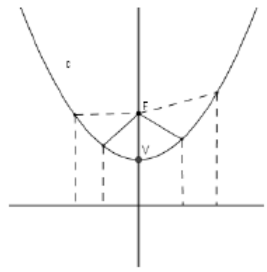

Figura 11 – Parábola

Exemplo 50. Considere uma parábola que no sistema catesiano tem seu foco no eixo das abscissas, isto é, $F(p,0)$ e a diretriz de equação $x = -p$. Deduz-se a equação reduzida da parábola nesse caso partindo da definição:

$$d(P,F) = d(P,-q)$$

$$\sqrt{(x-p)^2 + y^2} = \|x+p\|$$

$$(x-p)^2 + y^2 = (x+p)^2$$

$$x^2 - 2px + y^2 = x^2 + 2px + p^2$$

$$-2px + y^2 = 2px$$

$$y^2 = 4px.$$

Nessas condições, se F está a direita de V, a equação da parábola será da forma

$$y^2 = 4px,$$

e, se F está à esquerda de V, a equação da parábola será da forma

$$y^2 = -4px,$$

De modo análogo, obtém-se a equação da parábola quando o vértice está na origem do sistema e o foco no eixo das ordenadas. Nesse caso tem-se

$$x^2 = 4py,$$

quando F está acima de V e,

$$x^2 = -4py,$$

quando F está abaixo de V.

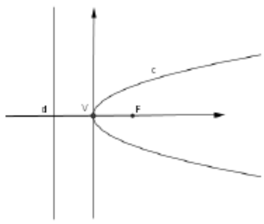

Figura 12 – $y^2 = 4px$

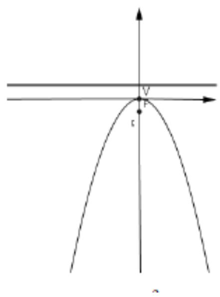

Figura 13 – $x^2 = -4px$

Se a parábola não estiver com seu vértice localizado na origem do sistema mas se o eixo que contém o foco for paralelo a abscissa tem-se a equação

$$(y - y_0)^2 = 2p(x - x_0).$$

E se o eixo que contém o foco estiver paralelo ao eixo da ordenada tem-se a equação

$$(x - x_0)^2 = 2p(y - y_0)$$

Exemplo 51. Vamos determinar a equação da parábola $\mathscr{P}$ com vértice V na origem, cujo foco é $F=(3,0)$. Temos $p=d(V,F)=3$ e reta focal $=$ eixo $-OX$. Como o foco F está à direita do vértice, temos que a diretriz é $\mathscr{L}: x=-3$ e a equação da parábola é $\mathscr{P}: y^2=12x$.

Exemplo 52. Um círculo C com centro no ponto $C=(4,-1)$ passa pelo foco F da parábola $\mathscr{P}: x^2=-16y$. Mostre que C é tangente à diretriz $\mathscr{L}$ de $\mathscr{P}$.

A reta focal da parábola $\mathscr{P}$ é o eixo $-OY$, o vértice é a origem, e o foco está abaixo da diretriz. Então, $F=(0,-4)$ e $\mathscr{L}: y=4$, pois $4p=16$. A equação do círculo é $C:(x-4)^2+(y+1)^2=r^2$. Sendo $F=(0,-4)\in C$, temos $r=5$. Logo $\mathscr{L}$ é tangente a C, pois $d(C,\mathscr{L})=d((4,-1),\mathscr{L})=|-1-4|=5=$ raio de C.

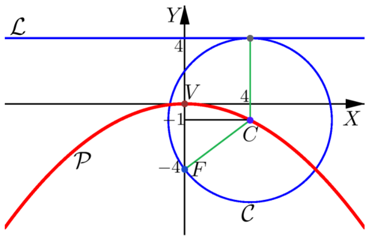

Figura 14 – Parabola $\mathscr{P}$ e círculo C

Exemplo 53. Vamos verificar se a equação $x^2-8y=0$ representa uma parábola ou uma parábola degenerada. Caso seja uma parábola, vamos determinar seus principais elementos.

Como $x^2=8y$, a equação representa uma parábola, com

- vértice: $V=(0,0)$.
- reta focal $=$ eixo $-OY: x=0$.
- parâmetro: $p=2$.
- foco: $F=(0,2)$, acima da diretriz.
- diretriz: $\mathscr{L}: y=-2$.

Cônicas Degeneradas

Como se vê, a elipse, a hipérbole e a parábola têm equações que serão representadas na forma $Ax^2+Bxy+Cy^2+Dx+Ey+F=0$. Entretanto, nem toda equação dessa forma, representa uma das curvas já citadas. Há equações que representam uma única solução, ou seja, um ponto; ou ainda, um par de retas. Essas representações são as cônicas degeneradas, que se verá a seguir.

1. Par de retas: o conjunto solução de uma equação do segundo grau que pode ser fatorada na forma $(a_1x+b_1y+c_1)(a_2x+b_2y+c_2)=0$, onde a_1,a_2,b_1,b_2,c_1,c_2 são reais e $a_1\neq 0$ ou $b_1\neq 0, a_2\neq 0$ ou $b_2\neq 0$ representa um par de retas, podendo ser paralelas, concorrentes ou mesmo coincidentes.

Exemplo 54. A equação $x^2+2xy+y^2-1=0$ representa um par de retas paralelas, pois pode ser fatorada como $(x+y+1)(x+y-1)=0$.

2. Um ponto: o conjunto solução de uma equação do segundo grau que pode ser escrito na forma $k_1(x-x_0)^2+(y-y_0)^2=0$ com $k_1 \neq 0$ e $k_2 \neq 0$ representa um ponto, pois só o ponto (x_0,y_0) satisfaz essa equação.

Exemplo 55. A equação $x^2+y^2=0$ representa um ponto, o $(0,0)$.

6.1 Reconhecimento de Cônicas

Agora mostraremos como por meio de Teorema Espectral é possível fazer o reconhecimento de cônicas. Considere a equageral do segundo grau nas duas variáveis x e y.

$$ax^2+bxy+cy^2+dx+ey+f=0$$

onde a,b,c,d,e e f são números reais dados. É possível mostrar que a equação acima representa uma cônica ou uma reta ou duas retas ou um ponto ou nenhum lugar geométrico em $\mathbb{R}^2$. Vejamos alguns exemplos

Exemplo 56. Vejamos que lugar geométrico em $\mathbb{R}^2$ cada uma das equações abaixo representa.

1. $x^2+y^2+1=0$,
2. $2x^2+4y^2=0$,
3. $x^2-9=0$,
4. $4x^2+9y^2-8x-36y+4=0$,
5. $y^2+6y^2-8x+1=0$

1. Está equação representa nenhum lugar geométrico em $\mathbb{R}^2$, pois:

$$\{(x,y)\in\mathbb{R}^2; x^2+y^2+1=0\}=\{(x,y)\in\mathbb{R}^2; x^2+y^2=-1\}=\emptyset$$

2. Esta equação representa a origem do plano cartesiano, pois:

$$2x^2+4y^2=0$$

equivale à equação $x^2=-2y^2$, que é verificada somente se $x=y=0$;

3. Esta equação depresenta duas retas do $\mathbb{R}^2$.Mais precisamente, as retas $x=3$ e $x=-3$;

4. Esta equação representa uma elipse.De fato, $(x,y) \in \mathbb{R}^2$ tal que $4x^2+9y^2-8x-36y+4=0$ que é equivalente à

$$4(x^2-2x)+9(y^2-4y)=-4$$

Completando os quadrados temos:

$$4(x-1)^2+9(y-2)^2=36$$

ou seja

$$\frac{(x-1)^2}{9}+\frac{(y-2)^2}{4}=1$$

que é a equação reduzida da elipse de centro $(1,2)$ e eixo maior e menor medindo 6 e 4 respectivamente.

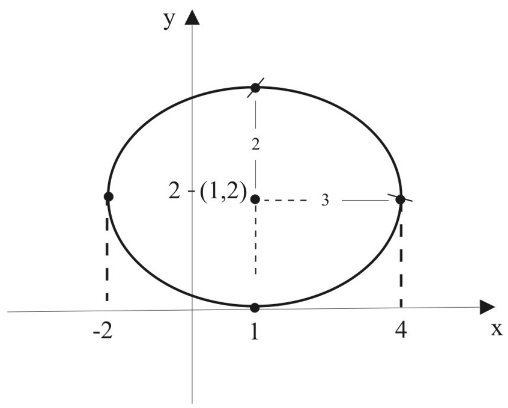

Figura 15 – A elipse $\frac{(x-1)^2}{9}+\frac{(x-2)^2}{4}=1$

5. Esta equação representa uma parabola. De fato, $(x,y) \in \mathbb{R}^2$ tal que $y^2+6y-8x+1=0$.Que é equivalente à:

$$(y^2+6y)=8x-1$$

Completando os quadrados

$$(y+3)^2=8(x+1)$$

que é a equação reduzida da parabola de vértice $(-1,-3)$ e parâmetro 2.

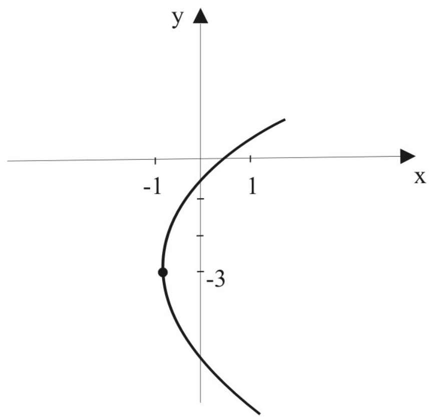

Figura 16 – $\{(x,y) \in \mathbb{R}^2 : (y+3)^2 = 8(x+1)\}$

Observamos que em todos os exemplos anteriores o termo xy, o chamado termo misto da equação não aparece. A técnica usualmente utilizada nestes casos é a técnica de completar quadrados. Porém em equações em que o termo misto aparece, precisamos de uma algebra mais avançada para reduzirmos a equação dada.

Exemplo 57. Vamos determinar o lugar geométrico em $\mathbb{R}^2$ representado pela equação:

$$2x^2 + 2xy + 2y^2 + 7\sqrt{2}x + 5\sqrt{2}y + 10 = 0?$$

Para respondermos a esta pergunta vamos usar o Teorema Espectral. Primeiramente a equação acima equivale a equação matricial.

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 7\sqrt{2} & 5\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 10 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix}$$

Sendo $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Como A é uma matriz simétrica e pelo Teorema Espectral, A é ortogonalmente diagonalizável.

Assim vamos encontrar os autovalores.

O polinômio característico:

$$\det(tI_2 - A) = \det \begin{pmatrix} t-2 & -1 \\ -1 & t-2 \end{pmatrix} = (t-2)^2 - 1 = t^2 - 4t + 3 \Longrightarrow t_1 = 3 \;\; t_2 = 1$$

Agora os autovetores associados aos autovalores t_1 e t_2

$$t_1 = 3 \Longrightarrow \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Longrightarrow x = y$$

$$t_2 = 1 \Longrightarrow \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Longrightarrow x = -y$$

Assim os vetores $\left(\text{sendo } x = \dfrac{1}{\sqrt{2}}\right)$ unitários $v_1 = \left(\dfrac{1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$ e o vetor unitário $v_2 = \left(\dfrac{-1}{\sqrt{2}}, \dfrac{1}{\sqrt{2}}\right)$ são autovetores de t_1 e t_2, respectivamente.

Assim $\beta = \{v_1, v_2\}$ é uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T_A. Seja $P = [I_{\mathbb{R}^2}]_\alpha^\beta$ onde α é a base canônica de $\mathbb{R}^2$. Chame $D = P^{-1}AP$.

Temos: $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ e $D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$.

Como $A = PDP^t$, já que $P^{-1} = P^t$, segue que

$$\begin{pmatrix} x & y \end{pmatrix} P \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} P^t \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 7\sqrt{2} & 5\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 10 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix}$$

.

Observamos que $P^t \begin{pmatrix} x \\ y \end{pmatrix} = [I_{\mathbb{R}^2}]_\beta^\alpha v_\alpha$.

Chamaremos $[v]_\beta$ de $\begin{pmatrix} x' \\ y' \end{pmatrix}$. Substituindo na equação acima obtemos

$$\begin{pmatrix} x' & y' \end{pmatrix} P \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} P^t \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 7\sqrt{2} & 5\sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 10 \end{pmatrix} = \begin{pmatrix} 0 \end{pmatrix}$$

ou seja

$$3x'^2 + y'^2 + 12x' - 2y' + 10 = 0$$

Com a mudança da base canônica α para a base β, reduzimos a primeira equação a última equação, que não apresenta termos mistos $x''y'$

Agora vamos reduzir a ultima equação completando os quadrados

$$3x'^2 + y'^2 + 12x' - 2y' + 10 = 0$$

equivale a equação

$$3(x'+2)^2 + (y'-1)^2 = 3$$

ou seja

$$(x'+2)^2 + \frac{(y'-1)^2}{3} = 1$$

Portanto a primeira equação representa uma elipse. para esboçarmos o gáfico dessa elipse, precisamos considerar as novas coordenadas x' e y'. Assim nesse sistema de coordenadas, a elipse tem centro $(-2, 1)$, semi-eixo menor medindo 1 e semi-eixo maior medindo $\sqrt{3}$, sendo este semi-eixo paralelo ao eixo y'.

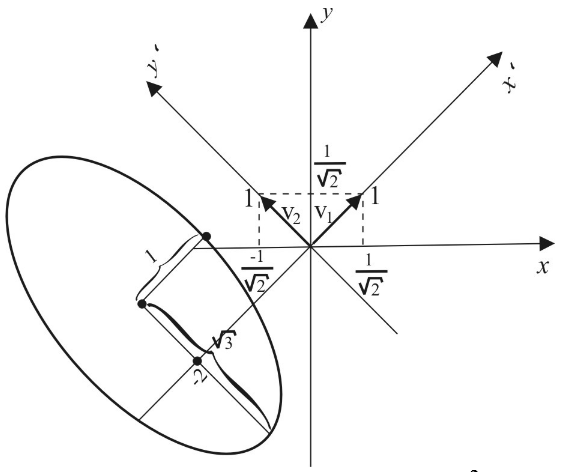

Figura 17 – Elipse $(x'+2)^2 + \dfrac{(y'-1)^2}{3} = 1$

Exemplo 58. Considere a seguinte equação quadratica

$$x^2 - 6xy - 7y^2 + 10x + 2y + 9 = 0.$$

A matriz associada a forma quadrática é

$$A = \begin{pmatrix} 1 & -3 \\ -3 & -7 \end{pmatrix}.$$

Os autovalores t_1 e t_2 podem ser encontrados calculando as raízes do polinômio característico, que neste caso é dado por

$$(t-1)(t+7)(3)^2 = t^2 + 6t - 17 \Longrightarrow t_1 = -8, \quad t_2 = 2.$$

Agora os autovalores associados.

$$t_1 = -8 \Longrightarrow \begin{pmatrix} 1 & -3 \\ -3 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -8 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} x - y = -8x \\ -3x - 7y = -8y \end{cases} \Longrightarrow x = x, \quad y = 3x.$$

$$t_2 = 2 \Longrightarrow \begin{pmatrix} 1 & -3 \\ -3 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} x - y = 2x \\ -3x - 7y = 2y \end{cases} \Longrightarrow x = x, \quad y = -\frac{1}{3}x.$$

Portanto os autovetores s e r associados respectivamente a t_1 e t_2, são dados por

$$s = \begin{bmatrix} x \\ 3x \end{bmatrix}, \; r = \begin{bmatrix} x \\ -\frac{1}{3}x \end{bmatrix}.$$

Para $x = 1$, temos que os autovetores normalizados associados respectivamente a t_1 e t_2 são

$$v_1 = \begin{bmatrix} \frac{\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} \end{bmatrix}, \; v_2 = \begin{bmatrix} \frac{3\sqrt{10}}{10} \\ -\frac{\sqrt{10}}{10} \end{bmatrix}.$$

Assim $\beta = \{v_1, v_2\}$ é uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T_A. Seja $P = [I\mathbb{R}^2]_\alpha^\beta$ onde α é a base canônica de $\mathbb{R}^2$. Chame $A = P^{-1}DP$. Então:

A matriz P que diagonaliza A ortogonalmente é

$$P = \begin{pmatrix} \frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} & -\frac{\sqrt{10}}{10} \end{pmatrix}.$$

Sendo $D = P^{-1}AP$, temos que $D = \begin{pmatrix} -8 & 0 \\ 0 & 2 \end{pmatrix}$. Como $A = P^{-1}DP$ e $P^{-1} = P^t$, temos

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} -8 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 10 & 2 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{10}}{10} & \frac{3\sqrt{10}}{10} \\ \frac{3\sqrt{10}}{10} & -\frac{\sqrt{10}}{10} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 9 \end{pmatrix} = 0$$

$$\Longrightarrow -8x'^2 + 2y'^2 + \frac{8\sqrt{10}}{5}x' + \frac{14\sqrt{10}}{5}y' + 9 = 0$$

Com a mudança da base canônica α para a base β, reduzimos a primeira equação a última equação, que não apresenta termos mistos xy'.

$$-8x"^2 + 2y"^2 = 0 \Longleftrightarrow y = \pm 4x.$$

Neste caso a cônica é um par de retas. Como representado na figura 18.

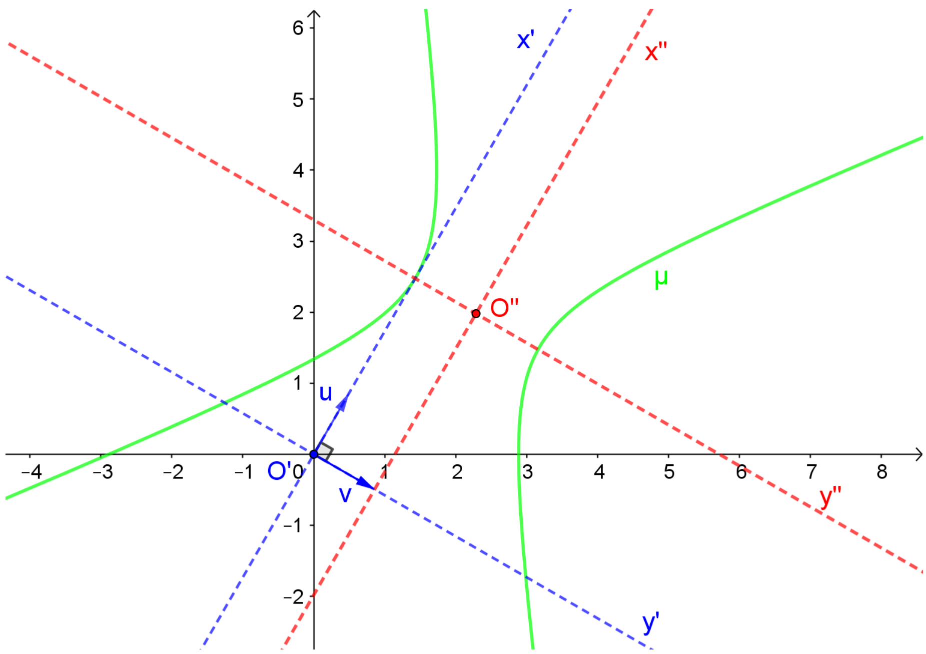

Figura 18 – Hiperbole - $\mathscr{H} = 3x^2 - 4\sqrt{3}xy - y^2 + 20y = 25$.

Exemplo 59. Considere a seguinte equação quadratica

$$3x^2 - 4\sqrt{3}xy - y^2 + 20y = 25.$$

A matriz associada a forma quadrática é

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

Os autovalores t_1 e t_2 podem ser encontrados calculando as raízes do polinômio característico, que neste caso é dado por

$$(t-3)(t+1)(-2\sqrt{3})^2 = t^2 - 2t - 15 \Longrightarrow t_1 = -3, \quad t_2 = 5.$$

Agora os autovalores associados.

$$t_1 = -3 \Longrightarrow \begin{pmatrix} 3 & -2\sqrt{3} \\ 2\sqrt{3} & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -3 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} 3x - 2\sqrt{3}y = -3x \\ -2\sqrt{3}x - y = -3y \end{cases} \Longrightarrow x = x, \quad y = \sqrt{3}x.$$

$$t_2 = 5 \Longrightarrow \begin{pmatrix} 3 & -2\sqrt{3} \\ 2\sqrt{3} & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} 3x - 2\sqrt{3}y = 5x \\ -2\sqrt{3}x - y = 5y \end{cases} \Longrightarrow x = x, \quad y = -\frac{\sqrt{3}}{3}x.$$

Portanto os autovetores s e r associados respectivamente a t_1 e t_2, são dados por

$$s = \begin{bmatrix} x \\ \sqrt{3}x \end{bmatrix}, \; r = \begin{bmatrix} x \\ -\frac{\sqrt{3}}{3}x \end{bmatrix}.$$

Para $x = 1$, temos que os autovetores normalizados associados respectivamente a t_1 e t_2 são

$$v_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}, \; v_2 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Assim $\beta = \{v_1, v_2\}$ é uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T_Q. Seja $P = [I_{\mathbb{R}^2}]_\alpha^\beta$ onde α é a base canônica de $\mathbb{R}^2$. Chame $A = P^{-1}DP$. Então A matriz P que diagonaliza A ortogonalmente é

$$P = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

Sendo $D = P^{-1}AP$, temos que $D = \begin{pmatrix} -3 & 0 \\ 0 & 5 \end{pmatrix}$. Como $A = P^{-1}DP$ e $P^{-1} = P^t$, temos

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} -3 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} x' \\ y'' \end{pmatrix} = \begin{pmatrix} 0 & 20 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} -25 \end{pmatrix} = 0$$

$$\Longrightarrow -3x'^2 + 5y'^2 + 10\sqrt{3}x' - 10y' - 25 = 0$$

Com a mudança da base canônica α para a base β, reduzimos a primeira equação a última equação, que não apresenta termos mistos xy'.

$$-3x\text{"}^2 + 5y\text{"}^2 = 5 \Longleftrightarrow -\frac{x\text{"}^2}{\left(\sqrt{\frac{5}{3}}\right)^2} + \frac{y\text{"}^2}{1^2} = 1.$$

Neste caso a cônica é uma hipérbole cuja a diretriz é paralela ao eixo x”. Como representado na figura 19.

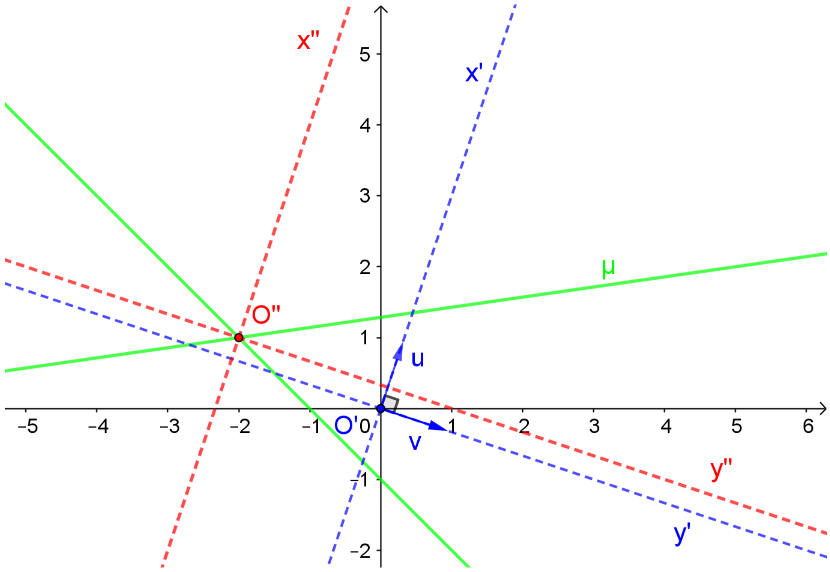

Figura 19 – Par de retas: $x^2 - 6xy - 7y^2 + 10x + 2y + 9 = 0$.

Exemplo 60. Considere a seguinte equação quadratica

$$x^2 + 2xy + y^2 + -2x + 2y + 3 = 0.$$

A matriz associada a forma quadrática é

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

Os autovalores t_1 e t_2 podem ser encontrados calculando as raízes do polinômio característico, que neste caso é dado por

$$(t-1)(t-1) - 1^2 = t^2 - 2t \Longrightarrow t_1 = 2, \quad t_2 = 0.$$

Agora os autovalores associados.

$$t_1 = 2 \Longrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} x + y = 2x \\ x + y = 2y \end{cases} \Longrightarrow x = x, \quad y = x.$$

$$t_2 = 0 \Longrightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \begin{pmatrix} x \\ y \end{pmatrix} \Longleftrightarrow \begin{cases} x + y = 0 \\ x + y = 0 \end{cases} \Longrightarrow x = x, \quad y = -x.$$

Portanto os autovetores s e r associados respectivamente a t_1 e t_2, são dados por

$$r = \begin{bmatrix} x \\ x \end{bmatrix}, \; r = \begin{bmatrix} x \\ -x \end{bmatrix}.$$

Para $x = 1$, temos que os autovetores normalizados associados respectivamente a t_1 e t_2 são

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \ v_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Assim $\beta = \{v_1, v_2\}$ é uma base ortonormal de $\mathbb{R}^2$ formada por autovetores de T_Q. Seja $P = [I\mathbb{R}^2]_\alpha^\beta$ onde α é a base canônica de $\mathbb{R}^2$. Chame $A = P^{-1}DP$. Então A matriz P que diagonaliza A ortogonalmente é

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Sendo $D = P^{-1}AP$, temos que $D = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Como $A = P^{-1}DP$ e $P^{-1} = P^t$, temos

$$\begin{pmatrix} x' & y' \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x' \\ y'' \end{pmatrix} = \begin{pmatrix} -2 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 3 \end{pmatrix} = 0$$

$$\Longrightarrow -3x'^2 + 5y'^2 + 10\sqrt{3}x' - 10y' - 25 = 0$$

Com a mudança da base canônica α para a base β, reduzimos a primeira equação a última equação, que não apresenta termos mistos xy'.

$$\frac{x"^2}{\sqrt{2}} = y"^2 \Longleftrightarrow x"^2 = \sqrt{2}y".$$

Neste caso a cônica é uma hipérbole cuja a diretriz é paralela ao eixo x". Como representado na figura 20.

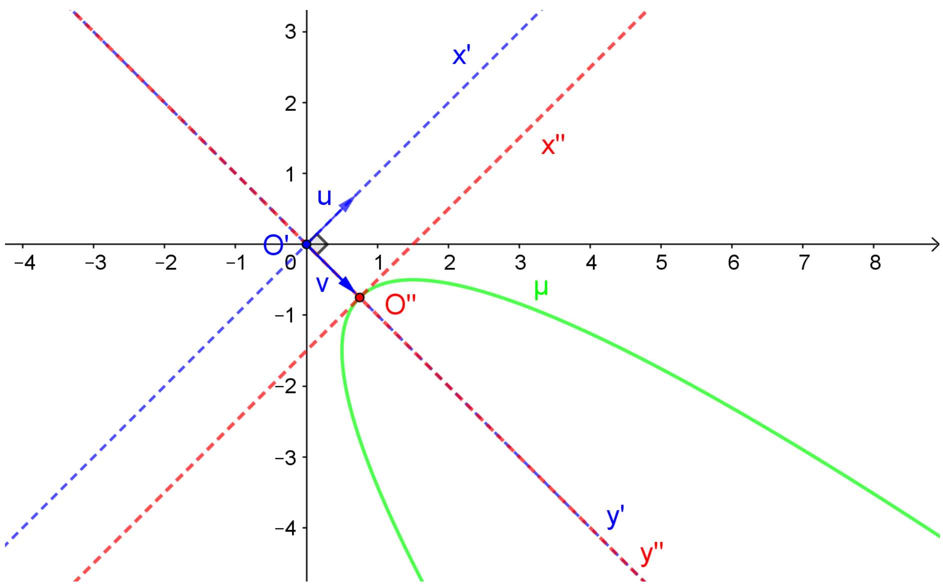

Figura 20 – Parábola $\mathscr{P} : x^2 + 2xy + y^2 - 2x + 2y + 3 = 0$.

CAPÍTULO

7

SISTEMAS DE EQUAÇÕES DIFERENCIAIS ORDINÁRIAS

Neste capítulo, introduziremos o conceito de sistemas de equações diferenciais ordinárias lineares e, utilizando os resultados de diagonalização de operadores, exibiremos uma forma de encontrar as soluções de tais sistemas quando a matriz é diagonalizável. A bibliografia usada para este capítulo pode ser encontrada em (LADEIRA L. A. C.; JUNIOR, 2011), (LADEIRA, 2004) e (BOYCE, 2002)

Primeiramente, vamos descrever o que é uma equação diferencial ordinária. Podemos dizer, sem muito formalismo, que uma equação diferencial é uma relação que envolve uma "função incógnita"e suas derivadas ou diferenciais. Assim, são exemplos de equações diferenciais

- $\dot{y}(t) = f(t)$, onde $\dot{y}(t) = \dfrac{dy}{dt}$,
- $\ddot{y}(t) + y(t) = 0$,
- $M(x,y)dx + N(x,y)dy = 0$.

Definição 33. Uma equação diferencial ordinária (E.D.O.) é uma equação diferencial na qual a função incógnita depende apenas de uma variável.

As equações do exemplo acima são equações diferenciais ordinárias. A ordem de uma equação diferencial é a ordem da mais alta derivada da função incógnita. Uma solução de uma equação diferencial é uma função definida num intervalo que, juntamente com suas derivadas, satisfaz a equação diferencial dada. Por exemplo, a função $y(t) = \sin t$ é uma solução da E.D.O. de segunda ordem $\ddot{y} + y = 0$, pois $\dfrac{d^2 \sin t}{dt^2} + \sin t = -\sin t + \sin t = 0$.

Um sistema de equações diferenciais ordinárias de primeira ordem pode, geralmente, ser escrito sob a forma:

$$\begin{cases} \dot{x_1} = F_1(t,x_1,x_2,\cdots,x_n) \\ \dot{x_2} = F_2(t,x_1,x_2,\cdots,x_n) \\ \vdots \\ \dot{x_n} = F_n(t,x_1,x_2,\cdots,x_n) \end{cases} \tag{7.1}$$

Uma solução de um sistema de equações diferenciais em um intervalo J é constituída por n funções $x_1(t), x_2(t), \cdots, x_n(t)$ tais que, cada uma destas funções são diferenciáveis em J e satisfazem o sistema (7.1) para todo $t \in J$.

Sistemas de equações diferenciais ocorrem em muitas aplicações como circuitos elétricos, mistura química de vários ingredientes, crescimento de duas ou mais populações interadas, vibrações de estruturas, etc.

Como exemplo, podemos citar o par de funções $x_1(t) = \sin t$ e $x_2(t) = \cos t$ como solução do sistema

$$\begin{cases} \dot{x_1} = x_2, \\ \dot{x_2} = -x_1. \end{cases}$$

Um problema de valor inicial (PVI) para uma equação diferencial ordinária pode ser colocado da seguinte maneira:

$$\begin{cases} \dot{y}(t) = f(t,y(t)) \\ y(t_0) = y_0. \end{cases} \tag{7.2}$$

Um PVI para um sistema de equações diferenciais de primeira ordem é dado por:

$$\begin{cases} \dot{x_1} = F_1(t,x_1,x_2,\cdots,x_n) \\ \dot{x_2} = F_2(t,x_1,x_2,\cdots,x_n) \\ \vdots \\ \dot{x_n} = F_n(t,x_1,x_2,\cdots,x_n) \\ x_1(t_0) = x_1^0, x_2(t_0) = x_2^0, \cdots, x_n(t_0) = x_n^0 \end{cases} \tag{7.3}$$

onde $x_1^0, x_2^0, \cdots, x_n^0 \in \mathbb{R}$ são as condições iniciais do sistema.

Existem questões fundamentais a serem respondidas referentes a um PVI de uma equação diferencial ordinária e, consequentemente, a um sistema de equações diferenciais ordinárias com condições iniciais, como, por exemplo, se a equação tem solução e, se esse é o caso, se a solução encontrada é única.

Vamos observar alguns exemplos.

Exemplo 61. Seja $f : [a,b] \to \mathbb{R}$ uma função contínua. Pelo Teorema Fundamental do Cálculo temos que a função $F(t) = \int_a^t f(s)ds$, com $a \leq t \leq b$ e $F(a) = 0$ é uma primitiva para a função f e assim, $F(t)$ é uma solução do problema de valor inicial

$$\begin{cases} \dot{y}(t) = f(t) \\ y(a) = 0 \end{cases}.$$

Este PVI possui uma solução, mas surge a pergunta, $F(t)$ é a única solução para este PVI? Neste caso a resposta é positiva, pois, se $G(t)$ for uma outra solução, temos que $G'(t) = f(t) = F'(t)$ e isso implica que $(F-G)'(t) = 0$, ou seja, $(F-G)(t)$ =constante. Mas, $(F-G)(a) = F(a) - G(a) = 0 - 0 = 0$. Portanto, $G(t) = F(t)$ para todo $t \in [a,b]$.

Exemplo 62. No entanto, há problemas de valor inicial que possuem mais de uma solução. O problema de valor inicial

$$\begin{cases} \dot{y}(t) = |y|^{\frac{1}{2}} \\ y(0) = 0 \end{cases}$$

não têm unicidade de soluções, pois $y_1 \equiv 0$ é uma solução e

$$y_2(t) \begin{cases} \frac{t^2}{4}, t \leq 0 \\ -\frac{t^2}{4}, t < 0 \end{cases}$$

também é solução. Portanto, temos duas soluções para o problema.

Para respondermos se um PVI tem solução, podemos utilizar o Método de Picard. Para entendermos o método, consideramos o PVI

$$\begin{cases} \dot{y}(t) = f(t,y) \\ y(t_0) = y_0 \end{cases},$$

onde f é uma função definida num aberto A de $\mathbb{R}^2$. Suponhamos que $f(t,x)$ seja uma função contínua em (t,x) e continuamente derivável em x. Obsevamos que $y(t)$ é solução do PVI se, e somente se, $y(t)$ é solução da equação integral

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))ds.$$

Considere agora a sequência $y_n(t)$, dada da seguinte forma:

$$\begin{aligned}
y_0(t) &= y_0 \\
y_1(t) &= y_0 + \int_{t_0}^{t} f(s, y_0(s))ds, \\
y_2(t) &= y_0 + \int_{t_0}^{t} f(s, y_1(s))ds, \\
&\vdots \\
y_n(t) &= y_0 + \int_{t_0}^{t} f(s, y_{n-1}(s))ds
\end{aligned}$$

As funções $y_n(t)$ são chamadas iteradas de Picard. Pode-se mostrar que $y_n(t) \to y(t)$, quando $n \to \infty$, para t num intervalo conveniente. A função limite $y(t)$ será a solução do PVI. Este processo é conhecido por Método de Picard.

Para exemplificar o método, vamos considerar o PVI

$$\begin{cases} \dot{y} = y \\ y(0) = 1. \end{cases}$$

Observamos que, neste caso, $f(t,y) = y$, $t_0 = 0$ e $y_0 = 1$. A equação integral equivalente ao PVI dado é:

$$y(t) = 1 + \int_0^t y(s)ds$$

Considerando, inicialmente, $y_0(t) = 1$, obtemos, pelo processo de interação a próxima função, $y_1(t) = 1 + \int_0^t 1ds = 1 + t$. Usando $y_1(t)$ no integrando, encontramos a próxima função, $y_2(t) = 1 + \int_0^t y_1(s)ds = 1 + \int_0^t (1+s)ds = 1 + t + \frac{t^2}{2!}$. Continuando o processo, obtemos:

$$y_3(t) = 1 + \int_0^t y_2(s)ds = 1 + \int_0^t (1 + s + \frac{s^2}{2!})ds = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!}$$

$$y_n(t) = 1 + \int_0^t y_{n-1}(s)ds = 1 + \int_0^t (1 + s + \frac{s^2}{2!} + \cdots + \frac{s^{n-1}}{(n-1)!})ds = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!}$$

Como $e^t = \lim_{n\to\infty}(1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!}) = \lim_{n\to\infty} y_n(t)$, vemos que as iteradas de Picard, $y_n(t)$, convergem para a função $y(t) = e^t$, que é solução do PVI dado.

Observação 5. Pode acontecer de a solução de um PVI não estar definida para todo $t \in \mathbb{R}$. Por exemplo, a função $y(t) = \tan(t + /4)$ é solução de $\dot{y} = 1 + y^2, y(0) = 1$, mas está definida somente no intervalo $(\frac{-3\pi}{4}, \frac{\pi}{4})$.

Por este fato, não podemos esperar que as iteradas de Picard convirjam para todo t. Para sabermos onde as iteradas de Picard convergem, tentamos encontrar um intervalo no qual todas as $y_n(t)$ são uniformemente limitadas, isto é, para quais valores de t existe uma constante $k > 0$ tal que $|y_n(t)| \leq k$ para todo $t \in (a,b)$. Este é o conteúdo do próximo lema.

Lema 2. Sejam $a, b \in \mathbb{R}$ e consideremos o retângulo

$$R = \{(t,y)/t_0 \leq t \leq t_o + a \text{ e } |y - y_0| \leq b\}.$$

Defina $M = max\{|f(t,y)|, (t,y) \in R\}$ e $\alpha = min\{a, \frac{b}{M}\}$. Então: $|y_n(t) - y_0| \leq M|t - t_0|$, para $t_0 \leq t \leq t_0 + \alpha$.

O próximo teorema nos apresenta as condições para a existência e unicidade de soluções para o PVI (7.2).

Teorema 19. (Existência e Unicidade Local). Suponha que f e $\frac{\partial f}{\partial y}$ sejam funções contínuas no retângulo R e suponha ainda que M e α são considerados como no lema anterior. Então o PVI

$$\begin{cases} \dot{y} = f(t,y) \\ y(t_0) = y_0 \end{cases}$$

possui uma e somente uma solução $y(t)$ no intervalo $t_0 \leq t \leq t_0 + \alpha$.

Exemplo 63. Consideramos a equação $\dot{y} = y^2 + \cos t^2$ com $y(0) = 0$. As funções $f(t,y) = y^2 + \cos t^2$ e $\frac{\partial f}{\partial y}(t,y) = 2y$, são contínuas em todo $\mathbb{R}^2$ e se $M = max\{|f(t,y)|, (t,y) \in R\} = max\{|y^2 + \cos t^2|, |y| \leq b \text{ e } 0 \leq t \leq a\} = b^2 + 1$, vemos, pelo teorema (19), que $y(t)$ existe para $0 \leq t \leq \alpha$, em que $\alpha = min\{a, \frac{b}{b^2+1}\}$. Como, a priori, podemos tomar qualquer valor de a, temos que o valor α será $\frac{b}{b^2+1}$ e este por sua vez atinge valor máximo $\frac{1}{2}$, logo $\alpha = \frac{1}{2}$. Portanto, a solução $y(t)$ existe e é única para $0 \leq t \leq \frac{1}{2}$.

Se cada uma das funções $F_1, \cdots, F_n$ em (7.3) for linear em $x_1, \cdots, x_n$, então dizemos que o sistema de equações diferenciais é linear. O sistema mais geral de n equações lineares de 1ª ordem possui a forma:

$$\begin{cases} \dot{x_1} = a_{11}(t)x_1 + \cdots + a_{1n}(t)x_n + g_1(t) \\ \vdots \\ \dot{x_n} = a_{n1}(t)x_1 + \cdots + a_{nn}(t)x_n + g_n(t) \end{cases} \tag{7.4}$$

Se $g_j(t) = 0$ para todo $1 \leq j \leq n$, então dizemos que o sistema de equações diferenciais lineares acima é homogêneo. Caso contrário, ele é não homogêneo. Para simplificar a representação do sistema, usaremos a notação matricial

$$A(t) = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}, g(t) = \begin{pmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{pmatrix} \text{ e } x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Assim, o sistema (7.4) pode ser expresso na forma compacta

$$\dot{x} = A(t)x + g(t) \tag{7.5}$$

onde $x = (x_1, \cdots, x_n)^t$.

Nosso principal objetivo neste capítulo é utilizar os resultados obtidos sobre diagonalização de operadores para encontrar as soluções do sistema (7.4) quando $A(t) = A$ e $g_j(t) = 0$, para $j = 1, \dots, n$. Para isso, vamos descrever a teoria necessária para descrevermos essas soluções.

Teorema 20. (Existência e Unicidade de Soluções para Sistemas). Suponha que as funções $a_{ij}(t)$ e $g_i(t)$, $1 \leq i, j \leq n$, sejam contínuas num intervalo J. Então dados $t_0 \in J$ e $x_0 \in \mathbb{R}^n$, existe uma única solução $x(t)$ de (7.5), definida em J, tal que $x(t_0) = x_0$.

Proposição 30. Se $u(t) = (x_1^1(t), \cdots, x_n^1(t))$ e $v(t) = (x_1^2(t), \cdots, x_n^2(t))$ são soluções do sistema homogêneo $\dot{x} = A(t)x$, então qualquer combinação linear $c_1u(t) + c_2v(t)$, em que c_1 e c_2 são constantes arbitrárias, também é solução do sistema linear homogêneo. Ou seja, o conjunto S de todas as soluções do sistema linear homogêneo é um espaço vetorial.

Teorema 21. Sejam $x^1(t), \cdots, x^k(t)$, soluções do sistema linear homogêneo $\dot{x} = A(t)x$ e seja $t_0 \in J$. Então $x^1(t), \cdots, x^k(t)$ são soluções linearmente independentes se, e somente se, os vetores $x^1(t_0), \cdots, x^k(t_0)$ são linearmente independentes em $\mathbb{R}^n$.

Demonstração. Suponhamos que $x^1(t), \cdots, x^k(t)$ sejam linearmente dependentes. Então, existem constantes $c_1, \cdots, c_k$ não todas nulas, tais que $c_1x^1(t) + \cdots + c_kx^k(t) = 0$, para todo $t \in J$.

Logo, $c_1x^1(t_0) + \cdots + c_kx^k(t_0) = 0$, com constantes $c_1, \cdots, c_k$ não todas nulas. Portanto, $x^1(t_0), \cdots, x^k(t_0)$ são linearmente dependentes em $\mathbb{R}^n$.

Reciprocamente, suponhamos que $x^1(t_0), \cdots, x^k(t_0)$ sejam linearmente dependentes em $\mathbb{R}^n$. Então, existem constantes $c_1, \cdots, c_k$ não todas nulas, tais que $c_1x^1(t_0) + \cdots + c_kx^k(t_0) = 0$. Temos que a função $u(t) = c_1x^1(t) + \cdots + c_kx^k(t)$,em que $c_1, \cdots, c_k$ são as constantes dadas acima, satisfaz o sistema linear homogêneo, pois é uma combinação linear de soluções. Além disso, $u(t_0) = 0$. Portanto, pelo Teorema (19), $u(t) = 0$ para todo t. Logo, $x^1(t), \cdots, x^k(t)$ são soluções linearmente dependentes. □

Teorema 22. A dimensão do espaço S de todas as soluções do sistema linear homogêneo $\dot{x} = A(t)x$ é n.

Demonstração. Vamos mostrar que sistema linear homogêneo possui n soluções linearmente independentes. Para isto, consideremos os vetores do $\mathbb{R}^n$, $e_1 = (1, 0, 0, \dots, 0)^t$, $e_2 = (0, 1, 0, \dots, 0)^t$

$, \ldots, e_n = (0,0,0,\ldots,1)$ e os PVI's

$$\begin{cases} \dot{x} = A(t)x \\ x^i(t_0) = e_i, i = 1,\cdots,n \ \text{ e } t_o \in J. \end{cases} \tag{7.6}$$

Pelo Teorema (20), temos que cada PVI possui uma única solução $x^i(t)$. Como os vetores $e_1,\cdots,e_n$ são linearmente independentes em $\mathbb{R}^n$, segue que $x^1(t),\cdots,x^n(t)$ são soluções linearmente independentes do sistema de equações diferenciais linear homogêneo. Resta mostrar que qualquer solução deste sistema pode ser escrita como combinação linear de $x^1(t),\cdots,x^n(t)$.

Seja $x(t)$ uma solução do sistema de equações diferenciais linear homogêneo tal que $x(t_0) = (c_1,\cdots,c_n)^t$. Com estas constantes $c_1,\cdots,c_n$, construímos a função $u(t) = c_1x^1(t) + \cdots + c_nx^n(t)$. Temos que $u(t)$ satisfaz o sistema linear homogêneo pois é combinação linear de soluções e além disso

$$u(t_0) = c_1x^1(t_0) + \cdots + c_nx^n(t_0) = c_1e_1 + c_2e_2 + \cdots + c_ne_n = (c_1,c_2,\ldots,c_n)^t = x(t_0).$$

Novamente, pelo Teorema (19), $u \equiv x$. Portanto,

$$x(t) = c_1x^1(t) + \cdots + c_nx^n(t)$$

$\square$

Observação 6. O teorema anterior diz que se conhecermos n soluções linearmente independentes $x^1(t),\cdots,x^n(t)$ do sistema linear homogêneo, então toda solução deste sistema será da forma $x(t) = c_1x^1(t) + \cdots + c_nx^n(t)$. Por esta razão, esta expressão é chamada solução geral do sistema linear homogêneo.

Definição 34. Dizemos que uma matriz $n \times n$, $X(t)$ é matriz solução do sistema $\dot{x} = A(t)x$, se cada coluna de $X(t)$ é solução do sistema.

Exemplo 64. Temos que $X(t) = \begin{pmatrix} e^t & 0 \\ 0 & e^{2t} \end{pmatrix}$ é uma matriz solução de $\dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} x$, pois $x^1(t) = \begin{pmatrix} e^t \\ 0 \end{pmatrix}$ e $x^2(t) = \begin{pmatrix} 0 \\ e^{2t} \end{pmatrix}$ são soluções do sistema.

Definição 35. Dizemos que uma matriz $n \times n$, $X(t)$ é matriz fundamental (M.F.) para o sistema $\dot{x} = A(t)x$ se $X(t)$ é uma matriz solução e $det\,X(t) \neq 0$ para todo t no intervalo de existência. Ou seja, suas colunas são soluções linearmente independentes de $\dot{x} = A(t)x$.

No exemplo anterior $X(t) = \begin{pmatrix} e^t & 0 \\ 0 & e^2t \end{pmatrix}$ é uma M.F. de $\dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} x$ pois, como vimos acima, ela é matriz solução e além disso $\det X(t) = e^{3t} \neq 0$ para todo t.

Teorema 23. Se $X(t)$ é uma M.F. do sistema linear homogêneo, então a solução geral será dada por $x(t) = X(t)c$, em que $c = (c_1,\ldots,c_n)^t$.

Com o auxílio do próximo resultado:

Lema 3. Se $X(t)$ é uma matriz solução do sistema linear homogêneo em algum intervalo J e se $t_0 \in J$, então

$$\det X(t) = \det X(t_0)\exp(\int_{t_0}^{t} trA(s)ds)$$

onde $trA(s) =$ soma dos elementos da diagonal principal de $A(s)$.

podemos provar o seguinte teorema:

Teorema 24. Seja $X(t)$ uma matriz solução do sistema de equações diferenciais linear homogêneo em J. $X(t)$ é M.F. se, e somente se, $\det X(t_0) \neq 0$ para algum $t_0 \in J$.

Vamos considerar o sistema $\dot{x} = Ax$, onde $A = (a_{ij}), i, j = 1, 2, \cdots, n$ é uma matriz constante e a partir da teoria descrita anteriormente vamos construir a solução geral do sistema. Para tanto, vamos procurar por soluções da forma $x(t) = e^{\lambda t}v$, em que o número λ e o vetor constante $v = (v_1, v_2, \cdots, v_n)^t$ não nulo devem ser determinados. Substituindo a solução no sistema, obtemos $\lambda e^{\lambda t}v = Ae^{\lambda t}v$ ou, equivalentemente, $Av = \lambda v$.

Logo, temos uma solução de $\dot{x} = Ax$ se, e somente se, λ é um autovalor de A e v é um autovetor associado a λ. A natureza dos autovalores e autovetores associados determinam a natureza da solução do sistema.

Proposição 31. Seja A uma matriz quadrada $n \times n$ e sejam $\lambda_1, \ldots, \lambda_n$ os autovalores de A, reais e distintos, com $v_1, v_2, \ldots v_n$ autovetores associados a esses autovalores. Se os autovetores forem linearmente independentes, então as soluções $x^1(t) = e^{\lambda_1 t}v_1, \ldots x^n(t) = e^{\lambda_n t}v_n$ são soluções linearmente independentes de $\dot{x} = Ax$.

Demonstração. Só precisamos mostrar que as soluções são linearmente independentes e, para isso, vemos que, calculadas em $t_0 = 0$, temos os vetores $v_1, \ldots, v_n$ que são linearmente independentes e usamos o teorema anterior. □

Teorema 25. Se $A \in M_n$ for simétrica, então a solução geral do sistema linear homogêneo $\dot{x} = Ax$ é dada por $x(t) = c_1 e^{\lambda_1(t)}v_1 + \ldots + c_n e^{\lambda_n t}v_n$, onde $\lambda_1, \lambda_2 \ldots, \lambda_n$ são os autovalores de A (não necessariamente distintos) e $v_1, v_2, \ldots, v_n$ são os autovetores de A associados aos seus respectivos autovalores.

Demonstração. De fato, A sendo simétrica, todos os autovalores serão reais e, mesmo que sejam repetidos, haverá uma base ortonormal de autovetores associados a esses autovalores, ou sejam, os autovetores serão linearmente independentes e a solução geral do sistema será dada como combinação linear das funções $e^{\lambda_i(t)}v_i$, para $i = 1, \ldots, n$.

□

Exemplo 65. Vamos considerar o sistema $\dot{x} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} x$, onde, obviamente, A é simétrica. Os autovalores da matriz são, calculando $det(A - \lambda I) = 0$, $\lambda_1 = 2$ e $\lambda_2 = -1$, com multiplicidade algébrica igual a 2. O autovetor associado ao autovalor λ_1 é $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ e os autovetores associados ao autovalor λ_2 são $v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ e $v_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

Portanto, a solução geral do sistema é dada por

$$x(t) = c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{2t} + c_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{-t} + c_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} e^{-t}.$$

Vamos, agora, olhar para as soluções dos sistema $\dot{x} = Ax$ utilizando a matriz fundamental, fazendo uma associação com as soluções da equação $\dot{x} = ax$, com $x(0) = x_0$, dadas por $x(t) = x_0 e^{at}$.

Para isso, consideramos a matriz solução $\Phi(t)$ do sistema $\dot{x} = Ax$ com $x(0) = x^0$, obtida como no teorema (22) considerando $\Phi(0) = I$.

Comparando a equação com o sistema, podemos sugerir que tal matriz $\Phi(t)$ tenha carater exponencial.

Podemos verificar que $\Phi(t) = e^{At}$, one A é a matriz do sistema. Tem-se várias considerações a fazer sobre essa afirmação. Primeiro, qual o significado de e^{At}? Definimos e^{At} da seguinte forma:

$$e^{At} = I + \sum_{n=1}^{\infty} \frac{A^n t^n}{n!}$$

É possível mostrar que cada elemento dessa soma de matrizes converge para todo t quando $n \to \infty$ e mais, que e^{At} satisfaz a equação $\dot{x} = Ax$, com $x(0) = I$ e, pela unicidade de soluções $\Phi(t) = e^{At}$. Mais ainda, é possível mostrar que essa função matricial tem as mesmas propriedades "convenientes"que tem a função exponencial real.

Quando a matriz A é diagonalizável, o cálculo desta exponencial é relativamente mais simples. A ideia que está por trás disso é transformar o sistema $\dot{x} = Ax$ em um sistema $\dot{y} = Dy$, onde D é diagonal e, dessa forma, as n equações que aparecem no sistema são desacopladas.

Se A for diagonalizável, o que é o caso quando A é simétrica, temos que existe uma base de autovetores associados aos autovalores de A tal que $D = P^{-1}AP$. onde P é formada pelos autovetores e D é diagonal, contendo em sua diagonal principal os autovaloes de A. Desta forma, se definirmos $x = Py$, temos

$$P\dot{y} = APy$$

e,

$$\dot{y} = (P^{-1}AP)y = Dy. \tag{7.7}$$

Uma matriz fundamental para o sistema (7.7) é a matriz diagonal

$$e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & e^{\lambda_n t} \end{pmatrix}$$

Logo, uma matriz fundamental para o sistema $\dot{x} = Ax$ é dada por

$$\Phi(t) = Pe^{Dt}.$$

Desta forma, podemos observar que resolver um sistema de equações diferenciais com essas condições e diagonalizar uma matriz estão estritamente relacionados.

Vamos fazer um exemplo para descrever todo esse processo.

Exemplo 66. Consideramos o sistema $\dot{x} = Ax$, do exemplo (65), que já apresentamos a solução geral, mas vamos utilizar a matriz exponencial para resolvê-lo. Obviamente, A é diagonalizável, pois é simétrica e temos $D = P^{-1}AP$, onde $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$ e $P^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix}$.

Observamos que

$$D^2 = \begin{pmatrix} 2^2 & 0 & 0 \\ 0 & (-1)^2 & 0 \\ 0 & 0 & (-1)^2 \end{pmatrix}, D^3 = \begin{pmatrix} 2^3 & 0 & 0 \\ 0 & (-1)^3 & 0 \\ 0 & 0 & (-1)^3 \end{pmatrix}, \dots, D^n = \begin{pmatrix} 2^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & (-1)^n \end{pmatrix},$$

e,

$$e^{Dt} = I + \sum_{n=1}^{\infty} \frac{1}{n!} \begin{pmatrix} (2t)^n & 0 & 0 \\ 0 & (-1t)^n & 0 \\ 0 & 0 & (-1t)^n \end{pmatrix} = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-t} \end{pmatrix}.$$

Portanto, a solução será dada por

$$\Phi(t) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & 0 & e^{-t} \end{pmatrix} = \begin{pmatrix} e^{2t} & e^{-t} & 0 \\ e^{2t} & 0 & e^{-t} \\ e^{2t} & -e^{-t} & -e^{-t} \end{pmatrix},$$

conforme foi também demonstrado no exemplo (65).

CAPÍTULO

8

CONCLUSÃO

O domínio dos conceitos matemáticos, das demonstrações, das definições é importante para a construção de novos conceitos e isso permite ao estudante a validação de intuições na construção de técnicas aplicadas em diversas situações.

A Matemática, diante disso, tem um papel importante no Ensino Médio, pois cabe a ela a apresentação de novas informações e instrumentos que deem condições ao estudante de continuar aprendendo.

Entende-se a importância do estudo das matrizes, determinantes e sistemas lineares.

No entanto, os conhecimentos adquiridos pelos alunos ficam limitados em sua maioria a cálculos abstratos.

Portanto, a perspectiva é que o trabalho aqui apresentado seja usado para futuras aplicações no último ano do Ensino Médio no estudo das cônicas. Isso representará um ganho para os alunos uma vez que serão recordados conceitos de matrizes, determinantes e sistemas lineares no enfoque da geometria analítica. Além disso, buscar técnicas para a identificação de cônicas de modo geral aumenta a quantidade de conteúdos adquiridos pelos alunos, possibilitando, assim, uma visão geral do conteúdo de cônicas que, sem dúvida, aumentarão seus conhecimentos de Matemática de um modo geral.

REFERÊNCIAS

BOYCE, R. C. D. W. E. **Equações Diferenciais Elementares e Problema de Valores de Contorno**. [S.l.]: LTC, 2002. Citado na página 99.

FRENSEL K.; DELGADO, J. **Geometria Analítica**. [S.l.]: UFMA, 2011. Citado nas páginas 41 e 81.

GOMEZ KATIA R. FRENSEL, L. S. C. J. J. D. **Geometria Analítica**. [S.l.]: SBM, 2017. Citado na página 81.

HEFEZ A.; SOUZA FERNANDES, C. **Introdução à Álgebra Linear**. [S.l.]: SBM, 2012. ISBN 978-85-85818-61-6. Citado nas páginas 23, 41, 55, 67 e 81.

LADEIRA, L. A. C. **Álgebra Linear e Equações Diferenciais**. [S.l.]: ICMC, 2004. Citado nas páginas 41 e 99.

LADEIRA L. A. C.; JUNIOR, H. C. **Notas de Aula : Equações Diferenciais Ordinárias**. [S.l.], 2011. Citado na página 99.

LIPSCHUTZ, S. **Álgebra Linear**. [S.l.]: São Paulo: Makron Books, 1994. Citado na página 23.

STEINBRUCH, A. **Álgebra Linear**. [S.l.]: São Paulo: Pearson Makron Books, 1987. Citado na página 23.

ZANI, S. L. **Álgebra Linear**. [S.l.]: ICMC- USP. Citado nas páginas 23, 41, 55 e 67.

ICMC USP
SÃO CARLOS