

Universidade Estadual do Sudoeste da Bahia
Departamento de Ciências Exatas e Tecnológicas

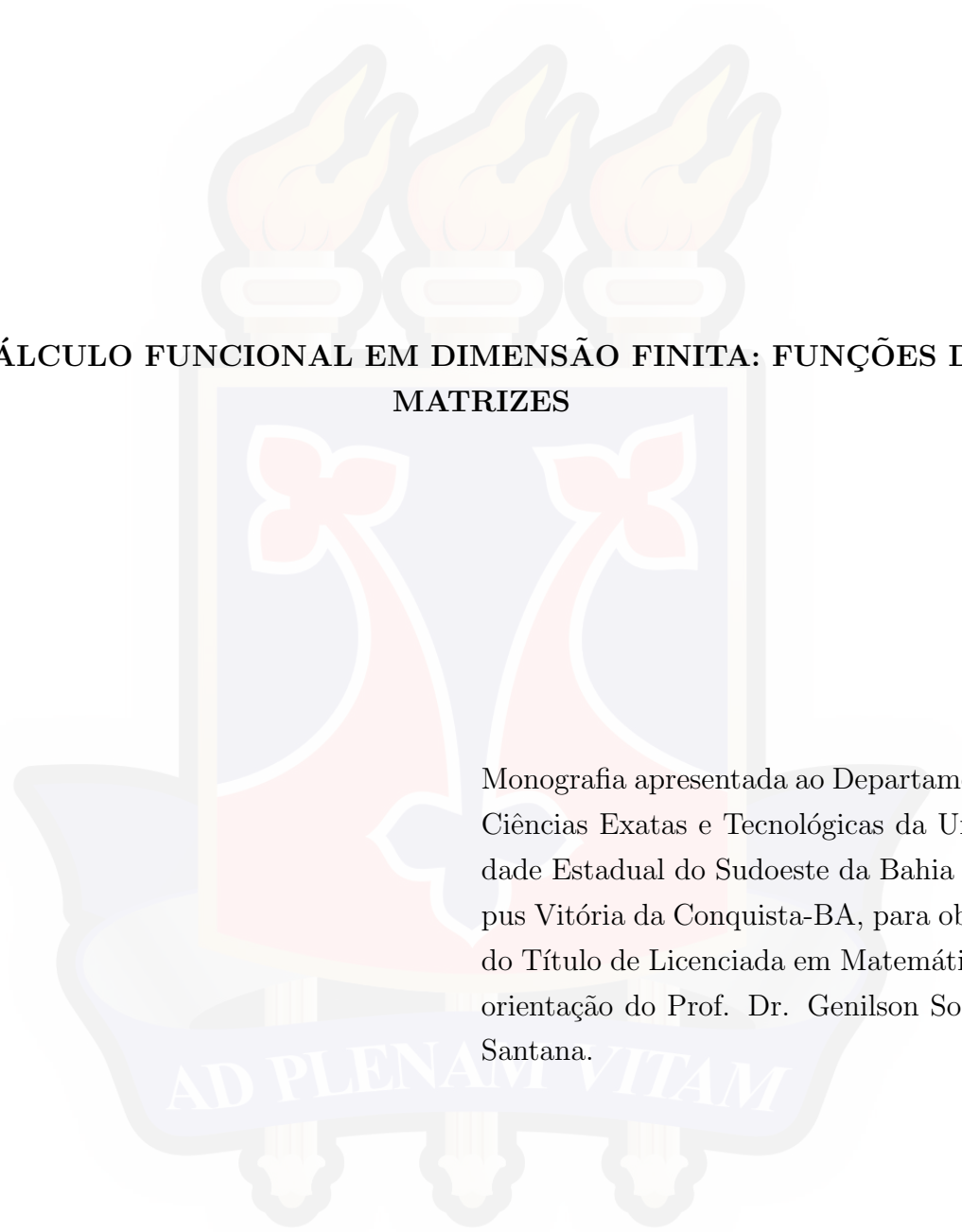
Licenciatura em Matemática

Franciele Almeida Campos

CÁLCULO FUNCIONAL EM DIMENSÃO FINITA:
FUNÇÕES DE MATRIZES

Vitória da Conquista - BA
2025

Franciele Almeida Campos



**CÁLCULO FUNCIONAL EM DIMENSÃO FINITA: FUNÇÕES DE
MATRIZES**

Monografia apresentada ao Departamento de Ciências Exatas e Tecnológicas da Universidade Estadual do Sudoeste da Bahia - Campus Vitória da Conquista-BA, para obtenção do Título de Licenciada em Matemática, sob orientação do Prof. Dr. Genilson Soares de Santana.

**Vitória da Conquista - BA
2025**

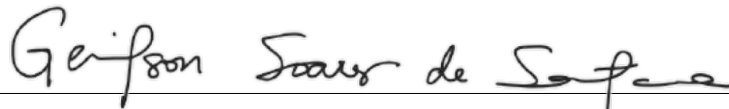
Franciele Almeida Campos

CÁLCULO FUNCIONAL EM DIMENSÃO FINITA: FUNÇÕES DE MATRIZES

Monografia apresentada ao Colegiado do Curso de Matemática como requisito parcial para aprovação na disciplina Seminário de Pesquisa II do Curso de Licenciatura em Matemática.

Trabalho aprovado em 02 de dezembro de 2025.

BANCA EXAMINADORA



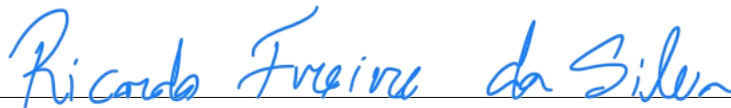
Prof. Dr. Genilson Soares de Santana - UESB

Orientador



Prof. Dr.^a Alessandra Oliveira Andrade - UESB

Convidada



Prof. Dr. Ricardo Freire da Silva - UESB

Convidado

Vitória da Conquista - BA

2025

”A única forma de chegar ao impossível é acreditar que é possível.”.
Alice no País das Maravilhas

AGRADECIMENTOS

Agradeço a todos que, de alguma forma, tornaram esta trajetória possível. Meu sincero e eterno obrigada. Mais do que a conclusão de uma etapa acadêmica, este trabalho é o reflexo do apoio e incentivo que recebi ao longo destes anos.

À Universidade Estadual do Sudoeste da Bahia (UESB), meu profundo agradecimento por ter sido o palco do meu crescimento intelectual e profissional.

Ao meu orientador, Professor Genilson, minha eterna gratidão por sua paciência, sabedoria, doçura e orientação. Agradeço os momentos em que acalmou minhas ansiedades, me acolheu e acreditou em mim. Obrigada por não apenas orientar um projeto, mas por se fazer presente em outros âmbitos da minha vida. Tenho muito orgulho e admiração pela pessoa e professor que tu és.

Aos demais professores, aos que, ao longo desta trajetória, compartilharam não apenas o conteúdo das ementas, mas também a paixão pelo conhecimento. Cada aula, conselho e questionamento contribuiu para a minha formação e para a construção do pensamento crítico necessário para chegar até aqui. Em especial, à banca avaliadora, por participar e contribuir nesse momento ímpar.

Aos meus colegas de curso, obrigada pelas trocas. Todos, de alguma forma, deixaram aprendizados que me fizeram crescer.

Aos meus amigos de curso, Alison, Luane, Maria Clara, Maria Luiza e Wéllington, agradeço pelo ombro amigo nos momentos de estresses, choros e de surtos pelas inúmeras demandas do curso. Foi uma honra compartilhar esta etapa com vocês.

Às minhas amigas, Julia e Marcella, por sempre me incentivar, aconselhar, motivar e acreditar em mim. Nos meus maiores momentos de dúvida e cansaço, vocês estavam lá, não apenas com palavras de incentivo, mas com uma crença inabalável em mim em que eu mesma não tinha. Ouvir de vocês que eu era capaz, fez toda a diferença. Amo vocês infinitamente.

À minha família, não existem palavras que possam expressar totalmente o que devo a vocês. Obrigada pelo amor incondicional e pelo apoio em cada decisão. Esta conquista é tão minha quanto é de vocês.

Ao meu amor, Julio César, obrigada pela paciência pelas diversas vezes que eu disse que iria sumir pois tinha muita coisa da faculdade para fazer, por cada abraço que acalmou e por acreditar em mim. Sua presença tornou os dias pesados mais leves. Obrigada por caminhar comigo e fazer dos meus sonhos os nossos. Te amo.

Por fim, agradeço a mim mesma por não ter desistido. Este trabalho é o fechar de um ciclo e a esperançosa abertura de muitos outros.

RESUMO

O presente trabalho consiste em uma apresentação detalhada sobre o cálculo funcional e operadores lineares definidos em espaços vetoriais de dimensão finita e aplicações desta teoria. O cálculo funcional constitui uma importante ferramenta dentro da Teoria de Operadores Lineares. Para espaços vetoriais V de dimensão finita, sabemos que o conjunto de operadores lineares de V em V , $\mathcal{L}(V)$, é isomorfo a um espaço de matrizes. Neste sentido, este cálculo nos permite definir e computar as funções de matrizes. Assim, o objetivo desta monografia é demonstrar como calculamos funções do tipo e^A , $\sin(A)$ e $\cos(A)$ em que A é uma matriz quadrada de ordem n . E ainda, sob algumas condições adicionais em A , mais precisamente sobre os autovalores, pode-se definir também $\log(A)$ e \sqrt{A} . Além disso, são exploradas aplicações do cálculo funcional na resolução de sistemas de equações diferenciais ordinárias lineares, particularmente na construção do fluxo $e^{A\tau}$ e na análise de suas propriedades.

Palavras-chave: Cálculo Funcional; Funções de Matrizes; Espaços Vetoriais de Dimensão Finita.

ABSTRACT

This work consists of a detailed presentation on functional calculus and linear operators defined on finite-dimensional vector spaces, as well as applications of this theory. Functional calculus constitutes an important tool within the Theory of Linear Operators. For finite-dimensional vector spaces V , we know that the set of linear operators from V to V , $\mathcal{L}(V)$, is isomorphic to a space of matrices. In this sense, this calculus allows us to define and compute functions of matrices. Thus, the objective of this monograph is to demonstrate how we compute functions such as e^A , $\sin(A)$ and $\cos(A)$, where A is a square matrix of order n . Furthermore, under some additional conditions on A , more precisely concerning its eigenvalues, it is also possible to define $\log(A)$ and \sqrt{A} . In addition, applications of functional calculus in solving systems of linear ordinary differential equations are explored, particularly in the construction of the flow e^{At} and the analysis of its properties.

Keywords: Functional Calculus; Functions of Matrices; Finite-Dimensional Vector Spaces.

Sumário

Introdução	6
1 Preliminares	8
1.1 Álgebra Linear	8
1.1.1 Matrizes	8
1.1.2 Alguns operadores matriciais	13
1.1.3 Espaços vetoriais	15
1.1.4 Subespaços	16
1.1.5 Bases e dimensão	17
1.1.6 Transformações lineares	19
1.1.7 Representação de transformações lineares por matrizes	22
1.1.8 Autovalores e Autovetores	26
1.1.9 Operadores diagonalizáveis	31
1.2 Análise	33
1.2.1 Séries	33
1.2.2 Definições e resultados auxiliares	36
2 O Cálculo Funcional	38
2.1 Polinômio interpolador	38
2.2 Funções de matrizes	42
2.3 Aplicações do cálculo funcional	46
2.3.1 O fluxo	46
2.3.2 Funções trigonométricas	51
2.3.3 Logaritmo	51
2.3.4 Raiz quadrada	53
2.3.5 A inversa	53
3 Sistemas de Equações Diferenciais Ordinárias Lineares	55
REFERÊNCIAS	63

Introdução

O estudo de funções de matrizes é tão antigo quanto a própria álgebra matricial. A sua origem remonta aos trabalhos pioneiros de matemáticos do século XIX, com destaque para Arthur Cayley. Em sua obra seminal, "A Memoir on the Theory of Matrices" (1858), Cayley não só estabeleceu bases fundamentais para a teoria das matrizes, como investigou um problema específico e não trivial: o cálculo da raiz quadrada de uma matriz. Pouco depois, esta linha de investigação foi ampliada por James Joseph Sylvester e outros contemporâneos, que começaram a propor definições gerais para $f(A)$, onde f é uma função aplicada a uma matriz A .

Inicialmente, sendo um tópico de Matemática Pura, o cálculo funcional matricial se expandiu rapidamente e tornou um pilar da Matemática Aplicada. Hoje, as suas aplicações permeiam diversas áreas da ciência e da engenharia, desde a solução de sistemas de equações diferenciais e equações de Riccati algébricas, até problemas mais recentes em ciência de dados e física computacional. Esta evolução transformou a pesquisa em funções de matrizes interdisciplinar, que combina teoria de matrizes, análise numérica, teoria da aproximação e o desenvolvimento de algoritmos e softwares.

Neste contexto, a presente monografia tem como objetivo apresentar de forma detalhada o cálculo funcional em dimensão finita, com ênfase na definição e no cálculo de funções de matrizes, tais como e^A , $\sin(A)$, $\cos(A)$, $\log(A)$ e \sqrt{A} , quando estas estão bem definidas. Vale ressaltar que a teoria aqui apresentada se restringe ao cálculo funcional de operadores (transformações) lineares definidos em espaços de dimensão finita, ou seja, cálculo de matrizes (lembramos que existe um isomorfismo entre o espaço de operadores e o espaços de matrizes neste caso), mas uma teoria similar é desenvolvida para operadores definidos em espaços de dimensão infinita (com propriedades adicionais sobre o espaço e sobre o operador). Vale ressaltar que em dimensão infinita não tem sentido a identificação de operadores e matrizes, uma vez que não há representação matricial finita para operadores em espaços de dimensão infinita.

Para tanto, o trabalho está estruturado em três capítulos: No Capítulo 1, são revisados os fundamentos de álgebra linear, análise e teoria de operadores necessários para a compreensão do cálculo funcional, incluindo autovalores, autovetores, diagonalização e séries de potências. No Capítulo 2 é introduzido formalmente o cálculo funcional, partindo do polinômio interpolador de Hermite e da noção de funções euclidianas, para então

definir funções de matrizes e explorar aplicações como o fluxo exponencial $e^{A\tau}$, funções trigonométricas, logaritmo, raiz quadrada e a inversa de matrizes. No Capítulo 3, o cálculo funcional é aplicado à resolução de sistemas de equações diferenciais lineares, com destaque para a construção do fluxo $e^{A\tau}$ e da análise de suas propriedades.

Capítulo 1

Preliminares

Neste capítulo apresentaremos alguns resultados básicos de matrizes, álgebra linear, cálculo, análise e variáveis complexas. Todos os resultados discutidos neste capítulo são imprescindíveis para a compreensão do texto.

1.1 Álgebra Linear

Nesta seção apresentaremos alguns resultados de Álgebra Linear que serão ferramentas úteis para os resultados principais deste trabalho. As definições e resultados aqui apresentados foram baseados em [1, 2, 7].

1.1.1 Matrizes

Definição 1. Uma $m \times n$ matriz sobre o corpo F é a imagem de uma função A que associa cada par de números inteiros (i, j) com $1 \leq i \leq m$, $1 \leq j \leq n$ a um elemento do corpo F .

Observação 1. Os elementos (as imagens) da matriz A são chamados de escalares, isto é, são elementos do corpo tais que $A(i, j) = a_{ij} \in F$. É usual descrever uma matriz exibindo seus elementos numa tabela retangular com m linhas e n colunas, como a seguir:

$$A = A_{m \times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Definição 2. Duas matrizes $A_{m \times n} = (a_{ij})_{m \times n}$ e $B_{r \times s} = (b_{ij})_{r \times s}$ são iguais, $A = B$, se $m = r$, $n = s$ e $a_{ij} = b_{ij}$ para todo $i \in \{1, \dots, m\}$ e $j \in \{1, \dots, n\}$.

Exemplo 1.

$$\begin{bmatrix} e^0 & \log 1 \\ 2^2 & \text{sen } 90^\circ \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}.$$

No decorrer deste texto, denotaremos o conjunto das matrizes $m \times n$ sobre o corpo F por $M_{m \times n}(F)$.

Agora, classifiquemos as matrizes quanto ao seu tipo:

- (i) **Matriz quadrada:** É aquela cujo número de linhas é igual ao número de colunas ($m = n$).

Observação 2. No decorrer do trabalho, denotaremos o conjunto das matrizes quadradas sobre o corpo F por $M_n(F)$.

Exemplo 2. $A \in M_{3 \times 3}(\mathbb{R})$, com

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

- (ii) **Matriz nula:** É aquela em que $a_{ij} = 0 \in F$ (zero do corpo), para todo i, j .

Exemplo 3.

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{5 \times 2}.$$

- (iii) **Matriz linha:** É aquela em que possui uma única linha ($m = 1$).

Exemplo 4.

$$A = \begin{bmatrix} 7 & 1 & 2003 \end{bmatrix}_{1 \times 3}.$$

- (iv) **Matriz coluna:** É aquela em que possui uma única coluna ($n = 1$).

Exemplo 5.

$$A = \begin{bmatrix} 24 \\ 4 \\ 1982 \end{bmatrix}_{3 \times 1}.$$

- (v) **Matriz diagonal:** É uma matriz quadrada onde $a_{ij} = 0 \in F$ (zero do corpo), para $i \neq j$. De modo geral,

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} := \text{diag}(a_{11} \cdots a_{nn}).$$

(vi) **Matriz identidade:** Matriz quadrada em que:

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

Ou seja,

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Observação 3. Aqui $1 \in F$ é a unidade do corpo e $0 \in F$ é o zero do corpo.

(vii) **Matriz triangular superior:** Matriz quadrada onde $a_{ij} = 0 \in F$ (zero do corpo), para $i > j$. De modo geral,

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

(viii) **Matriz triangular inferior:** Matriz quadrada onde $a_{ij} = 0$, para $i < j$.

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

(ix) **Matriz simétrica:** Matriz quadrada $A = (a_{ij})_{n \times n}$ tal que $a_{ij} = a_{ji}$ para todo $i, j \in \{1, \dots, n\}$.

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

No conjunto de matrizes $M_{m \times n}(F)$ definimos duas operações como apresentamos a seguir.

Definição 3. Sejam $A, B \in M_{m \times n}(F)$. Define-se a soma de A com B por

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

Dessa forma, temos que a soma é dada por:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Proposição 1. Dadas as matrizes $A, B, C \in M_{m \times n}(F)$, são válidas as seguintes propriedades:

- (i) $A + B = B + A$ (comutatividade);
- (ii) $A + (B + C) = (A + B) + C$ (associatividade);
- (iii) $A + 0 = 0 + A = A$ (elemento neutro, onde 0 denota a matriz nula $0_{m \times n}$);
- (iv) $A + (-A) = -A + A = 0$ (elemento simétrico).

Demonstração: Sejam $A, B, C \in M_{m \times n}$ tais que $A = (a_{ij})$, $B = (b_{ij})$ e $C = (c_{ij})$, por definição temos

- (i) $A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A$, onde usamos a comutatividade do corpo F .
- (ii) $A + (B + C) = (a_{ij} + (b_{ij} + c_{ij})) = ((a_{ij} + b_{ij}) + c_{ij}) = (A + B) + C$, aqui usamos a associatividade dos elementos de F .
- (iii) $A + 0 = (a_{ij} + 0) = (0 + a_{ij}) = 0 + A$.
- (iv) $A + (-A) = (a_{ij} + (-a_{ij})) = (-a_{ij} + a_{ij}) = (0_{ij}) = 0$.

□

Observação 4. Pode-se definir a operação diferença $A - B$ como a matriz resultante da soma de A com a matriz formada pelos elementos simétricos da matriz B .

Definição 4 (Multiplicação por escalar). Seja $A \in M_{m \times n}(F)$ e $\alpha \in F$, define-se o produto por escalar de α pela matriz A por:

$$\alpha \cdot A = (\alpha a_{ij})_{m \times n}.$$

Exemplo 6. A multiplicação por escalar significa multiplicar cada elemento da matriz por esse número (escalar), como mostra abaixo:

$$(-7) \begin{bmatrix} 3 & 7 & 11 \\ 1 & 4 & 9 \end{bmatrix} = \begin{bmatrix} -21 & -49 & -77 \\ -7 & -28 & -63 \end{bmatrix}.$$

Proposição 2. Dadas as matrizes $A, B, C \in M_{m \times n}(F)$ e $\alpha, \beta \in F$, são válidas as seguintes propriedades:

- (i) $\alpha(A + B) = \alpha A + \alpha B$
- (ii) $(\alpha + \beta)A = \alpha A + \beta A$
- (iii) $0 \cdot A = 0_{m \times n}$
- (iv) $\alpha(\beta A) = (\alpha\beta)A$

Demonstração: Sejam $A, B, C \in M_{m \times n}$ e $\alpha, \beta \in F$ tais que $A = a_{ij}$, $B = b_{ij}$ e $C = c_{ij}$, temos que:

- (i) $\alpha(A + B) = \alpha(a_{ij} + b_{ij}) = \alpha a_{ij} + \alpha b_{ij} = \alpha A + \alpha B.$
- (ii) $(\alpha + \beta)A = (\alpha + \beta)a_{ij} = \alpha a_{ij} + \beta a_{ij} = \alpha A + \beta A.$
- (iii) $0 \cdot A = 0_{ij} \cdot a_{ij} = 0_{ij} = 0$
- (iv) $\alpha(\beta A) = \alpha(\beta a_{ij}) = (\alpha\beta)a_{ij} = (\alpha\beta)A$

□

Definição 5. Seja $A \in M_{m \times n}(F)$ e $B \in M_{n \times p}(F)$. O produto da matriz A pela matriz B é a $m \times p$ matriz C cujo elemento i, j é:

$$c_{ij} = \sum_{r=1}^n a_{ir} b_{rj}.$$

Exemplo 7. Dadas as matrizes $A \in M_{2 \times 3}(\mathbb{R})$ e $B \in M_{3 \times 2}(\mathbb{R})$, tal que $A = \begin{bmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \end{bmatrix}$ e

$B = \begin{bmatrix} -3 & -2 \\ 6 & -5 \\ -1 & 4 \end{bmatrix}$, então o produto de AB vai ser a matriz C tal que:

$$C = \begin{bmatrix} 3 \cdot (-3) + 2 \cdot 6 + 1 \cdot (-1) & 3 \cdot (-2) + 2 \cdot (-5) + 1 \cdot 4 \\ 6 \cdot (-3) + 5 \cdot 6 + 4 \cdot (-1) & 6 \cdot (-2) + 5 \cdot (-5) + 4 \cdot 4 \end{bmatrix} = \begin{bmatrix} 2 & -12 \\ 8 & -21 \end{bmatrix}.$$

Proposição 3. Dadas as matrizes $A, B, C \in M_{m \times n}(F)$ e $\alpha \in F$, são válidas as seguintes propriedades:

- (i) $(AB)C = A(BC);$
- (ii) $(A + B)C = AC + BC;$
- (iii) $C(A + B) = CA + CB;$

$$(iv) (\alpha A)B = \alpha(AB).$$

Demonstração: Ver a demonstração do Teorema 49 do capítulo 4 em [8]. □

Observação 5. A multiplicação de matrizes não é comutativa. Quando $AB = BA$, dizemos que A e B comutam.

Definição 6. Dada a matriz $A \in M_{m \times n}(F)$, define-se a transposição de $A = (a_{ij})_{m \times n}$ como sendo a matriz $A^t \in M_{n \times m}(F)$ com $A^t = (a_{ji})_{n \times m}$ tal que $a_{ji} = a_{ij}$.

Exemplo 8. Seja $A \in M_{3 \times 2}(\mathbb{R})$ com $A = \begin{bmatrix} 1 & 6 \\ 6 & -1 \\ 6 & 1 \end{bmatrix}_{3 \times 2}$, então $A^t = \begin{bmatrix} 1 & 6 & 6 \\ 6 & -1 & 1 \end{bmatrix}_{2 \times 3}$.

Proposição 4. Dadas as matrizes $A, B \in M_{m \times n}(F)$ e $\alpha \in F$, são válidas as seguintes propriedades:

- (i) $(A + B)^t = A^t + B^t$;
- (ii) $(\alpha A)^t = \alpha A^t$;
- (iii) $(AB)^t = B^t A^t$;
- (iv) $(A^t)^t = A$.

Demonstração: Ver a demonstração do Teorema 53 do capítulo 4 em [8]. □

Observação 6. Note que, sendo $A \in M_{n \times n}(F)$. Assim A é simétrica se, e somente se, $A^t = A$.

Definição 7. Seja $A \in M_{n \times n}(F)$ e seja $B \in M_{n \times n}(F)$. Se $BA = I_n$, então B é dita uma inversa à esquerda de A . Se $AB = I_n$, então B é dita uma inversa à direita de A . Se $AB = BA = I_n$, então A é dita invertível, com B sendo sua matriz inversa.

1.1.2 Alguns operadores matriciais

Definição 8. Seja $A \in M_{n \times n}(F)$, definimos o operador traço $Tr : M_{n \times n}(F) \rightarrow F$ de A como sendo:

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

Exemplo 9. Seja $A \in M_{3 \times 3}(F)$, tal que $A = \begin{bmatrix} -3 & 3 & 1 \\ 5 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, temos que o $Tr(A) = (-3) + 2 + 1 = 0$.

Proposição 5. Dadas as matrizes $A, B \in M_{n \times n}(F)$ e $\alpha \in F$, são válidas as seguintes propriedades:

- (i) $Tr(A + B) = Tr(A) + Tr(B)$;
- (ii) $Tr(\alpha A) = \alpha Tr(A)$;
- (iii) $Tr(AB) = Tr(BA)$;
- (iv) $Tr(A) = Tr(A^t)$.

Demonstração: Sejam $A, B \in \mathcal{M}_{m \times n}$ e $\alpha \in F$ tais que $A = a_{ij}$ e $B = b_{ij}$.

- (i) $Tr(A + B) = \sum_{i=1}^n (a_{ii} + b_{ii}) = \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} = Tr(A) + Tr(B)$;
- (ii) $Tr(\alpha A) = \sum_{i=1}^n \alpha a_{ii} = \alpha \sum_{i=1}^n a_{ii} = \alpha Tr(A)$;
- (iii) $Tr(AB) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}$ e $Tr(BA) = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ik}$.

Como em F a multiplicação é comutativa, $a_{ik} b_{ki} = b_{ki} a_{ik}$, assim

$$Tr(BA) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}.$$

Portanto, $Tr(AB) = Tr(BA)$.

- (iv) $Tr(A^t) = \sum_{i=1}^n [A^t]_{ii} = \sum_{i=1}^n a_{ii} = Tr(A)$.

□

Definição 9. Dada uma permutação dos inteiros $1, 2, \dots, n$, dizemos que existe uma inversão quando um inteiro precede outro menor que ele, mais especificamente, há uma inversão quando para um par de posições (i, j) em que $i < j$, tem-se que o valor m posição i é maior que o da posição j .

Exemplo 10. Consideremos as permutações de $1, 2, 3$ e vejamos em cada permutação o número de inversões.

Tabela 1.1: Permutações de $1, 2, 3$ e número de inversões

Permutação	Número de inversões
(1 2 3)	0
(1 3 2)	1
(2 1 3)	1
(2 3 1)	2
(3 1 2)	2
(3 2 1)	3

Fonte: Boldrini (1980, p. 66).

Definição 10. O determinante de uma matriz $A = (a_{ij})_{m \times n}$ é definido por: $\det[a_{ij}] = \sum_{\rho} (-1)^J a_{1j_1} a_{2j_2} \dots a_{nj_n}$ onde $J = J(j_1, \dots, j_n)$ é o número de inversões da permutação (j_1, j_2, \dots, j_n) , e ρ indica que a soma é estendida a todas as $n!$ permutações de $(1, 2, \dots, n)$.

Propriedades:

- (i) Se todos os elementos de uma linha (coluna) de uma matriz A são nulos, $\det A = 0$.
- (ii) $\det A = \det A^t$.
- (iii) Se multiplicarmos uma linha da matriz por uma constante, o determinante fica multiplicado por esta constante.
- (iv) Uma vez trocada a posição de duas linhas, o determinante troca de sinal.
- (v) O determinante de uma matriz que tem duas linhas (colunas) iguais é zero.
- (vi)
$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ b_{i1} + c_{i1} & \dots & b_{in} + c_{in} \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ b_{i1} & \dots & b_{in} \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ c_{i1} & \dots & c_{in} \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$
- (vii) O determinante não se altera se somarmos a uma linha outra linha multiplicada por uma constante.
- (viii) $\det(A \cdot B) = \det A \cdot \det B$.

1.1.3 Espaços vetoriais

Definição 11. Seja F um corpo e V um conjunto não vazio. Sobre o conjunto V defina-se duas operações

$$(\text{adição}) \quad + : V \times V \rightarrow V \quad (\text{produto por escalar}) \quad \cdot : F \times V \rightarrow V.$$

A terna $(V, +, \cdot)$ é um espaço vetorial sobre F se os axiomas a seguir são verificados:

- (i) A adição é comutativa, $u + v = v + u$, para todos $u, v \in V$;
- (ii) A adição é associativa, $u + (v + w) = (u + v) + w$, para todos $u, v, w \in V$;
- (iii) Existe um único vetor 0 em V , denominado o vetor nulo, tal que $u + 0 = u$ para todo u em V ;
- (iv) Para cada vetor u em V existe um único vetor $-u$ em V tal que $u + (-u) = 0$;
- (v) $1 \cdot u = u$ para todo $u \in V$;
- (vi) $(\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot u)$, para todo $u \in V$ e $\alpha, \beta \in F$;

(vii) $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$, para todos $u, v \in V$ e $\alpha \in F$;

(viii) $(\alpha + \beta) \cdot u = \alpha u + \beta u$, para todo $u \in V$ e $\alpha, \beta \in F$.

Neste caso, os elementos de V são chamados de vetores.

Exemplo 11. Seja V o conjunto de todas as $m \times n$ matrizes sobre o corpo F . Pode-se definir naturalmente em V a adição como apresentado na Definição 3 e o produto por escalar como na Definição 4. Combinando as Proposições 1 e 2, tem-se que $(M_{m \times n}(F), +, \cdot)$, com estas operações, é um espaço vetorial sobre F .

Tendo em vista as operações em V , uma forma de encontrarmos novos vetores em V é fazermos a adição de produtos por escalar de uma quantidade finita de vetores V .

Definição 12. Um vetor $v \in V$ é dito uma combinação linear dos vetores $u_1, \dots, u_n \in V$ se existirem escalares $\alpha_1, \dots, \alpha_n \in F$ tais que:

$$v = \alpha_1 u_1 + \dots + \alpha_n u_n = \sum_{i=1}^n \alpha_i u_i.$$

1.1.4 Subespaços

Definição 13. Seja $(V, +, \cdot)$ um espaço vetorial sobre o corpo F . Um subconjunto $W \subset V$ é um subespaço V , se $(W, +, \cdot)$ é um espaço vetorial sobre F .

Teorema 1. Um subconjunto $\emptyset \neq W \subset V$ é um subespaço de V se, e somente se, para cada $u, v \in W$ e cada escalar $\alpha \in F$, o vetor $\alpha u + v \in W$.

Demonstração: Seja $\emptyset \neq W \subset V$ tal que $\alpha u + v \in W$ para todos $u, v \in W$ e todo $\alpha \in F$. Sendo $W \subset V$ e V um espaço vetorial, as propriedades (i), (ii), (v), (vi), (vii) e (viii) da definição 11 valem para todos os vetores de V e escalares de F , conseqüentemente valem para os vetores de W (são herdados de V). Portanto, os únicos axiomas que precisam de demonstração são (iii) e (iv), além de verificar que as operações $+$ e \cdot estão bem definidas em W .

Temos que o vetor nulo $0 \in V$ está em W , pois sendo W não vazio, existe $p \in W$. Assim, tome $\alpha = -1$ e $u = v = p$. Então,

$$(-1)p + p = -p + p = 0 \in W.$$

Seja $u \in W$ e $\alpha \in F$. Tome $v = 0 \in W$. Então,

$$\alpha u + 0 = \alpha u \in W.$$

Em particular, se $\alpha = -1$, então $-u \in W$ para cada $u \in W$. Se $u, v \in W$, tome $\alpha = 1$, então

$$1 \cdot u + v = u + v \in W.$$

Portanto, W é um subespaço de V .

Reciprocamente, se W é subespaço de V , então é fechado sob adição e multiplicação escalar, já que é um espaço vetorial sobre F . Logo, para quaisquer $u, v \in W$ e $\alpha \in F$, temos $\alpha u \in W$ e $\alpha u + v \in W$. \square

Exemplo 12. Sendo $V = M_{n \times n}(F)$ e $S = \{A \in M_{n \times n}(F); A^t = A\}$ (conjunto das matrizes simétricas). Como consequência da Proposição 4 tem-se que S é um subespaço de V .

1.1.5 Bases e dimensão

Nesta seção estamos interessados em apresentar um modo de descrevermos os elementos de um espaço vetorial V conhecendo-se um subconjunto de V . Assim, buscamos descrever o "menor" subconjunto para o qual conseguimos descrever todos os elementos de V .

Definição 14. Seja V um espaço vetorial sobre F . Um subconjunto $S \subset V$ é dito linearmente dependente se existem vetores distintos $u_1, \dots, u_n \in S$ e escalares $\alpha_1, \dots, \alpha_n \in F$, não todos nulos, tais que

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0.$$

Um conjunto que não é linearmente dependente é dito linearmente independente.

Observação 7. Um conjunto $S \subset V$ de vetores é linearmente independente se e somente se todo subconjunto finito de S é linearmente independente, isto é, se e somente se para quaisquer vetores distintos $u_1, \dots, u_n \in S$ $\alpha_1 u_1 + \dots + \alpha_n u_n = 0$ implica que cada $\alpha_i = 0$.

Definição 15. Seja V um espaço vetorial sobre F . Um subconjunto $S \subset V$ é uma base de V , se S é um conjunto linearmente independente de vetores em V e se qualquer elemento de V é descrito como combinação linear dos elementos de S . Neste caso dizemos que S gera o espaço vetorial V .

Exemplo 13. Seja $V = M_{2 \times 2}(F)$ e

$$S = \left\{ e_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, e_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, e_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Sejam $\alpha_1, \dots, \alpha_4$ escalares em F e $u = \begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} \in M_{2 \times 2}(F)$. Note que $u = \alpha_1 e_1 + \dots + \alpha_4 e_4$. Isto mostra que S gera $M_{2 \times 2}(F)$. Como $u = 0 \in M_{2 \times 2}(F)$ se e somente se $\alpha_1 = \dots = \alpha_4 = 0$, os vetores e_1, \dots, e_4 são linearmente independentes. O conjunto S é, portanto, uma base de V .

Observação 8. A definição de base se aplica a qualquer espaço vetorial. Para todo o espaço das matrizes $M_{m \times n}(F)$, basta generalizarmos o exemplo anterior.

Teorema 2. Seja V um espaço vetorial sobre o corpo F gerado por um conjunto finito de vetores $\{v_1, \dots, v_m\}$, isto é, $V = \{\alpha_1 v_1 + \dots + \alpha_n v_n; \alpha_1, \dots, \alpha_n \in F\}$. Então, todo conjunto independente de vetores em V é finito e contém no máximo m elementos.

Demonstração: Ver a demonstração do Teorema 4 em [7]. □

O Teorema 2 estabelece uma condição sobre a quantidade de vetores que deve compor uma base V (no caso em que V é gerado por uma quantidade finita de vetores). Neste sentido, somos motivados a estudar esta quantidade.

Definição 16. Um espaço vetorial V é de dimensão finita se ele possui uma base finita.

Corolário 1. Se V é um espaço vetorial de dimensão finita, então duas quaisquer bases de V tem o mesmo número (finito) de elementos.

Demonstração: Como V tem dimensão finita, admite uma base finita $\mathcal{B} = \{v_1, v_2, \dots, v_m\}$. E seja $\mathcal{B}' = \{u_1, u_2, \dots, u_n\}$ outra base de V . Pelo Teorema 2, toda base de V é finita e contém no máximo m elementos. Assim, $n \leq m$. Pela mesma razão, $m \leq n$. Logo $m = n$. □

Definição 17. Se V é um espaço vetorial sobre F de dimensão finita, a dimensão de V é definida como sendo o número de elementos de uma base de V . Indicaremos a dimensão de um espaço vetorial V de dimensão finita por $\dim_F V$ ou $\dim V$.

Exemplo 14. O espaço vetorial matrizes $M_{m \times n}(F)$ tem dimensão $m \cdot n$, pois a base canônica de $M_{m \times n}(F)$ contém $m \cdot n$ vetores.

Corolário 2. Seja V um espaço vetorial sobre F n -dimensional (ou seja, $\dim_F V = n$). Então:

- (a) Todo conjunto de vetores em V que contém mais de n vetores é linearmente dependente.
- (b) Nenhum conjunto contendo menos de n de vetores pode gerar V .

Demonstração: Ver a demonstração em [7]. □

Corolário 3. Num espaço vetorial V de dimensão finita todo conjunto não vazio de vetores linearmente independentes é parte de uma base.

Demonstração: Ver a demonstração do Corolário 1 em [7]. □

Corolário 4. Seja $A \in M_{n \times n}(F)$ e suponhamos que os vetores linhas de A formem um conjunto de vetores de F^n linearmente independentes. Então A é invertível.

Demonstração: Ver demonstração do Corolário 3 em [7]. □

1.1.6 Transformações lineares

Definição 18. Sejam V e W espaços vetoriais sobre o corpo F . Uma transformação linear de V em W é uma aplicação T de V em W tal que:

$$T(\alpha u + v) = \alpha T(u) + T(v)$$

para todos $u, v \in V$ e todos escalares $\alpha \in F$.

Observação 9. Qualquer que seja a transformação linear $T : V \rightarrow W$ vale $T(0) = 0$. Com efeito, $T(0) = T(0 + 0) = T(0) + T(0) \implies T(0) = 0$. Isto estabelece uma condição para que uma aplicação seja uma transformação linear.

Exemplo 15. Considere o espaço vetorial $M_{n \times n}(F)$, o traço $Tr : M_{n \times n}(F) \rightarrow F$, é uma transformação linear. De fato, sejam $A, B \in M_{n \times n}(F)$ e $\alpha \in F$, pela Proposição 5,

$$\begin{aligned} Tr(\alpha A + B) &= \sum_{i=1}^n \alpha a_{ii} + b_{ii} \\ &= \alpha \sum_{i=1}^n a_{ii} + \sum_{i=1}^n b_{ii} \\ &= \alpha Tr(A) + Tr(B). \end{aligned}$$

Observação 10. O conjunto de todas as transformações lineares de V em W será denotado por $\mathcal{L}(V, W)$. Por convenção, o conjunto $\mathcal{L}(V, V)$ será denotado por $\mathcal{L}(V)$.

Teorema 3. Seja V um espaço vetorial de dimensão finita sobre o corpo F e seja $\{u_1, \dots, u_n\}$ uma base ordenada de V . Seja W um espaço vetorial sobre o mesmo corpo F e sejam v_1, \dots, v_n vetores arbitrários em W . Então, existe exatamente uma transformação linear $T : V \rightarrow W$ tal que:

$$T(u_j) = v_j, \quad j = 1, \dots, n.$$

Demonstração: Ver demonstração do Teorema 1 em [7]. □

Teorema 4. Sejam V e W espaços vetoriais sobre o corpo F e seja T uma transformação linear $T : V \rightarrow W$. Então, a imagem de T é um subespaço de W . O conjunto de vetores u em V tais que $T(u) = 0$ também é um subespaço de V , dito o núcleo de T .

Demonstração: Tome $w_1, w_2 \in Im(T)$. Note que existem $v_1, v_2 \in V$ tal que $T(v_1) = w_1$ e $T(v_2) = w_2$. Assim $\alpha w_1 + w_2 = \alpha T(v_1) + T(v_2) = T(\alpha v_1 + v_2)$. Pelo Teorema 1, é um subespaço vetorial de W .

Agora, tome u, v pertencentes ao núcleo de T . Segue que $T(u) = 0$ e $T(v) = 0$. Logo, $T(\alpha u + v) = \alpha T(u) + T(v) = \alpha \cdot 0 + 0 = 0$. Novamente, pelo Teorema 1, o núcleo é um subespaço vetorial de V . □

Definição 19. Seja T uma transformação linear de V em W , sendo V de dimensão finita. O posto de T é a dimensão da imagem de T . A nulidade de T é a dimensão do núcleo de T .

Teorema 5 (do núcleo e da imagem). Sejam V e W espaços vetoriais de dimensão finita sobre o corpo F e seja T uma transformação linear de V em W . Suponhamos que V seja de dimensão finita. Então

$$\text{posto}(T) + \text{nulidade}(T) = \dim V.$$

Demonstração: Seja n a dimensão de V e $\{u_1, \dots, u_k\}$ uma base para o núcleo de T . Note que a dimensão do núcleo é k . Podemos completar o conjunto com $n - k$ vetores linearmente independentes para obtermos uma base para V , ou seja, considere $\{u_{k+1}, \dots, u_n\}$. Mostraremos que $S = \{Tu_{k+1}, \dots, Tu_n\}$ é uma base da $\text{Im}(T)$. Para isso, precisamos mostrar que S gera $\text{Im}(T)$ e, é um conjunto linearmente independentes. Como $\{u_1, \dots, u_n\}$ é uma base de V , podemos escrever qualquer $v \in V$ como combinação linear, assim, $v = \alpha_1 u_1 + \dots + \alpha_n u_n$. Aplicando a T e usando a linearidade, obtemos: $T(v) = \alpha_1 T(u_1) + \dots + \alpha_n T(u_n)$. Mas perceba que, para cada $j \in \{1, \dots, k\}$, $T(u_j) = 0$, pois $\{u_1, \dots, u_k\}$ pertencem ao núcleo. Dessa forma, $T(v) = \alpha_{k+1} T(u_{k+1}) + \dots + \alpha_n T(u_n)$, ou seja, $\{Tu_{k+1}, \dots, Tu_n\}$ geram a imagem de T . Agora, vamos verificar que $\{Tu_{k+1}, \dots, Tu_n\}$ são linearmente independentes. Suponha que existam escalares c_{k+1}, \dots, c_n tais que $c_{k+1} T(u_{k+1}) + \dots + c_n T(u_n) = 0$. Pela linearidade de T , $T(c_{k+1} u_{k+1} + \dots + c_n u_n) = 0$. Assim, $u = c_{k+1} u_{k+1} + \dots + c_n u_n$ pertence ao núcleo. Como o núcleo tem base $\{u_1, \dots, u_k\}$, existem escalares b_1, \dots, b_k tais que $u = b_1 u_1 + \dots + b_k u_k$. Temos, assim, que $c_{k+1} u_{k+1} + \dots + c_n u_n = u = b_1 u_1 + \dots + b_k u_k \implies c_{k+1} u_{k+1} + \dots + c_n u_n - b_1 u_1 - \dots - b_k u_k = 0$. Como $\{u_1, \dots, u_n\}$ é uma base de V , então $b_1 = \dots = b_k = c_{k+1} = \dots = c_n = 0$, assim $T(u_{k+1}), \dots, T(u_n)$ são linearmente independentes e, portanto, uma base da imagem de T . Dessa forma, obtemos que $(n - k) + k = n = \dim V$. \square

Teorema 6. Sejam V e W espaços vetoriais de dimensão finita sobre o corpo F e seja T uma transformação linear de V em W . Se T é bijetora então a função inversa $T^{-1} : W \rightarrow V$ é uma transformação linear.

Demonstração: Sejam v_1 e v_2 vetores em W e seja $\alpha \in F$. Queremos mostrar que

$$T^{-1}(\alpha v_1 + v_2) = \alpha T^{-1} v_1 + T^{-1} v_2.$$

Como T é bijetora, existem únicos $u_1, u_2 \in V$ tais que $T(u_1) = v_1$ e $T(u_2) = v_2$, ou seja, $u_1 = T^{-1} v_1$ e $u_2 = T^{-1} v_2$. Pela linearidade de T ,

$$T(\alpha u_1 + u_2) = \alpha T u_1 + T u_2 = \alpha v_1 + v_2.$$

Assim $\alpha u_1 + u_2$ é o único vetor em V que é levado por T em $\alpha v_1 + v_2$, portanto

$$T^{-1}(\alpha v_1 + v_2) = \alpha u_1 + u_2 = \alpha(T^{-1}v_1) + T^{-1}v_2$$

e T^{-1} é linear. □

Observação 11. Seja T uma transformação linear de V em W . Se o núcleo de T tiver somente o vetor nulo, diremos que T é não singular. Note que T é injetora, tome $u, v \in V$ tais que $Tu = Tv$. Como T é uma aplicação linear, temos $Tu - Tv = 0 = T(u - v) = 0$. Logo $u - v$ pertence ao núcleo, assim $u - v = 0 \implies u = v$. Deste modo, T é invertível (bijetora) se e somente se T é não singular e a imagem é todo o espaço W .

Teorema 7. Seja T uma transformação linear de V em W . Então T é não singular se, e somente se, T leva todo subconjunto linearmente independente de V sobre um conjunto linearmente independente de W .

Demonstração: (\implies) Suponhamos primeiro que T seja não singular. Seja S um subconjunto linearmente independente de V . Tomemos $u_1, \dots, u_k \in S$ e suponha que

$$\alpha_1 T(u_1) + \dots + \alpha_k T(u_k) = 0$$

para escalares $\alpha_1, \dots, \alpha_k \in F$. Pela linearidade de T ,

$$T(\alpha_1 u_1 + \dots + \alpha_k u_k) = 0.$$

Como T é não singular, $\alpha_1 u_1 + \dots + \alpha_k u_k = 0$. Mas S é linearmente independente, logo $\alpha_1 = \dots = \alpha_k = 0$. Portanto, $\{T(u_1), \dots, T(u_k)\}$ é linearmente independente.

(\impliedby) Suponhamos que T leva subconjuntos linearmente independentes de V em subconjuntos linearmente independentes de W . Seja $u \in V$ não nulo. O conjunto $\{u\}$ é linearmente independente. Então $\{T(u)\}$ é linearmente independente em W , o que implica $T(u) \neq 0$. Assim, o único vetor que T leva em 0 é o vetor nulo. Logo, T é não singular. □

Teorema 8. Sejam V e W espaços vetoriais de dimensão finita sobre o corpo F tais que $\dim V = \dim W$. Se T é uma transformação linear de V em W , as seguintes afirmações são equivalentes:

- (i) T é invertível.
- (ii) T é não singular.
- (iii) A imagem de T é W .
- (iv) Se $\{u_1, \dots, u_n\}$ é uma base arbitrária de V , então $\{Tu_1, \dots, Tu_n\}$ é uma base de W .

(v) Existe pelo menos uma base $\{u_1, \dots, u_n\}$ de V tal que $\{Tu_1, \dots, Tu_n\}$ seja uma base de W .

Demonstração: Ver demonstração do Teorema 9 em [7]. □

Definição 20. Dois espaços vetoriais V e W são ditos isomorfos se existe uma transformação linear $T : V \rightarrow W$ bijetora.

Teorema 9. Dois espaços vetoriais de dimensão finita V e W são isomorfos se, e somente se, $\dim V = \dim W$.

Demonstração: (\Rightarrow) Sendo V e W isomorfos, existe uma transformação linear $T : V \rightarrow W$ bijetora. Pelo Teorema 5 e pela Observação 11, segue que $\dim V = \text{Nulidade}(T) + \text{posto}(T) = 0 + \dim W = \dim W$.

(\Leftarrow) Como $\dim V = \dim W$ as bases de V e W têm a mesma quantidade de vetores. Assim, sejam $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ bases de V e W respectivamente. Pelo Teorema 3 existe uma transformação linear $T : V \rightarrow W$ tal que, para cada $i \in \{1, \dots, n\}$, $T(v_i) = w_i$. Sabemos que, para cada $v \in V$, $v = \sum_{i=1}^n \alpha_i v_i$, assim

$$T(v) = \sum_{i=1}^n \alpha_i T(v_i) = \sum_{i=1}^n \alpha_i w_i. \quad (1.1)$$

Agora, seja $v = \sum_{i=1}^n \alpha_i v_i \in V$ tal que $0 = T(v) = \sum_{i=1}^n \alpha_i w_i$ (onde usamos a identidade (1.1)). Daí concluímos que $\alpha_1 = \dots = \alpha_n = 0$, pois $\{w_1, \dots, w_n\}$ é uma base de W , então $v = 0$. Pela Observação 11 segue que T é não singular e por consequência do Teorema 8 segue que T é uma bijeção. □

Corolário 5. Todo espaço vetorial n -dimensional sobre o corpo F é isomorfo ao espaço $F^n := \{(\alpha_1, \dots, \alpha_n); \alpha_i \in F, \forall i = 1, \dots, n\}$.

Demonstração: Sendo V um espaço vetorial sobre F tal que $\dim V = n$. Como $\dim F^n = n$, segue do Teorema 9 segue que V e F^n são isomorfos. □

1.1.7 Representação de transformações lineares por matrizes

As transformações lineares definidas entre espaços vetoriais de dimensão finita e as matrizes possuem uma estreita relação. Nesta seção apresentaremos esta relação. Mais precisamente, veremos que as transformações lineares podem ser identificadas com matrizes.

Seja V um espaço vetorial n -dimensional sobre o corpo F e seja W um espaço vetorial m -dimensional sobre F . Fixemos $\mathcal{B} = \{u_1, \dots, u_n\}$ uma base ordenada de V e $\mathcal{B}' =$

$\{v_1, \dots, v_m\}$ uma base ordenada de W . Seja $T : V \rightarrow W$ uma transformação linear arbitrária, então T é determinada por sua ação sobre os vetores u_j (veja o Teorema 3), isto é, como cada Tu_j é um vetor de W , podemos escrever em termos de \mathcal{B}' , de maneira única, como uma combinação linear:

$$Tu_j = \sum_{i=1}^m A_{ij}v_i$$

onde os escalares A_{ij}, \dots, A_{mj} são as coordenadas de Tu_j na base \mathcal{B}' e que definem uma matriz $A_{m \times n}$, denominada matriz de T em relação às bases \mathcal{B} e \mathcal{B}' . Agora, nosso intuito é entender como essa matriz A determina a transformação linear T .

Para qualquer $v \in V$, podemos escrevê-lo como a seguinte combinação linear

$$v = \alpha_1u_1 + \dots + \alpha_nu_n,$$

onde $\alpha_1, \dots, \alpha_n \in F$.

Agora, aplicando a T , obtemos

$$\begin{aligned} Tv &= T\left(\sum_{j=1}^n \alpha_j u_j\right) \\ &= \sum_{j=1}^n \alpha_j (Tu_j) \\ &= \sum_{j=1}^n \alpha_j \sum_{i=1}^m A_{ij} v_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n A_{ij} \alpha_j v_i\right). \end{aligned}$$

Sendo X é a matriz das coordenadas de v em relação à base ordenada \mathcal{B} , o cálculo acima mostra que AX é a matriz das coordenadas do vetor Tv em relação à base ordenada \mathcal{B}' , uma vez que o escalar

$$\sum_{j=1}^n A_{ij} v_j$$

é o elemento da i -ésima linha da matriz coluna AX . Observemos também que se A é uma matriz $m \times n$ arbitrária sobre o corpo F então

$$T\left(\sum_{j=1}^n \alpha_j v_j\right) = \sum_{i=1}^m \beta_i \left(\sum_{j=1}^n A_{ij} \alpha_j\right)$$

define uma transformação linear T de V em W , cuja matriz é A , em relação a \mathcal{B} e \mathcal{B}' . Resumindo formalmente:

Teorema 10. Seja V um espaço vetorial n -dimensional sobre o corpo F e W um espaço vetorial m -dimensional sobre F . Seja \mathcal{B} uma base ordenada de V e \mathcal{B}' uma base ordenada de W . Para cada transformação linear T de V em W , existe uma $m \times n$ matriz A sobre o corpo F , a matriz de T em relação a \mathcal{B} e \mathcal{B}' , tal que

$$[Tv]_{\mathcal{B}'} = A[v]_{\mathcal{B}}$$

Para todo vetor v em V .

Teorema 11. Seja V um espaço vetorial n -dimensional sobre o corpo F e seja W um espaço vetorial m -dimensional sobre o corpo F . Para cada par de bases ordenadas, \mathcal{B} e \mathcal{B}' de V e W , respectivamente. O operador $\Phi_{\mathcal{B},\mathcal{B}'} : \mathcal{L}(V, W) \rightarrow M_{m \times n}(F)$, $\Phi_{\mathcal{B},\mathcal{B}'}(T) = A$, é linear e bijetor, onde A é matriz dada pelo Teorema 10.

Demonstração: Ver demonstração do Teorema 12 em [7]. □

Estaremos particularmente interessados na representação por matrizes de transformações lineares $T : V \rightarrow V$. Neste caso, é mais conveniente usar a mesma base ordenada em cada caso, isto é, tomar $\mathcal{B} = \mathcal{B}'$. A matriz representante será então denominada simplesmente a matriz de T em relação à base ordenada \mathcal{B} . Como este conceito será muito importante para nós, recordaremos sua definição.

Se T é um operador linear sobre o espaço vetorial V de dimensão finita e $\mathcal{B} = \{u_1, \dots, u_n\}$ é uma base ordenada de V , a matriz de T em relação a \mathcal{B} é a $n \times n$ matriz A cujos elementos A_{ij} são definidos pelas equações

$$Tu_j = \sum_{i=1}^n A_{ij}u_i, \quad j = 1, \dots, n.$$

Deve-se ter sempre em mente que esta matriz que representa T depende da base ordenada \mathcal{B} e que existe uma matriz que representa T em relação a cada base ordenada de V . Para não esquecermos esta dependência, usaremos a notação $[T]_{\mathcal{B}}$ para a matriz do operador linear T em relação à base ordenada \mathcal{B} .

A maneira como esta matriz e a base ordenada descrevem T é que, para cada u em V ,

$$[Tu]_{\mathcal{B}} = [T]_{\mathcal{B}}[u]_{\mathcal{B}}.$$

Exemplo 16. Seja F um corpo e seja T o operador sobre F^2 definido por

$$T(x_1, x_2) = (x_1, 0).$$

Pode-se verificar que T é um operador linear sobre F^2 . Seja \mathcal{B} a base ordenada canônica

de F^2 , $\mathcal{B} = \{e_1, e_2\}$ em que $e_1 = (1, 0)$ e $e_2 = (0, 1)$. Ora,

$$\begin{cases} Te_1 = T(1, 0) = (1, 0) = 1e_1 + 0e_2 \\ Te_2 = T(0, 1) = (0, 0) = 0e_1 + 0e_2 \end{cases}$$

de modo que a matriz de T em relação à base ordenada \mathcal{B} é

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Teorema 12. Seja V um espaço vetorial de dimensão finita sobre o corpo F e sejam $\mathcal{B} = \{u_1, \dots, u_n\}$ e $\mathcal{B}' = \{u'_1, \dots, u'_n\}$ bases ordenadas de V . Seja $T : V \rightarrow V$ uma transformação linear e $P \in M_{n \times n}(F)$ a matriz que exprime as coordenadas de cada vetor de V em relação a \mathcal{B} em termos de suas coordenadas em relação a \mathcal{B}' , P é chamada matriz de mudança de base. Então

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P.$$

Demonstração: Ver a demonstração do Teorema 14 em [7]. □

Exemplo 17. Seja T o operador linear sobre \mathbb{R}^2 definido por $T(x_1, x_2) = (x_1, 0)$. No Exemplo 16 mostramos que a matriz de T em relação à base ordenada canônica $\mathbf{a} = \{e_1, e_2\}$ é

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Suponhamos que \mathcal{B}' seja a base ordenada de \mathbb{R}^2 formada pelos vetores $e'_1 = (1, 1)$, $e'_2 = (2, 1)$. Então

$$\begin{cases} e'_1 = e_1 + e_2 \\ e'_2 = 2e_1 + e_2 \end{cases}$$

de modo que P é a matriz

$$P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

Efetuada cálculos simples obtemos

$$P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}.$$

Assim

$$\begin{aligned}
 [T]_{\mathcal{B}'} &= P^{-1}[T]_{\mathcal{B}}P \\
 &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} -1 & -2 \\ 1 & 2 \end{bmatrix}.
 \end{aligned}$$

Podemos verificar facilmente que isto está correto, pois,

$$\begin{cases} Te'_1 = (1, 0) = -e'_1 + e'_2 \\ Te'_2 = (2, 0) = -2e'_1 + 2e'_2. \end{cases}$$

Definição 21. Sejam $A, B \in M_{n \times n}(F)$. Dizemos que a matriz B é semelhante a matriz A se existe uma matriz invertível $P \in M_{n \times n}(F)$ tal que $B = P^{-1}AP$. Neste caso, denotamos $A \sim B$.

Observação 12.

1. A relação \sim é uma relação de equivalência.
2. Sejam V um espaço de dimensão finita e $T : V \rightarrow V$ uma transformação linear. Pelo Teorema 12, $[T]_{\mathcal{B}'} \sim [T]_{\mathcal{B}}$ para quaisquer que sejam as bases \mathcal{B} e \mathcal{B}' de V .

1.1.8 Autovalores e Autovetores

Definição 22. Seja $T : V \rightarrow V$ uma aplicação linear. Se existirem $0 \neq v \in V$ e $\lambda \in F$ tais que $T(v) = \lambda v$, então v é chamado de **autovetor** de T e λ é chamado de **autovalor** de T . O conjunto de todos os autovalores de T é chamado **espectro** de T e denotado por $\sigma(T)$.

Teorema 13. Dada uma aplicação linear $T : V \rightarrow V$ e um autovetor v associado a um autovalor λ , qualquer vetor $w = \alpha v$ ($\alpha \neq 0$) também é autovetor de T associado a λ .

Demonstração: Se $Tv = \lambda v$ e $w = \alpha v$, tem-se $Tw = T(\alpha v) = \alpha Tv = \alpha(\lambda v) = \lambda(\alpha v) = \lambda w$. □

Assim, sendo λ autovalor de T o Teorema 13 estabelece que o conjunto $\{v \in V \mid T(v) = \lambda v\} \subset V$ é um subespaço.

Definição 23. O subespaço $V_\lambda = \{v \in V; T(v) = \lambda v\} \subset V$ é chamado o subespaço associado ao autovalor λ ou autoespaço associado a λ .

Observação 13. O subespaço $V_\lambda \subset V$ é invariante por T , ou seja, $T(V_\lambda) \subset V_\lambda$.

Teorema 14. Autovalores dois a dois distintos da mesma aplicação linear correspondem autovetores linearmente independentes.

Demonstração: Ver a demonstração do Teorema 5.1.5 em [2]. □

Observação 14 (Autovalor e autovetor de uma matriz). Dada uma matriz $A \in M_{n \times n}(\mathbb{R})$, estaremos entendendo por autovalor e autovetor de A , os autovalores e autovetores da transformação linear $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$, tal que a matriz A em relação às bases canônicas é A , isto é, $T_A(v) = A \cdot v$ (na forma coluna). Assim, um autovalor $\lambda \in \mathbb{R}$ de A , e um autovetor $v \in \mathbb{R}^n$, são soluções da equação $A \cdot v = \lambda v$, $v \neq 0$. Considere a matriz diagonal

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} = \text{diag}(a_{11}, \dots, a_{nn}).$$

e dados os vetores $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 0, 1)$, temos

$$A \cdot e_1 = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = a_{11}e_1 \quad \text{e em geral,}$$

$A \cdot e_i = a_{ii}e_i$. Então, estes vetores da base canônica de \mathbb{R}^n são autovetores para A , e o autovetor e_i é associado ao autovalor a_{ii} .

Corolário 6. Seja V um espaço vetorial n -dimensional. Se a aplicação linear $T : V \rightarrow V$ possui n autovalores distintos, então existe uma base de V tal que a matriz de T relativa a essa base é diagonal.

Demonstração: Seja $\mathcal{B} = \{v_1, \dots, v_n\}$ um conjunto tal que $Tv_1 = \lambda_1 v_1, \dots, Tv_n = \lambda_n v_n$, com $\lambda_i \neq \lambda_j, i \neq j, i, j \in \{1, \dots, n\}$. Do Teorema 14, \mathcal{B} é linearmente independente. Como $\dim V = n$, o conjunto \mathcal{B} é uma base (n elementos linearmente independentes). Note que $Tv_1 = \lambda_1 v_1 = \lambda_1 v_1 + 0v_2 + \dots + 0v_n, \dots, Tv_n = \lambda_n v_n = 0v_1 + \dots + \lambda_n v_n$. Note que $[T(v_1)]_{\mathcal{B}} = (\lambda_1, 0, \dots, 0), \dots, [T(v_n)]_{\mathcal{B}} = (0, 0, \dots, \lambda_n)$. Segue que

$$[T]_{\mathcal{B}} = \left(\begin{array}{cccc} [T(v_1)]_{\mathcal{B}} & [T(v_2)]_{\mathcal{B}} & \cdots & [T(v_n)]_{\mathcal{B}} \end{array} \right) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

□

Teorema 15. Seja V um espaço vetorial de dimensão finita, $T : V \rightarrow V$ linear e A uma matriz associada a T em relação a uma base qualquer V . As seguintes condições são equivalentes:

- (i) λ é um autovalor de T .
- (ii) O operador $(T - \lambda I)$ é não invertível.
- (iii) $\det(\lambda I - A) = 0$.

Demonstração: Vamos mostrar que $(i) \implies (ii)$, $(ii) \implies (iii)$ e $(iii) \implies (i)$.

$(i) \implies (ii)$

Suponha que λ é um autovalor de T . Então, por definição, existe um vetor não nulo $u \in V$ tal que $Tu = \lambda u$. Reescrevendo: $Tu - \lambda u = 0 \implies (T - \lambda I)u = 0$. Ou seja, $u \neq 0$ pertence ao núcleo de $T - \lambda I$. Assim T é singular, então, pelo Teorema 8 não é invertível.

$(ii) \implies (iii)$

Suponha que $T - \lambda I$ é não invertível. Assim, a matriz $\lambda I - A$ não é invertível, logo $\det(\lambda I - A) = 0$.

$(iii) \implies (i)$

Suponha $\det(\lambda I - A) = 0$. Então $\lambda I - A$ não é invertível. Assim, $T - \lambda I$ também não é invertível, então pelo Teorema 8, não é injetivo. Se não é injetivo, existe $u \neq 0$ tal que $(T - \lambda I)u = 0$. Logo, $Tu = \lambda u$, e $u \neq 0$, ou seja, λ é autovalor de T .

□

Definição 24. Seja V um espaço vetorial n -dimensional, $T : V \rightarrow V$ um operador linear e $A \in M_{n \times n}(F)$ uma matriz associada T em uma base qualquer fixada. O polinômio $p(x) = \det(xI - A)$ é chamado polinômio característico associado ao operador T .

Exemplo 18. Consideremos o operador linear $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definido por

$$T(x_1, x_2) = (-x_2, x_1).$$

Ao representar T na base canônica $\{e_1, e_2\}$ do \mathbb{R}^2 , obtemos a matriz

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

cujos polinômio característico é $p(\zeta) = \zeta^2 + 1$, que não possui raízes reais. Portanto, T não possui autovalores. Mas a matriz A pode ser considerada como uma matriz com entradas complexas; nesse caso, o polinômio característico $p(\zeta) = \zeta^2 + 1$ pode ser fatorado como $p(\zeta) = (\zeta - i)(\zeta + i)$ e T possui dois “autovalores complexos” distintos. Isso mostra que a análise do espectro de um operador $T : V \rightarrow V$ depende muito do corpo F sobre o qual V é espaço vetorial.

Observação 15. Sendo um V um espaço vetorial de dimensão finita sobre \mathbb{C} e $T : V \rightarrow V$ uma transformação linear. O Teorema Fundamental da Álgebra garante que $\sigma(T) \neq \emptyset$.

Lema 1. Matrizes semelhantes têm o mesmo polinômio característico.

Demonstração: Se $B = P^{-1}AP$, então

$$\begin{aligned} \det(xI - B) &= \det(xI - P^{-1}AP) \\ &= \det(P^{-1}(xI - A)P) \\ &= \det P^{-1} \cdot \det(xI - A) \cdot \det P \\ &= \det(xI - A). \end{aligned}$$

□

Observação 16. Como consequência direta do Lema anterior, o polinômio associado ao operador T independe da matriz associada a T para cada base fixada (veja a Observação 12).

Exemplo 19. Seja a matriz

$$A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 4 \\ 3 & -6 & -4 \end{bmatrix}.$$

Então o polinômio característico de A é:

$$\det(xI - A) = \begin{vmatrix} x-5 & 6 & 6 \\ 1 & x-4 & -2 \\ -3 & 6 & x+4 \end{vmatrix} = x^3 - 5x^2 + 8x - 4 = (x-1)(x-2)^2.$$

O teorema a seguir é um importante resultado da Álgebra linear e estabelece uma característica importante para o polinômio característico, fornecendo um exemplo de Cálculo Funcional.

Teorema 16 (Cayley-Hamilton). Seja T um operador linear sobre o espaço vetorial n -dimensional V e seja p o polinômio característico de T , então $p(T) = 0$.

Demonstração: Seja T um operador linear sobre o espaço n -dimensional V . Se g é um polinômio sobre o corpo F , então $g(T) : V \rightarrow V$ é uma função de T que associa a cada $u \in V$, um vetor, o qual indicaremos por $g(T)u$. Se h é um outro polinômio sobre F , então

$$(g + h)(T)u = g(T)u + h(T)u \quad \text{e} \quad [g(T)h(T)]u = (gh)(T)u.$$

Estas afirmações definem uma soma e produto dados por

$$g(T) + h(T) = (g + h)(T) \quad \text{e} \quad g(T)h(T) = (gh)(T).$$

Seja $\{u_1, \dots, u_n\}$ uma base de V e seja A a matriz de T nessa base, assim:

$$Tu_j = \sum_{i=1}^n A_{ij}u_i.$$

Definimos a matriz $B(x) = xI - A$, que é uma matriz $n \times n$ com entradas em $F[x]$ (anel dos polinômios com coeficientes em F). O polinômio característico é $p(x) = \det(B(x))$. Rearranjando a expressão acima, nós obtemos

$$Tu_j - \sum_{i=1}^n A_{ij}u_i = 0.$$

Usando que T e I são operadores lineares, organizamos a última expressão na forma:

$$\sum_{i=1}^n (\delta_{ij}T - A_{ij}I)u_i = 0, \quad \text{onde } \delta_{ij} = 1 \text{ se } i = j \text{ e } \delta_{ij} = 0, \text{ se } i \neq j.$$

Mas $\delta_{ij}T - A_{ij}I = B_{ij}(x)$, onde $B(x) = xI - A$. Portanto, a expressão $(\delta_{ij}T - A_{ij}I)$ dá o operador linear cuja matriz é $B = B(T) := (B_{ij})$ e satisfaz

$$\sum_{i=1}^n B_{ij}u_i = 0, \quad j = 1, \dots, n. \quad (1.2)$$

Desejamos mostrar que $p(T) = 0$. Assim, queremos mostrar que $p(T)u_k = 0$ para cada $k \in \{1, \dots, n\}$ (isso é suficiente porque $\{u_1, \dots, u_n\}$ é base). Seja $\bar{B} = \text{adj}(B)$ a matriz adjunta clássica de B (adjunta clássica de B está bem-definida). Lembremos que, $B\bar{B} = \bar{B}B = (\det B)I = p(T)I$. Multiplicando (1.2) por \bar{B}_{jk} e somando em j

$$\sum_{j=1}^n \bar{B}_{jk} \left(\sum_{i=1}^n B_{ij}u_i \right) = 0. \quad (1.3)$$

Agora, trocando a ordem dos somatórios, ficamos com

$$\sum_{i=1}^n \left(\sum_{j=1}^n B_{ij}\bar{B}_{jk} \right) u_i = 0.$$

Como

$$(B\bar{B})_{ik} = \sum_{j=1}^n B_{ij}\bar{B}_{jk}$$

e lembrando que $B\bar{B} = p(T)I$, logo

$$(B\bar{B})_{ik} = p(T)(I)_{ik} = \delta_{ik}p(T).$$

Portanto,

$$\sum_{j=1}^n B_{ij} \bar{B}_{jk} = \delta_{ik} p(T).$$

Substituindo em 1.3, obtemos

$$\sum_{i=1}^n \left(\sum_{j=1}^n B_{ij} \bar{B}_{jk} \right) u_i = \sum_{i=1}^n \delta_{ik} p(T) u_i. \quad (1.4)$$

Pela definição de δ_{ik} ,

$$\sum_{i=1}^n \delta_{ik} p(T) u_i = p(T) u_k. \quad (1.5)$$

Combinando (1.3), (1.4) e (1.5), obtemos $p(T) u_k = 0$. Desta forma,

$$p(T) = 0.$$

□

Definição 25. O polinômio minimal de uma aplicação linear $T : V \rightarrow V$ é o polinômio mônico de menor grau tal que $p(T) = 0$.

1.1.9 Operadores diagonalizáveis

Definição 26. Seja V um espaço vetorial de dimensão finita e $T : V \rightarrow V$ um operador linear. Dizemos que T é **diagonalizável** se existe uma base de V formada pelos autovetores de T .

Sendo T diagonalizável, então existe uma base ordenada $\mathcal{B} = \{v_1, \dots, v_n\}$ de V na qual cada $i \in \{1, \dots, n\}$, o vetor v_i é um autovetor de T associado ao autovalor α_i , assim $Tv_i = \alpha_i v_i$, então a matriz

$$[T]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix}$$

é uma matriz diagonal, o que justifica a denominação "diagonalizável" aqui usada.

Teorema 17. Seja V um espaço vetorial de dimensão finita e $T : V \rightarrow V$ um operador diagonalizável e $\alpha_1, \dots, \alpha_k$ são os autovalores distintos de T . Então existem operadores lineares $E_1, \dots, E_k : V \rightarrow V$ tais que

- (a) Para cada $v \in V$, $T(v) = \alpha_1 E_1(v) + \cdots + \alpha_k E_k(v)$;

- (b) $I = E_1 + \cdots + E_k$;
- (c) $E_i \circ E_j = 0$, $i \neq j$, com $i \in \{1, \dots, k\}$ e $j \in \{1, \dots, k\}$;
- (d) $E_i^2 = E_i$ para cada $i \in \{1, \dots, k\}$;
- (e) O conjunto imagem de E_i é o espaço dos autovetores de T associados ao autovalor α_i .

Demonstração: Ver a demonstração do Teorema 8 do capítulo 6 em [7]. □

Teorema 18. Sejam V e T como no Teorema 17. Suponhamos que existam k escalares $\alpha_1, \dots, \alpha_k$, distintos e k operadores lineares não-nulos E_1, \dots, E_k sobre V que satisfaçam as condições (a), (b) e (c) do Teorema 17. Então T é diagonalizável, $\alpha_1, \dots, \alpha_k$ são os autovalores distintos de T e os E_i também satisfazem as condições (d) e (e) do Teorema 17.

Demonstração: Ver a demonstração do Teorema 9 do capítulo 6 em [7]. □

Teorema 19. Sejam V um espaço de dimensão finita e $T : V \rightarrow V$ um operador linear. Então, uma condição necessária e suficiente para que T seja diagonalizável é que o polinômio minimal de T seja da forma

$$p(x) = (x - \alpha_1) \cdots (x - \alpha_k), \quad x \in F,$$

onde $\alpha_1, \dots, \alpha_k$ são escalares distintos em F .

Demonstração: Ver a demonstração o Teorema 5.5.9 em [2]. □

Corolário 7. Seja $A \in M_{n \times n}(F)$. Então A é semelhante sobre F a uma matriz diagonal se, e somente se, o polinômio minimal p de A é da forma

$$p(x) = (x - \alpha_1) \cdots (x - \alpha_k)$$

onde $\alpha_1, \dots, \alpha_k$ são elementos distintos de F .

Demonstração: Segue imediatamente do Teorema anterior. □

Teorema 20. Sejam T e S operadores lineares diagonalizáveis sobre um espaço de dimensão finita V . Se T e S comutam, então eles são simultaneamente diagonalizáveis, isto é, existe uma base de V na qual cada vetor é um autovetor de T e também um autovetor de S .

Demonstração: Ver demonstração do Teorema 11 do capítulo 6 em [7]. □

Corolário 8. Sejam $A, B \in M_n(F)$, sendo cada uma semelhante sobre F a uma matriz diagonal. Se A e B comutam, existe uma $n \times n$ matriz invertível P sobre F tal que tanto $P^{-1}AP$ como $P^{-1}BP$ são diagonais.

Demonstração: Ver demonstração em [7]. □

1.2 Análise

Nesta seção apresentaremos alguns resultados de Análise que serão importantes para a compreensão dos capítulos que se seguem. Os resultados aqui apresentados foram baseados nas referências [4, 9, 10, 12, 13, 14].

1.2.1 Séries

Definição 27. Uma sequência numérica $a_1, a_2, a_3, \dots, a_n, \dots$ é uma função $f : \mathbb{N} \rightarrow F$ (com $F = \mathbb{R}$ ou $F = \mathbb{C}$), $\mathbb{N} \ni n \mapsto f(n) = a_n$. O número natural n é chamado o índice e a imagem a_n é o n -ésimo elemento da sequência, ou termo geral.

Definição 28. Dada uma série $\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \dots$, isto é, uma soma infinita de números reais (ou complexos), denota-se por s_n sua n -ésima soma parcial definida por

$$s_n = \sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_n.$$

Se a sequência $(s_n)_n$ for convergente e $\lim_{n \rightarrow \infty} s_n = s$ existir como um número real, então a série $\sum a_n$ é chamada **convergente**, e escrevemos

$$a_1 + a_2 + \dots + a_n + \dots = s \quad \text{ou} \quad \sum_{n=1}^{\infty} a_n = s.$$

O número s é chamado a **soma** da série. Se a sequência $(s_n)_n$ é divergente, então a série é chamada **divergente**.

Teorema 21. A série $\sum a_n$ converge se, e só se, dado $\varepsilon > 0$, existir n_0 (que pode depender de ε) tal que $|\sum_{k=n+1}^m a_k| < \varepsilon$ para todos $m > n \geq n_0$.

Demonstração: Ver a demonstração do Teorema 1.7 do capítulo 1 em [5]. □

Corolário 9. Se $\sum a_n$ for convergente, então $\lim a_n = 0$.

Demonstração: Note que $a_n = s_n - s_{n-1}$, sendo a série convergente. Então $\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} s_{n-1} = s$. Logo $\lim_{n \rightarrow \infty} a_n = 0$. □

É interessante estabelecermos critérios que possam auxiliar em identificar se uma série é convergente. Neste sentido, tem-se o chamado teste da razão (veja [10] para mais detalhes).

Teste da razão

- (i) Se $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$, então a série $\sum_{n=1}^{\infty} a_n$ é absolutamente convergente (e, portanto, convergente).

- (ii) Se $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1$ ou $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty$, então a série $\sum_{n=1}^{\infty} a_n$ é divergente.
- (iii) Se $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$, o Teste da Razão é inconclusivo, ou seja, nenhuma conclusão pode ser tirada sobre a convergência ou divergência de $\sum a_n$.

Definição 29. Sejam $z, a \in \mathbb{C}$. Uma série de potências centrada em a é uma série da forma

$$\sum_{n=0}^{\infty} c_n (z - a)^n = c_0 + c_1(z - a) + c_2(z - a)^2 + \dots \quad (1.6)$$

Note que a série (1.6) é uma função na variável z desde que consigamos garantir a convergência. Neste sentido é interessante buscar por quais valores $z \in \mathbb{C}$ esta série é convergente.

Teorema 22. Para dada série de potências $\sum_{n=0}^{\infty} c_n (z - a)^n$, existem apenas três possibilidades:

- (i) A série converge apenas quando $z = a$.
- (ii) A série converge para todo z .
- (iii) Existe um número positivo R tal que a série converge se $|z - a| < R$ e diverge se $|z - a| > R$. (o número R é chamado de raio de convergência).

Teorema 23. Se f tiver uma representação (expansão) em série de potências em a , isto é, se existe $R > 0$ tal que

$$f(z) = \sum_{n=0}^{\infty} c_n (z - a)^n, \quad |z - a| < R,$$

então seus coeficientes são dados pela fórmula

$$c_n = \frac{f^{(n)}(a)}{n!}.$$

Demonstração: Ver a demonstração do Corolário 2.13 em [11]. □

Substituindo essa fórmula para c_n de volta na série, vemos que, se f tiver uma expansão em série de potências em a , então ela deve ser da seguinte forma:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n = f(a) + \frac{f'(a)}{1!} (x - a) + \frac{f''(a)}{2!} (x - a)^2 + \frac{f'''(a)}{3!} (x - a)^3 + \dots \quad (1.7)$$

A série na equação (1.7) é chamada **série de Taylor** da função centrada em a . Para o caso especial $a = 0$, a série de Taylor torna-se

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \dots \quad (1.8)$$

Esse caso surge com frequência e lhe foi dado o nome especial de **série de Maclaurin**.

Exemplo 20. Vamos encontrar a série de potências que representa $f(x) = e^x$ em torno de $x = 0$. Como a derivada de e^x é ela mesma, temos: $f^{(n)}(x) = e^x$ para todo n , logo $f^{(n)}(0) = 1$. A série de Maclaurin fica:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Exemplo 21. De maneira análoga, podemos encontrar a série de Maclaurin para $f(x) = \cos(x)$ e $f(x) = \sin(x)$:

$$\begin{aligned} \sin(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \text{para todo } x \end{aligned}$$

e

$$\begin{aligned} \cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad \text{para todo } x. \end{aligned}$$

Exemplo 22. Seja $x \in \mathbb{R}$ com $|x| < 1$, então

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} x^n}{n}.$$

Definição 30. Diz-se que uma sequência de funções $f_n : D \subset \mathbb{C} \rightarrow \mathbb{C}$ converge uniformemente para uma função $f : X \rightarrow \mathbb{R}$ quando, para todo $\epsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ tal que $n > n_0 \Rightarrow |f_n(x) - f(x)| < \epsilon$, seja qual for $x \in D$.

Teorema 24 (M-teste de Weierstrass). Para cada $n \in \mathbb{N}$, considere a sequência de funções $f_n : D \subset \mathbb{C} \rightarrow \mathbb{C}$, tal que

$$\sup_{x \in D} |f_n(x)| \leq M_n, \quad \forall n \in \mathbb{N}.$$

Se $\sum_{n=1}^{\infty} M_n$ converge, então $\sum_{n=1}^{\infty} f_n(z)$ converge uniformemente em D .

Demonstração: Ver a demonstração do Teorema 4.4 em [14]. □

1.2.2 Definições e resultados auxiliares

Definição 31. Dados os números $r > 0$ e $z_0 \in \mathbb{C}$, chama-se **disco aberto** de raio r e centro z_0 ao conjunto $D_r(z_0)$ de todos os números complexos que estão a uma distância menor que r do centro z_0 .

$$D_r(z_0) = \{z \in \mathbb{C}; |z - z_0| < r\}.$$

Definição 32. Dizemos que o conjunto $C \subset \mathbb{C}$ é um conjunto aberto se para cada $z \in C$, existir $r > 0$ tal que $D_r(z) \subset C$.

Definição 33.

- (i) Dizemos que um conjunto $F \subset \mathbb{C}$ é fechado quando seu complementar é aberto.
- (ii) Diz-se que um conjunto $C \subset \mathbb{C}$ é limitado se existe um número positivo K tal que $|z| \leq K$ para todo z em C . Chama-se conjunto compacto a todo conjunto limitado e fechado de \mathbb{C} .

Definição 34. Diz-se que uma função f é analítica numa região R se ela é derivável em cada ponto de R ; f é analítica num ponto z_0 se f é analítica num disco aberto contendo z_0 .

Definição 35 (Função suave). Uma função $f : U \rightarrow \mathbb{R}$ que possui, em cada ponto de U , todas as derivadas parciais de ordem k , as quais são funções contínuas em U , chama-se uma função de **classe** C^k . Escreve-se então $f \in C^k$. Quando $f \in C^k$ para todo $k = 1, 2, 3, \dots$, diz-se que f é uma função de **classe** C^∞ .

Proposição 6 (Desigualdade de Cauchy-Schwarz). Seja $u, v \in F^n$ (os vetores u, v podem ser identificados com matrizes $1 \times n$ com coeficientes em F), note que a desigualdade a seguir é verdadeira

$$|uv^t| \leq \|u\| \|v\|,$$

onde $\|v\| = \|(v_1, \dots, v_n)\| = \sqrt{|v_1|^2 + \dots + |v_n|^2}$ é a norma euclidiana.

Demonstração: Ver a demonstração em [1]. □

Definição 36. Seja $A \in M_{n \times n}(F)$ e $\|\cdot\|$ uma norma em F^n . Definimos a **norma** da matriz A por

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Para mais detalhes a respeito da definição de normas e espaços vetoriais normados sugerimos [1, 2, 7].

A próxima proposição nos garante que a norma de uma matriz está bem definida no espaço $M_{n \times n}(F)$.

Proposição 7. Seja $A: F^n \rightarrow F^n$ a aplicação linear dada pela matriz $A \in M_{n \times n}(F)$. Então

- (i) A aplicação $\|\cdot\|: M_{n \times n}(F) \rightarrow [0, \infty)$ é uma norma;
- (ii) $\|AB\| \leq \|A\|\|B\|$ para quaisquer $A, B \in M_n(F)$;
- (iii) $\|A\| = \max_{\|x\|=\|y\|=1} |(Ax)y^t|$.

Demonstração: Ver demonstração em [2]. □

A seguinte proposição mostra que $M_{n \times n}(\mathbb{C})$ é um espaço normado que tem a propriedade de completude. Nestes casos, dizemos que o espaço é de Banach.

Proposição 8. Considere $(A_k)_k \subset M_{n \times n}(\mathbb{C})$ tal que $\lim_{l, k \rightarrow \infty} \|A_k - A_l\| = 0$, então existe $B \in M_{n \times n}(\mathbb{C})$ tal que

$$\lim_{k \rightarrow \infty} \|A_k - B\| = 0.$$

Demonstração: Mostraremos o resultado considerando $M_{n \times n}(\mathbb{R})$, mas o argumento é o mesmo para $M_n(\mathbb{C})$. Seja $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, \dots , $e_n = (0, 0, \dots, 1)$. Note que, para cada $i \in \{1, \dots, n\}$, $\|A_k e_i - A_l e_i\| \leq \|A_k - A_l\|$ (pela proposição 7 (iii)), tomando o limite temos que $\lim_{k, l} \|A_k e_i - A_l e_i\| = 0$. Observe que, para cada k , $A_k e_i$ fornece a i -ésima coluna da matriz A_k , que podemos identificar como um vetor $v_i^k \in \mathbb{R}^n$. Desta forma, para cada i fixado, construímos uma sequência $(v_i^k)_k \subset \mathbb{R}^n$, tal que $\lim_{l, k} \|v_i^k - v_i^l\| = 0$. Desta forma $(v_i^k)_k$ satisfaz o critério de Cauchy (veja o Teorema 4 em [9]), assim existe $v_i \in \mathbb{R}^n$ tal que $\lim_{k \rightarrow \infty} \|v_i^k - v_i\| = 0$. Defina B a matriz a qual as colunas sejam os vetores $v_i \cdots v_n$ (encontrados anteriormente). Agora, seja $x \in \mathbb{R}^n$, com $\|x\| = 1$, pela Definição 5,

$$\lim_{k \rightarrow \infty} [(A_k - B)x^t]_{ij} = \lim_{k \rightarrow \infty} \sum_{r=1}^n (a_{ir}^k - v_{ir})x_{rj} = \sum_{r=1}^n \lim_{k \rightarrow \infty} (a_{ir}^k - v_{ir})x_{rj} = 0.$$

Isso mostra que, para k suficientemente grande, cada entrada de A_k fica arbitrariamente próxima da entrada correspondente de B . Assim, tomando o máximo sobre $\|x\| = 1$, obtemos

$$\lim_{k \rightarrow \infty} \|A_k - B\| = 0.$$

□

Capítulo 2

O Cálculo Funcional

Neste capítulo apresentaremos o Cálculo Funcional para operadores lineares definido entre espaços vetoriais de dimensão finita. As definições e resultados que nos auxiliarão na compreensão deste cálculo podem ser consultadas em [2, 3].

2.1 Polinômio interpolador

Definição 37. Seja p um polinômio e $f : U \subset \mathbb{C} \rightarrow \mathbb{C}$ uma função complexa. A função f é dita *euclidiana* com relação ao polinômio p se:

- (i) todas as raízes de p pertencem a U .
- (ii) se ζ_0 for uma raiz de p com multiplicidade k , então f tem derivadas até a ordem k em ζ_0 .

Observação 17. Note que, se U é um aberto e f é analítica então, em particular, é infinitamente diferenciável, assim a condição (ii) é verificada imediatamente.

Proposição 9. Seja $f : U \subset \mathbb{C} \rightarrow \mathbb{C}$ uma função complexa e p um polinômio. Se f é euclidiana com relação ao polinômio p , então existe uma função $q : U \subset \mathbb{C} \rightarrow \mathbb{C}$, contínua em cada uma das raízes do polinômio p , e um polinômio r tais que $f = qp + r$, com o $gr(r) < gr(p)$. Aqui $gr(p)$ denota o grau do polinômio p .

Demonstração: Seja r um polinômio arbitrário. Considere a função q definida para cada $z \in U$ por

$$q(z) = \frac{f(z) - r(z)}{p(z)}. \quad (2.1)$$

Queremos mostrar que podemos escolher r com grau menor do que o de p , de modo que q possua extensão contínua em cada uma das raízes de p , ou seja, para cada raiz ζ de p , a função $q_\zeta : U \rightarrow \mathbb{C}$,

$$q_\zeta(z) = \begin{cases} q(z), & z \neq \zeta \\ \lim_{z \rightarrow \zeta} q(z), & z = \zeta \end{cases}$$

é contínua em U . Desta forma precisamos garantir que o limite existe

$$\lim_{z \rightarrow \zeta} q(z) = \lim_{z \rightarrow \zeta} \frac{f(z) - r(z)}{p(z)}$$

e seja finito. Perceba que q é tão suave quanto a f fora das raízes de p (onde p não se anula), pois a f é euclidiana e r, p são polinômios com $p(z) \neq 0$, portanto, a divisão de funções suaves preserva a suavidade.

Seja ζ_0 uma raiz de multiplicidade k do polinômio p , isto é, $p(\zeta) = (\zeta - \zeta_0)^k s(\zeta)$, sendo s um polinômio tal que $s(\zeta_0) \neq 0$. Queremos encontrar r de modo que o quociente, para cada $z \neq \zeta_0$,

$$g(z) := \frac{f(z) - r(z)}{(z - \zeta_0)^k}$$

possua extensão contínua em ζ_0 , ou seja, precisamos que

$$\lim_{z \rightarrow \zeta_0} g(z) = \lim_{z \rightarrow \zeta_0} \frac{f(z) - r(z)}{(z - \zeta_0)^k} \quad (2.2)$$

exista e seja finito.

Uma vez que f têm derivadas até a ordem k em ζ_0 , pois f é euclidiana em relação a p , de acordo com a regra de L'Hospital, podemos perceber que, para que o $\lim_{z \rightarrow \zeta_0} g(z)$ seja finito, precisamos

$$f(\zeta_0) = r(\zeta_0), \quad f'(\zeta_0) = r'(\zeta_0), \quad \dots, \quad f^{(k-1)}(\zeta_0) = r^{(k-1)}(\zeta_0). \quad (2.3)$$

Garantindo a existência do limite (2.2) conseguiremos uma extensão contínua para q em cada uma das raízes de p e por consequência o quociente (2.1) estará bem definido. \square

A seguir apresentaremos a demonstração da existência do polinômio r mencionado no resultado anterior. Denotaremos $f^{(0)} = f$.

Lema 2. Sejam dados os valores

$$\begin{array}{cccc} f(\zeta_1) & f'(\zeta_1) & \dots & f^{(\gamma_1-1)}(\zeta_1) \\ \vdots & & \vdots & \\ f(\zeta_j) & f'(\zeta_j) & \dots & f^{(\gamma_j-1)}(\zeta_j) \end{array}$$

em que ζ_1, \dots, ζ_j são distintos. Seja $n = \gamma_1 + \gamma_2 + \dots + \gamma_j \in \mathbb{N}$. Então, existe um único polinômio r , de grau menor do que ou igual a $n - 1$, satisfazendo

$$r^{(k)}(\zeta_i) = f^{(k)}(\zeta_i)$$

para todo $i = 1, \dots, j$ e $k_i = 0, \dots, \gamma_i - 1$.

Exemplo 23. Antes de apresentarmos o caso geral, vejamos em um exemplo o Lema 2. Suponhamos conhecidos os valores $f(\zeta_0), f(\zeta_1)$ e $f'(\zeta_1)$, isto é, $\gamma_1 = 1$, $\gamma_2 = 2$ e $n = 3$. Queremos encontrar um polinômio r de grau 2 que satisfaça

$$\begin{cases} f(\zeta_0) = r(\zeta_0) \\ f(\zeta_1) = r(\zeta_1) \\ f'(\zeta_1) = r'(\zeta_1) \end{cases} .$$

Seja $r(\zeta) = \lambda\zeta^2 + \mu\zeta + \nu$, então da relação anterior

$$\begin{cases} f(\zeta_0) = \lambda\zeta_0^2 + \mu\zeta_0 + \nu \\ f(\zeta_1) = \lambda\zeta_1^2 + \mu\zeta_1 + \nu \\ f'(\zeta_1) = 2\lambda\zeta_1 + \mu \end{cases} .$$

Então os coeficientes de r devem satisfazer o sistema matricial:

$$\begin{pmatrix} \zeta_0^2 & \zeta_0 & 1 \\ \zeta_1^2 & \zeta_1 & 1 \\ 2\zeta_1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \begin{pmatrix} f(\zeta_0) \\ f(\zeta_1) \\ f'(\zeta_1) \end{pmatrix} . \quad (2.4)$$

Se os valores $f(\zeta_0), f(\zeta_1)$ e $f'(\zeta_1)$ forem nulos, temos a solução trivial $\lambda = \zeta = \nu = 0$, então basta tomarmos $r \equiv 0$.

Afirmamos que o sistema (2.4) possui uma única solução, que implica que o determinante da matriz é 0. De fato, suponha que não possua solução ou que essa solução não seja única. Então, o sistema homogêneo associado possui uma solução não trivial $(\lambda_0 \ \mu_0 \ \nu_0)^t$. Consideremos o polinômio não nulo

$$g(\zeta) = \lambda_0\zeta^2 + \mu_0\zeta + \nu_0.$$

Das equações do sistema homogêneo

$$\begin{cases} \lambda\zeta_0^2 + \mu\zeta_0 + \nu = 0 \implies g(\zeta_0) = 0 \\ \lambda\zeta_1^2 + \mu\zeta_1 + \nu = 0 \implies g(\zeta_1) = 0 \\ 2\lambda\zeta_1 + \mu = 0 \implies g'(\zeta_1) = 0 \end{cases}$$

Isso nos mostra que ζ_1 é uma raiz de g , ou seja, $(\zeta - \zeta_1)$ é um fator de $g(\zeta)$. Assim,

$$g(\zeta) = (\zeta - \zeta_1) \cdot Q(\zeta).$$

Derivando, obtemos

$$g'(\zeta) = Q(\zeta) + (\zeta - \zeta_1) \cdot Q'(\zeta).$$

Se $g'(\zeta_1) = 0$, então $g'(\zeta_1) = Q(\zeta_1) + 0 \cdot Q'(\zeta_1) = Q(\zeta_1) = 0$. Assim, temos que $Q(\zeta_1) = 0$, então $(\zeta - \zeta_1)$ é também fator de $Q(\zeta)$, dessa maneira podemos escrever

$$Q(\zeta) = (\zeta - \zeta_1) \cdot W(\zeta).$$

Portanto,

$$g(\zeta) = (\zeta - \zeta_1) \cdot [(\zeta - \zeta_1) \cdot W(\zeta)] = (\zeta - \zeta_1)^2 \cdot W(\zeta).$$

Além disso, $g(\zeta_0) = 0$, assim $(\zeta - \zeta_0)$ é fator de $g(\zeta)$. Isso implica que $(\zeta - \zeta_1)^2$ e $(\zeta - \zeta_0)$ divide $W(\zeta)$, como $\zeta_1 \neq \zeta_0$, o produto

$$(\zeta - \zeta_0)(\zeta - \zeta_1)^2$$

divide $g(\zeta)$ e o grau desse polinômio é 3. Mas $g(\zeta)$ é não nulo e tem grau menor ou igual a 2, um absurdo. Logo o sistema (2.4) tem solução única para quaisquer valores de $f(\zeta_0)$, $f(\zeta_1)$ e $f'(\zeta_1)$.

Demonstração do Lema 2: O polinômio r procurado satisfaz a um sistema linear que pode ser escrito matricialmente como

$$Bz = b,$$

sendo z o vetor que tem como coordenadas os coeficientes procurados de r , b um vetor cujas n coordenadas são os valores conhecidos de f e B a matriz $n \times n$ do sistema linear assim formado. Se B não é uma matriz invertível, o sistema $Bz = 0$ tem solução não trivial

$$z_0 = (\lambda_0 \dots \lambda_{n-1})^t.$$

Consideremos, então, o polinômio na incógnita ζ

$$g(\zeta) = \lambda_0 + \lambda_1\zeta + \dots + \lambda_{n-1}\zeta^{n-1},$$

que é um polinômio de grau menor do que ou igual a $n - 1$. Como z_0 satisfaz o sistema homogêneo associado, por argumentos análogos ao que apresentamos no Exemplo anterior, temos que $g(\zeta)$ deve ser um múltiplo de

$$(\zeta - \zeta_1)^{\gamma_1} \dots (\zeta - \zeta_j)^{\gamma_j} \tag{2.5}$$

o que é um absurdo, pois o grau do polinômio 2.5 é n . Assim, B possui inversa e o sistema $Bz = b$ tem solução única, $z = B^{-1}b$, para qualquer que seja o vetor b . \square

Observação 18. O polinômio r definido no Lema 2 é conhecido na literatura como polinômio interpolador de Hermite.

Proposição 10. Na divisão euclidiana, $f = qp + r$, ($gr(r) < gr(p)$) da função analítica $f : U \subset \mathbb{C} \rightarrow \mathbb{C}$ pelo polinômio p cujas raízes estão em U , o quociente q é analítico em U .

Demonstração: De acordo com a demonstração da Proposição 9, a função $q : U \rightarrow \mathbb{C}$, para cada $z \in U$, dada por

$$q(z) = \frac{f(z) - r(z)}{p(z)}$$

é analítica, pois o numerador e o denominador se anulam exatamente nos mesmos pontos e os zeros do numerador possuem multiplicidade maior do que ou igual à dos zeros do denominador. A conta a seguir verifica que essas singularidades são removíveis:

$$\lim_{z \rightarrow z_0} q(z) = \lim_{z \rightarrow z_0} \frac{f(z) - r(z)}{p(z)} = \lim_{z \rightarrow z_0} \frac{(z - z_0)^m \alpha(z)}{(z - z_0)^n \beta(z)} = \lim_{z \rightarrow z_0} (z - z_0)^{m-n} \frac{\alpha(z)}{\beta(z)} = 0,$$

com $m > n$, $\alpha(z_0) \neq 0$ e $\beta(z_0) \neq 0$. Assim, q possui uma expansão em série de potências em cada ponto de U . □

2.2 Funções de matrizes

Seja V um espaço vetorial (real ou complexo) de dimensão finita, $T : V \rightarrow V$ um operador linear e m polinômio mínimo de T . Se f é uma função complexa euclidiana em relação a m , ou seja, $f(z) = q(z)m(z) + r(z)$. Pelo Teorema de Cayley-Hamilton, $m(T) = 0$ (operador nulo), assim é natural pensarmos em definir $f(T) = q(T)m(T) + r(T) = r(T)$. O que motiva a definição a seguir.

Definição 38. Seja $m(\zeta) = (\zeta - \zeta_1)^{\gamma_1} \cdots (\zeta - \zeta_j)^{\gamma_j}$ o polinômio mínimo do operador T . Se estiverem definidos os valores

$$\begin{array}{cccc} f(\zeta_1) & f'(\zeta_1) & \cdots & f^{(\gamma_1-1)}(\zeta_1) \\ \vdots & \vdots & & \vdots \\ f(\zeta_j) & f'(\zeta_j) & \cdots & f^{(\gamma_j-1)}(\zeta_j), \end{array}$$

dizemos que f é **euclidiana com respeito a T** e definimos

$$f(T) = r(T),$$

sendo r o polinômio interpolador dado pelo Lema 2.

Exemplo 24. Consideremos a matriz

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

cujo polinômio característico é

$$\det(\zeta I - A) = \det \begin{pmatrix} \zeta - 2 & -1 & -1 \\ -1 & \zeta - 2 & -1 \\ -1 & -1 & \zeta - 2 \end{pmatrix} = (\zeta - 4)(\zeta - 1)^2$$

e cujo polinômio mínimo é $m(\zeta) = (\zeta - 4)(\zeta - 1)$. Caso queiramos calcular A^{1000} , definimos a função $f(\zeta) = \zeta^{1000}$ e consideramos o polinômio $r(\zeta) = \lambda\zeta + \mu$ (pois pelo algoritmo da divisão Euclidiana de polinômios: $f(\zeta) = q(\zeta)m(\zeta) + r(\zeta)$, como $gr(r) < gr(m) = 2$, o polinômio $r(\zeta)$ tem grau 1). Vamos encontrar agora os valores de λ e μ .

Para $\zeta = 1$

$$\left\{ \begin{array}{l} f(1) = r(1) \\ \end{array} \right. \implies \left\{ \begin{array}{l} 1^{1000} = \lambda + \mu \end{array} \right. \quad (2.6)$$

Para $\zeta = 4$

$$\left\{ \begin{array}{l} f(4) = r(4) \\ \end{array} \right. \implies \left\{ \begin{array}{l} 4^{1000} = 4\lambda + \mu \end{array} \right. \quad (2.7)$$

De (2.6): $\mu = 1 - \lambda$. Substituindo (2.6) em (2.7) e fazendo os devidos algebrismos, obtemos $\lambda = \frac{4^{1000}-1}{3}$ e $\mu = \frac{4-4^{1000}}{3}$. Assim, $r(\zeta) = \frac{4^{1000}-1}{3}\zeta + \frac{4-4^{1000}}{3}$.

Como $A^{1000} = r(A) = \frac{4^{1000}-1}{3}A + \frac{4-4^{1000}}{3}I$. Faça $\gamma = 4^{1000}$, assim

$$A^{1000} = r(A) = \begin{bmatrix} \frac{\gamma+2}{3} & \frac{\gamma-1}{3} & \frac{\gamma-1}{3} \\ \frac{\gamma-1}{3} & \frac{\gamma+2}{3} & \frac{\gamma-1}{3} \\ \frac{\gamma-1}{3} & \frac{\gamma-1}{3} & \frac{\gamma+2}{3} \end{bmatrix}.$$

Exemplo 25. Considere a matriz

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

Queremos calcular o $\cos A$. Temos que o polinômio característico de A é

$$\det(\zeta I - A) = \det \begin{pmatrix} \zeta & 0 & -1 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta \end{pmatrix} = -\zeta^3$$

e o polinômio mínimo $p(\zeta) = \zeta^2$ (faça A^2 e veja que é o menor polinômio que anula A). Vejamos que $p(\zeta)$ tem raiz $\zeta = 0$ com multiplicidade 3. Dessa maneira, queremos encontrar $r(\zeta)$ com grau menor que 3 tal que $f(0) = r(0)$; $f'(0) = r'(0)$ e $f''(0) = r''(0)$, sendo $f(\zeta) = \cos \zeta$. Utilizando p , obtemos o polinômio interpolador $r_p(\zeta) = \lambda \zeta^2 + \mu \zeta + \nu$. Agora, vamos encontrar os valores de λ, μ, ν .

$$\begin{cases} f(0) = r(0) \\ f'(0) = r'(0) \\ f''(0) = r''(0) \end{cases} \implies \begin{cases} 1 = \nu \\ 0 = 2\lambda\zeta + \mu = \mu \\ -1 = 2\lambda \end{cases}$$

daí, tiramos que $\lambda = -\frac{1}{2}, \mu = 0, \nu = 1$. Logo $r_p(\zeta) = -\frac{1}{2}\zeta^2 + 1$. Aplicando a matriz A ao polinômio interpolador, obtemos: $r_p(A) = -\frac{1}{2}A^2 + I$, mas $A^2 = 0$, assim, $r_p(A) = I$.

Agora, perceba que de uma maneira mais simples, podemos encontrar o mesmo resultado. Para essa conta, vamos utilizar o polinômio minimal. Como $r_m(\zeta)$ possui grau 2, o polinômio interpolador terá grau 1 e $\zeta = 0$ possui multiplicidade 2, precisamos igualar a função à sua primeira derivada. Segue assim que $r_m(\zeta) = \lambda\zeta + \mu$, dessa forma:

$$\begin{cases} f(0) = r(0) \\ f'(0) = r'(0) \end{cases} \implies \begin{cases} 1 = \mu \\ 0 = \lambda \end{cases}$$

Obtemos o polinômio interpolador $r_m(\zeta) = 1 \implies r_m(A) = I$ e $\cos(A) = r_p(A) = r_m(A) = I$.

Observação 19. Seja V um espaço de dimensão finita e $T : V \rightarrow V$ uma transformação linear diagonalizável. Sendo A uma representação para T , pelo Corolário 8 existe $P \in M_n(F)$ invertível tal que $A = P^{-1} \text{diag}(\lambda_1, \dots, \lambda_n) P$. Seja $r(z) = a_{n-1}z^{n-1} + \dots + a_1z + a_0$ o polinômio interpolador, assim

$$\begin{aligned} f(A) &= a_{n-1}A^{n-1} + \dots + a_1A + a_0I \\ &= a_{n-1}(P^{-1} \text{diag}(\lambda_1, \dots, \lambda_k) P)^{n-1} + \dots + a_1P^{-1} \text{diag}(\lambda_1, \dots, \lambda_k) P + a_0I \\ &= P^{-1} \text{diag}(a_{n-1}\lambda_1^n, \dots, a_{n-1}\lambda_n^{n-1}) P + \dots + P^{-1} \text{diag}(a_1\lambda_1, \dots, a_1\lambda_n) P + a_0P^{-1}P \\ &= P^{-1} \text{diag}(a_{n-1}\lambda_1^n + \dots + a_1\lambda_1 + a_0, \dots, a_{n-1}\lambda_n^{n-1} + \dots + a_1\lambda_n + a_0) P \\ &= P^{-1} \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) P = P^{-1} f(A) P, \end{aligned}$$

onde usamos que r e f coincidem nos autovalores de T .

O resultado a seguir estabelece a relação entre os autovalores de T e os autovalores de $f(T)$.

Teorema 25 (Mapeamento Espectral). Seja V um espaço vetorial n -dimensional sobre \mathbb{C} e f uma função euclidiana com relação ao operador linear $T : V \rightarrow V$. Dessa forma,

1. Se $\lambda \in \sigma(T)$, então $f(\lambda) \in \sigma(f(T))$ (ou seja, $f(\sigma(T)) = \sigma(f(T))$).
2. Se $w \in \sigma(f(T))$, então $w = f(\lambda)$, com $\lambda \in \sigma(T)$.

Mais precisamente, $f(\sigma(T)) = \sigma(f(T))$.

Demonstração: 1. Seja v autovetor de T associado ao autovalor λ , isto é, $Tv = \lambda v$, $v \neq 0$. Como f é euclidiana com relação à T , $f(T) = r(T) = a_k T^k + \dots + a_1 T + a_0 I$ (com $k \in \mathbb{N}$, $k < n$).

Agora, note que, $T(v) = \lambda v$, assim $T^2(v) = T(T(v)) = T(\lambda v) = \lambda(\lambda v) = \lambda^2 v$. Além disso, de modo geral, $T^m(v) = \lambda^m v$, para qualquer $m \in \mathbb{N}$.

Aplicando $r(T)$ em v e combinando a relação anterior, obtém-se

$$\begin{aligned} r(T)v &= a_k T^k v + a_{k-1} T^{k-1} v + \dots + a_1 T v + a_0 I v \\ &= a_k \lambda^k v + a_{k-1} \lambda^{k-1} v + \dots + a_1 \lambda v + a_0 v \\ &= (a_k \lambda^k + a_{k-1} \lambda^{k-1} + \dots + a_1 \lambda + a_0) v \end{aligned}$$

Ora, $a_k \lambda^k + \dots + a_0 = r(\lambda)$, e $r(\lambda) = f(\lambda)$, pois, f coincide com r nos autovalores de T . Assim,

$$f(T)v = f(\lambda)v.$$

Dessa forma, para cada $\lambda \in \sigma(T)$, temos $f(\lambda) \in \sigma(f(T))$. Ou seja, $f(\sigma(T)) \subset \sigma(f(T))$.

2. Agora, seja μ um autovalor de $f(T)$. Ou seja, existe um vetor não nulo w tal que $f(T)w = \mu w \iff f(T)w - \mu w = 0 \iff (f(T) - \mu I)w = 0$, temos $(r(T) - \mu I)w = 0$ (lembre-se que $r(T) = f(T)$). Considere o polinômio $g(z) = r(z) - \mu$. Pelo Teorema Fundamental da Álgebra, em \mathbb{C} , ele pode ser fatorado como:

$$g(z) = r(z) - \mu = a_k (z - \lambda_1) \cdots (z - \lambda_k),$$

onde os λ_i são as raízes complexas de $r(\zeta) - \mu$. Consequentemente,

$$g(T) = r(T) - \mu I = a_k (T - \lambda_1 I) \circ \dots \circ (T - \lambda_k I). \quad (2.8)$$

Se $g(T)$ não é invertível, então o produto de operadores em (2.8) não é invertível. Portanto, pelo menos um dos fatores $T - \lambda_i I$ não é invertível. Isso significa que λ_i é autovalor de

T , ou seja, $\lambda_i \in \sigma(T)$. Além disso, λ_i é raiz de $g(z)$, logo

$$g(\lambda_i) = r(\lambda_i) - \mu = 0 \quad \Rightarrow \quad r(\lambda_i) = \mu.$$

Mas $r(\lambda_i) = f(\lambda_i)$, já que λ_i é autovalor de T (lembre-se que r e f coincidem nos autovalores de T). Portanto,

$$f(\lambda_i) = \mu, \quad \lambda_i \in \sigma(T).$$

Assim, $\mu \in f(\sigma(T))$. □

2.3 Aplicações do cálculo funcional

O cálculo funcional nos permite aplicar funções matemáticas conhecidas diretamente a matrizes e operadores. Essa técnica nos ajuda a definir o fluxo de sistemas de equações diferenciais ordinárias, o qual é a solução desses sistemas por meio da exponencial de matriz. Nesta seção apresentaremos interessantes aplicações de funções de matrizes e algumas de suas propriedades.

2.3.1 O fluxo

Seja $A \in M_{n \times n}(\mathbb{R})$ e o sistema de equações diferenciais

$$\begin{cases} x'(t) = Ax(t), & t \in \mathbb{R} \\ x(0) = x_0 \in \mathbb{R}^n, \end{cases} \quad (2.9)$$

Exemplo 26. Seja $a \in \mathbb{R}$ e considere o problema de valor inicial (PVI)

$$\begin{cases} x'(t) = ax(t), & t \in \mathbb{R} \\ x(0) = x_0 \in \mathbb{R}, \end{cases}$$

sabemos que a solução (única) deste PVI é dada por $x(t) = x_0 e^{at}$. Isso nos motiva a definir o que seria as soluções de um sistema de equações diferenciais lineares de primeira ordem, mais precisamente o sistema (2.9).

O fluxo de um sistema linear de equações diferenciais que tem a forma $\frac{dx}{dt} = Ax$, em que A é uma matriz $n \times n$ e $x(t)$ a solução geral, pode ser escrita como $x(t) = e^{At}x(0)$. Começamos com a definição usual do fluxo e^{At} . Essa depende da noção de convergência uniforme e da noção de norma de uma matriz quadrada. Essa primeira definição pode ser omitida.

Considere a função exponencial $\exp : \mathbb{C} \rightarrow \mathbb{C}$, cuja representação em série de potências

$$\exp(zw) = e^{zw} = 1 + \sum_{n=1}^{\infty} \frac{z^n w^n}{n!}.$$

Agora, seja $\|A\|$ a norma usual em $M_{n \times n}(\mathbb{C})$

Afirmamos que a série

$$I + \sum_{n=1}^{\infty} \frac{A^n w^n}{n!}$$

é um elemento de $M_{n \times n}(\mathbb{C})$ para todo $A \in M_{n \times n}(\mathbb{C})$ e $w \in \mathbb{C}$. De fato, defina a sequência de matrizes

$$S_k(w) = I + \sum_{n=1}^k \frac{A^n w^n}{n!} \in M_{n \times n}(\mathbb{C}).$$

A norma em $M_{n \times n}(\mathbb{C})$ goza da propriedade

$$\|AB\| \leq \|A\| \|B\|,$$

donde segue que $\|A^i\| \leq \|A\|^i$, para todo $i \in \mathbb{N}$, por um argumento indutivo. Assim, para cada $k, l \in \mathbb{N}$, sem perda de generalidade, assumamos $l > k$, donde decorre que

$$\|S_k(w) - S_l(w)\| = \left\| \sum_{n=k+1}^l \frac{A^n w^n}{n!} \right\| \leq \sum_{n=k+1}^l \frac{\|A\|^n |w|^n}{n!}, \quad (2.10)$$

onde usamos a desigualdade triangular de normas (veja a Definição 6.2.1 em [3]) e $\|A^i\| \leq \|A\|^i$ para obter a última desigualdade.

Para cada valor de $w \in \mathbb{C}$, note que

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\|A\|^n}{n!} &= \lim_{n \rightarrow \infty} \left| \frac{\|A\|^{n+1}}{(n+1)!} \cdot \frac{\|A\|^n}{n!} \right| \\ &= \lim_{n \rightarrow \infty} \frac{\|A\|^{n+1}}{(n+1)!} \cdot \frac{n!}{\|A\|^n} \\ &= \lim_{n \rightarrow \infty} \frac{\|A\|^n \cdot \|A\|}{(n+1)n!} \cdot \frac{n!}{\|A\|^n} \\ &= \lim_{n \rightarrow \infty} \frac{\|A\|}{(n+1)} = \|A\| \cdot \lim_{n \rightarrow \infty} \frac{1}{(n+1)} = 0. \end{aligned}$$

Assim, para cada $w \in \mathbb{C}$, o Teste da razão garante que a série $\sum_n \frac{\|A\|^n |w|^n}{n!}$ converge.

Donde, pelo Teorema 9, segue

$$\lim_{k, l \rightarrow \infty} \sum_{n=k+1}^l \frac{\|A\|^n |w|^n}{n!} = 0, \quad (2.11)$$

o que implica por (2.10) que

$$\lim_{k,l \rightarrow \infty} \|S_k(w) - S_l(w)\| = 0. \quad (2.12)$$

Agora, note que para cada w um conjunto compacto $D \subset \mathbb{C}$,

$$\sup_{z \in D} \frac{\|A\|^n |w|^n}{n!} < \infty,$$

assim, pelo Teste- M de Weierstrass, segue que o limite (2.11) é uniforme em $w \in D$.

E por consequência o limite (2.12) também, para $w \in D$ ($D \subset \mathbb{C}$ compacto). Então pela Proposição 8 segue que, para cada $w \in \mathbb{C}$, existe $B_w \in M_{n \times n}(\mathbb{C})$ de modo que $\|S_k(w) - B_w\| = 0$, ou seja,

$$M_{n \times n}(\mathbb{C}) \ni B_w = I + \sum_{n=1}^{\infty} \frac{A^n w^n}{n!} := \exp(Aw) = e^{Aw}. \quad (2.13)$$

Tomando $w = \tau \in \mathbb{R}$, definimos o fluxo $e^{A\tau}$ da matriz A . Como já observamos, a convergência da série é uniforme, se τ pertencer a um conjunto compacto. Logo, a diferenciação termo a termo produz sua derivada e

$$\begin{aligned} \frac{d}{d\tau} e^{A\tau} &= \frac{d}{d\tau} \left[I + A\tau + \frac{A^2\tau^2}{2!} + \frac{A^3\tau^3}{3!} + \dots \right] \\ &= 0 + A + A^2\tau + \frac{A^3\tau^2}{2!} + \dots \end{aligned}$$

Colocando A em evidência à direita (porque matrizes não comutam em geral, mas aqui A comuta com A^n):

$$= \left[I + A\tau + \frac{A^2\tau^2}{2!} + \dots \right] A = e^{A\tau} A.$$

Além disso, quando $\tau = 0$, temos

$$e^{A\tau} \Big|_{\tau=0} = e^0 = I.$$

Essas são as propriedades principais do fluxo $e^{A\tau}$. Em particular, vemos que $e^{A\tau}$ é uma solução fundamental do sistema matricial $X' = AX$, $X(0) = I$.

Exemplo 27. Seja

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -5 \\ 0 & 1 & -2 \end{pmatrix} \in M_{3 \times 3}(\mathbb{C}).$$

Queremos calcular $e^{A\tau}$. O polinômio característico de A (e também o seu polinômio

mínimo) é

$$p(\zeta) = (\zeta - 1)(\zeta + i)(\zeta - i).$$

Para obtermos $e^{A\tau}$, definimos a função $f(\zeta) = e^{\zeta\tau}$. Basta, então, encontrar um polinômio, de grau no máximo igual a 2. Da fórmula de Euler, sabemos que $e^{i\theta} = \cos(\theta) + i\text{sen}(\theta)$, então temos o seguinte:

$$\begin{cases} f(1) = r(1) \\ f(i) = r(i) \\ f(-i) = r(-i) \end{cases} \implies \begin{cases} e^\tau = \lambda + \mu + \nu \\ \cos \tau + i \sin \tau = -\lambda + \mu i + \nu \\ \cos \tau - i \sin \tau = -\lambda - \mu i + \nu \end{cases}.$$

Resolvendo o sistema, achamos $\lambda = (e^\tau/2) - (\cos \tau + \sin \tau)/2$, $\mu = \sin \tau$ e $\nu = (e^\tau/2) + (\cos \tau - \sin \tau)/2$. Assim,

$$e^{A\tau} = \left[\frac{e^\tau}{2} - \frac{\cos \tau + \sin \tau}{2} \right] A^2 + (\sin \tau)A + \left[\frac{e^\tau}{2} + \frac{\cos \tau - \sin \tau}{2} \right] I,$$

que é, para cada $\tau \in \mathbb{R}$, uma matriz real, embora tenhamos considerado a matriz A como uma matriz complexa.

Exemplo 28. Seja

$$A = \begin{pmatrix} 3 & -4 & -1 \\ -3 & 5 & 1 \\ 21 & -32 & -7 \end{pmatrix}.$$

O polinômio característico de A é

$$p(\zeta) = (\zeta - 1)\zeta^2.$$

Para calcularmos $e^{A\tau}$, obtemos os coeficientes de $r(\zeta) = \lambda\zeta^2 + \mu\zeta + \nu$ de modo que sejam satisfeitas as relações

$$\begin{cases} f(1) = r(1) \\ f(0) = r(0) \\ f'(0) = r'(0) \end{cases} \implies \begin{cases} e^\tau = \lambda + \mu + \nu \\ 1 = \nu \\ \tau = \mu \end{cases}.$$

Assim, $\nu = 1$, $\mu = \tau$ e $\lambda = e^\tau - \tau - 1$. Concluimos que

$$e^{A\tau} = (e^\tau - \tau - 1)A^2 + \tau A + I.$$

Para obtermos algumas propriedades do fluxo, definimos:

Definição 39. Seja $I \subset \mathbb{R}$ um intervalo fechado não degenerado (isto é, I não se reduz a

um ponto). Uma aplicação contínua $x : I \rightarrow \mathbb{R}^n$ é chamada de **caminho**. O caminho x é diferenciável, se existir o **vetor velocidade**

$$x'(t) = \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} \in \mathbb{R}^n.$$

Se t for um ponto de fronteira, o limite é o respectivo limite lateral. Também chamamos o vetor velocidade de **derivada** de $x(t)$. Em outras palavras, se $x(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$, então $x'(t) = (x'_1(t), \dots, x'_n(t))$.

Observação 20. Como podemos identificar o espaço $M_{m \times n}(F)$ com F^{mn} , esta mesma noção faz sentido para caminhos que tomam valores no espaço $M_{m \times n}(F)$. Assim, considere o caminho $A : I \rightarrow M_{m \times n}(F)$, $A(t)$ denota um caminho em $M_{m \times n}(F)$ e, sua derivada é obtida ao se derivar cada uma das mn entradas de $A(t)$.

Proposição 11. Seja $A \in M_n(F)$, então são válidas as seguintes propriedades:

(i) $e^{At} \Big|_{t=0} = I;$

(ii) $\frac{d}{dt} e^{At} = e^{At} A;$

(iii) Como consequência direta de (i) e (ii), tem-se que o

$$\lim_{t \rightarrow 0} \frac{e^{At} - I}{t} = A.$$

Observação 21. Embora a função $f(\zeta) = e^\zeta$ satisfaça a equação

$$e^{\zeta+\omega} = e^\zeta e^\omega,$$

não podemos deduzir que $e^{A+B} = e^A e^B$, uma vez que a substituição simultânea das variáveis ζ por A e ω por B não é permitida pelo cálculo funcional. Contudo, se A e B comutarem, o simples conhecimento de que $e^{A\tau}$ é um polinômio em A nos permite concluir que $e^{A\tau} B = B e^{A\tau}$, que é uma parte importante da demonstração de que $e^{(A+B)\tau} = e^{A\tau} e^{B\tau}$ se, e somente se, $AB = BA$.

Exemplo 29. Sejam as matrizes

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note $A^2 = 0 \in M_{2 \times 2}(\mathbb{C})$, então, $A^n = 0$ para cada $n \in \mathbb{N}$ com $n \geq 2$. Assim, pela relação (2.13)

$$e^A = I + A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Agora, sendo $B = \text{diag}(0, 1)$, então pela Observação 19 segue que $e^B = \text{diag}(e^0, e^1) = \text{diag}(1, e)$, ou seja,

$$e^B = \begin{pmatrix} 1 & 0 \\ 0 & e \end{pmatrix}.$$

Assim,

$$e^A e^B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e \end{pmatrix} = \begin{pmatrix} 1 & e \\ 0 & e \end{pmatrix}.$$

Agora, vamos calcular e^{A+B} . Primeiro, temos que

$$C = A + B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Note que $C^2 = C$. Pela definição de exponencial de matriz:

$$e^{A+B} = e^C = I + \sum_{n=1}^{\infty} \frac{C^n}{n!} = I + \left(\sum_{n=1}^{\infty} \frac{1}{n!} \right) C = I + \left(\sum_{n=0}^{\infty} \frac{1}{n!} - \frac{1}{0!} \right) C = I + (e - 1)C.$$

Isso implica que

$$e^{A+B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (e - 1) \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & e - 1 \\ 0 & e \end{pmatrix}.$$

Dessa maneira, podemos perceber que $e^{A+B} \neq e^{AB}$.

2.3.2 Funções trigonométricas

O estudo anterior permanece válido para o caso da exponencial e^{iAt} (ou seja, para a função $g(\zeta, i\tau)$, com $\tau \in \mathbb{R}$, a qual gera as funções trigonométricas $\sin At$ e $\cos At$. Essas funções também são fáceis de obter por meio do cálculo funcional. As mesmas observações também se aplicam a outras funções trigonométricas.

2.3.3 Logaritmo

Seja $A \in M_{n \times n}(\mathbb{C})$ tal que $\det A \neq 0$ (ou seja, $0 \notin \sigma(A)$). Considere $f(z) = \log(z)$, escolha um ramo de f que contenha $\sigma(A)$, então pelo cálculo funcional que apresentamos está definida bem definida a matriz $B = \log A$ por meio do polinômio interpolador.

Observação 22. A matriz $B = \log A$ depende do ramo escolhido, mas a relação $e^B = A$ segue-se sempre de $e^{\log z} = z$. Se todos os autovalores da matriz real A forem positivos, podemos então considerar a função real $f(z) = \ln z$ (logaritmo neperiano) e aplicar a

mesma técnica. No caso em que A é simétrica, a matriz $B = \ln A$, assim obtida, é a única solução real da equação $e^B = A$.

Exemplo 30. Seja

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

O polinômio característico (e também minimal) de A é

$$p(\zeta) = (\zeta - 1)(\zeta - 2).$$

Como $\sigma(A) = \{1, 2\}$ está contido em $U = \mathbb{C} \setminus (\infty, 0]$, logo faz sentido calcularmos o $\log A$. E para isso, definimos $f(\zeta) = \ln \zeta$ e obtemos os coeficientes de $r(\zeta) = \lambda\zeta + \mu$ de modo que sejam satisfeitas as relações

$$\begin{cases} f(1) = r(1) \\ f(2) = r(2) \end{cases} \implies \begin{cases} 0 = \lambda + \mu \\ \ln 2 = 2\lambda + \mu \end{cases}.$$

Assim, $\lambda = \ln 2$ e $\mu = -\ln 2$. Dessa maneira, obtemos que $r(\zeta) = (\ln 2)\zeta - \ln 2$. Portanto, $\log A = r(A) = \ln(2)A - (\ln 2)I = \begin{pmatrix} \ln 2 & 0 \\ \ln 2 & 2\ln 2 \end{pmatrix} - \begin{pmatrix} \ln 2 & 0 \\ 0 & \ln 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \ln 2 & \ln 2 \end{pmatrix}$.

Teorema 26. Seja $A \in M_{n \times n}(\mathbb{C})$ com autovalores com parte real positiva e $\log : U = \mathbb{C} \setminus (-\infty, 0] \rightarrow \mathbb{C}$ (ramo principal). Então,

(i)

$$\log A = \int_0^\infty (A - I)(A + (t - 1)I)^{-1} \frac{dt}{t};$$

(ii) $\sigma(\log A) \subset \{z \in \mathbb{C} \mid |\operatorname{im}(z)| \leq \pi\}$.

Demonstração: (i) Note que $g(z) = \log(1 + z) = \int_1^\infty \frac{z}{z + s} \frac{ds}{s}$. Note que $f(z) = \int_1^\infty \frac{z}{z + s} \frac{ds}{s}$ é uma função euclidiana (já é uma representação para o log) e r o polinômio interpolador. Assim, $r(A) = f(A) = \int_1^\infty A(A + s)^{-1} \frac{ds}{s}$ (interpretamos a integral como a integral imprópria de cada uma das funções em s que compõe as entradas de $s^{-1}A(A + s)^{-1}$). Agora, observe que o polinômio interpolador em relação a g também deve ser r (a unicidade de r é dado pelo Lema 2) já que $f = g$ em U .

(ii) Pelo Teorema do mapeamento espectral temos que $\log(\sigma(A)) = \sigma(\log(A))$. Seja $z \in \sigma(\log(A))$, então $z = \log w$ com $w \in \sigma(A) \subset U$. Pela definição de ramo principal do logaritmo $z = \ln |w| + i \arg(w)$ e $-\pi < \arg(w) < \pi$. Portanto, $|\operatorname{im}(z)| = |\arg(w)| < \pi$.

□

2.3.4 Raiz quadrada

Suponhamos que todos os autovalores da matriz real A sejam reais e não negativos. Adicionalmente, se 0 for um autovalor de A , supomos que ele seja uma raiz simples do polinômio mínimo m de A . Nesse caso, podemos utilizar a função $f : (0, \infty) \rightarrow \mathbb{R}$, $f(x) = \sqrt{x}$ para definir \sqrt{A} . Aqui, o cálculo funcional é utilizado em uma função que é apenas contínua no autovalor simples $\lambda = 0$ da matriz A .

Contudo, podemos definir \sqrt{A} mesmo que A e seus autovalores sejam complexos e não nulos. Apenas precisamos escolher um ramo da função logaritmo $f(\zeta) = \log \zeta$ para o qual a raiz quadrada de todos os autovalores da matriz A esteja definida. Então aplicamos o cálculo funcional à função complexa $f(\zeta) = \sqrt{\zeta}$.

Observação 23. A definição de \sqrt{A} não determina todas as soluções da equação $B^2 = A$. Se A for a matriz identidade 2×2 ,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

também são soluções de $B^2 = I$, além de $B = I$, a única solução que pode ser obtida por meio da função raiz quadrada real. (A matriz $-I$ pode ser obtida utilizando a raiz quadrada complexa). Além disso, se $A = I$, a equação $B^2 = A$ possui a família de soluções reais

$$\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad \theta \in [0, 2\pi],$$

que não vem da função \sqrt{A} . Note que, se $\sigma(A) \subset [0, \infty)$, a raiz quadrada \sqrt{A} é única.

2.3.5 A inversa

A maneira clássica de se obter a inversa por meio do polinômio característico p (ou mínimo) da matriz invertível A é a seguinte: se

$$p(\zeta) = \zeta^m + \lambda_{m-1}\zeta^{m-1} + \dots + \lambda_1\zeta + \lambda_0$$

temos pelo Teorema de Cayley-Hamilton,

$$0 = A^m + \lambda_{m-1}A^{m-1} + \dots + \lambda_1A + \lambda_0I.$$

Multiplicando essa relação por A^{-1} , obtemos

$$\lambda_0A^{-1} = -[\lambda_1I + \dots + \lambda_mA^{m-1}].$$

Como A possui inversa, $\lambda_0 \neq 0$. Obtemos A^{-1} dividindo o lado direito da igualdade anterior por λ_0 . Para uma matriz invertível arbitrária, esse procedimento não é vantajoso com relação ao cálculo da inversa por meio de Gauss-Jordan. Em geral, também o cálculo funcional não é vantajoso. Mas, por exemplo, se a matriz invertível A for simétrica e possuir poucos autovalores, o cálculo funcional pode ser útil.

Exemplo 31. Usando a mesma matriz do Exemplo 30, definimos $f(\zeta) = \frac{1}{\zeta}$. Sabemos que o polinômio característico é $p(\zeta) = (\zeta - 1)(\zeta - 2)$. Vamos obter os coeficientes de $r(\zeta) = \lambda\zeta + \mu$ de modo que

$$\begin{cases} f(1) = r(1) \\ f(2) = r(2) \end{cases} \implies \begin{cases} 1 = \lambda + \mu \\ \frac{1}{2} = 2\lambda + \mu \end{cases}.$$

Assim, $\lambda = -\frac{1}{2}$ e $\mu = \frac{3}{2}$. Assim, temos que $r(\zeta) = -\frac{1}{2}\zeta + \frac{3}{2}$. Portanto,

$$\frac{1}{A} = r(A) = -\frac{1}{2}A + \frac{3}{2}I = \begin{pmatrix} -\frac{1}{2} & 0 \\ -\frac{1}{2} & -1 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Se, ao procurarmos, soluções da equação (3.1) exigirmos, adicionalmente, que a solução $x(t)$ satisfaça a condição inicial

$$x(t_0) = x_0, \quad (3.2)$$

em que x_0 é um vetor do \mathbb{R}^n , estaremos lidando com um problema de valor inicial (essa discussão já foi iniciada na Seção 2.3.1). Mostraremos, inicialmente, a unicidade de solução do problema de valor inicial (3.1)-(3.2).

Teorema 27. O problema de valor inicial (3.1)-(3.2) possui, no máximo, uma solução.

Demonstração: Suponha que existam $x(t) = (\chi_1(t) \dots \chi_n(t))^t$ e $y(t) = (\gamma_1(t) \dots \gamma_n(t))^t$ soluções do problema de valor inicial. Definimos $z(\tau) = (\zeta_1(\tau) \dots \zeta_n(\tau))^t$ por $z(\tau) = x(\tau) - y(\tau)$ e a função auxiliar

$$u(\tau) = \int_{\tau_0}^{\tau} (|\chi_1(\sigma) - \gamma_1(\sigma)| + \dots + |\chi_n(\sigma) - \gamma_n(\sigma)|) d\sigma,$$

onde χ_i são as funções coordenadas de x e γ_i de y . Como cada uma das parcelas da integral anterior é uma função contínua (uma vez que as funções coordenadas de x e y são deriváveis), o Teorema Fundamental do Cálculo garante que, para cada $\tau \in \mathbb{R}$,

$$u'(\tau) = |\chi_1(\tau) - \gamma_1(\tau)| + \dots + |\chi_n(\tau) - \gamma_n(\tau)| = \sum_{i=1}^n |\chi_i(\tau) - \gamma_i(\tau)|. \quad (3.3)$$

Agora, note que, para cada $i \in \{1, \dots, n\}$, $\chi_i(\tau) - \chi_i(\tau_0) = \int_{\tau_0}^{\tau} \chi_i'(\sigma) d\sigma$ e $\gamma_i(\tau) - \gamma_i(\tau_0) = \int_{\tau_0}^{\tau} \gamma_i'(\sigma) d\sigma$. Logo, para cada $i \in \{1, \dots, n\}$,

$$|\chi_i(\tau) - \gamma_i(\tau)| = \left| \int_{\tau_0}^{\tau} (\chi_i'(\sigma) - \gamma_i'(\sigma)) d\sigma \right| \leq \int_{\tau_0}^{\tau} |\chi_i'(\sigma) - \gamma_i'(\sigma)| d\sigma, \quad (3.4)$$

uma vez que $\chi_i(\tau_0) = \gamma_i(\tau_0)$.

Portanto, combinando (3.3) com (3.4),

$$u'(\tau) \leq \sum_{i=1}^n \int_{\tau_0}^{\tau} |\chi_i'(\sigma) - \gamma_i'(\sigma)| d\sigma. \quad (3.5)$$

Note que, para cada $i \in \{1, \dots, n\}$, $\chi_i'(\sigma) = \ell_i x(\sigma)^t$, em que ℓ_i denota a i -ésima linha da matriz A , mais precisamente, sendo

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

Assim,

$$\begin{aligned}\ell_1 &= (\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n}) \\ \ell_2 &= (\alpha_{21}, \alpha_{22}, \dots, \alpha_{2n}) \\ &\vdots \\ \ell_n &= (\alpha_{n1}, \alpha_{n2}, \dots, \alpha_{nn}).\end{aligned}$$

Lembre-se que $x' = Ax$, dessa maneira,

$$\chi'_i(\sigma) = \alpha_{i1}\chi_1(\sigma) + \alpha_{i2}\chi_2(\sigma) + \dots + \alpha_{in}\chi_n(\sigma) = \ell_i x(\sigma)^t.$$

Assim, a desigualdade de Cauchy-Schwarz-Bunyakovsky (veja a Proposição 6) garante que:

$$|\chi'_i(\sigma) - \gamma'_i(\sigma)| = |\ell_i(x(\sigma) - y(\sigma))^t| \leq \|\ell_i\| \|x(\sigma) - y(\sigma)\| \leq \|\ell_i\| \|x(\sigma) - y(\sigma)\|_{\text{sum}}, \quad (3.6)$$

onde $\|\cdot\|$ denota a norma euclidiana em \mathbb{R}^n e $\|x(\sigma) - y(\sigma)\|_{\text{sum}} = \sum_{i=1}^n |\chi_i(\sigma) - \gamma_i(\sigma)|$.

Portanto, combinando (3.6) com (3.5),

$$u'(\tau) \leq \sum_{i=1}^n \|\ell_i\| \int_{\tau_0}^{\tau} \|x(\sigma) - y(\sigma)\|_{\text{sum}} ds \leq \rho u(\tau),$$

onde $\rho = \sum_{i=1}^n \|\ell_i\|$, e ainda $\|x(\sigma) - y(\sigma)\|_{\text{sum}} = u'(\sigma)$ e $\int_{\tau_0}^{\tau} u'(\sigma) d\sigma = u(\tau) - u(\tau_0) = u(\tau)$.

Multiplicando a desigualdade $u'(\tau) \leq \rho u(\tau)$ por $e^{-\rho\tau}$, obtemos

$$u'(\tau)e^{-\rho\tau} \leq \rho u(\tau)e^{-\rho\tau} \implies u'(\tau)e^{-\rho\tau} - \rho u(\tau)e^{-\rho\tau} \leq 0$$

e concluímos que

$$\frac{d}{d\tau} (e^{-\rho\tau} u(\tau)) \leq 0.$$

Daí,

$$\int_{\tau_0}^{\tau} \frac{d}{ds} (e^{-\rho s} u(s)) ds \leq 0 \iff e^{-\rho\tau} u(\tau) - e^{-\rho\tau_0} u(\tau_0) \leq 0 \iff e^{-\rho\tau} u(\tau) \leq 0 \iff u(\tau) \leq 0,$$

onde usamos o fato de $u(\tau_0) = x(\tau_0) - y(\tau_0) = x_0 - x_0 = 0$, pois, x e y resolvem o mesmo PVI. Desta forma $u(\tau) \leq 0$ para todo $\tau \in \mathbb{R}$, donde decorre que $u \equiv 0$ e portanto, $x = y$. \square

Definição 40. Sejam x_1, \dots, x_n soluções do sistema (3.1). Consideremos a matriz $X(\tau)$, cujas colunas são os vetores $x_1(\tau), \dots, x_n(\tau)$. Definimos o **wronskiano** $W(x_1, \dots, x_n)(\tau)$

das soluções x_1, \dots, x_n por

$$W(x_1, \dots, x_n)(\tau) = \det X(\tau).$$

Dizemos que as soluções x_1, \dots, x_n são **linearmente independentes**, se

$$W(x_1, \dots, x_n)(\tau) \neq 0 \quad \forall \tau \in \mathbb{R}.$$

Nesse caso, a matriz $X(\tau)$ é chamada **matriz fundamental** do sistema (3.1). Caso contrário, isto é, se $W(x_1, \dots, x_n)(\tau_1) = 0$ para algum $\tau_1 \in \mathbb{R}$, as soluções são **linearmente dependentes**.

Sabemos que os vetores $x_1(\tau_0), \dots, x_n(\tau_0) \in \mathbb{R}^n$ são linearmente independentes se, e somente se, $W(x_1, \dots, x_n)(\tau_0) \neq 0$. Assim, a independência linear das soluções x_1, \dots, x_n de (3.1) equivale à independência linear dos vetores $x_1(\tau), \dots, x_n(\tau)$ para todo $\tau \in \mathbb{R}$. Note que as combinações lineares $y_1x_1 + \dots + y_jx_j$ de soluções x_1, \dots, x_j de (3.1) também são soluções de (3.1).

Proposição 12. Sejam x_1, \dots, x_n soluções linearmente independentes de (3.1). Então elas constituem uma base do espaço de soluções, isto é, toda solução y de (3.1) é uma combinação linear $\alpha_1x_1 + \dots + \alpha_nx_n$, com $\alpha_1, \dots, \alpha_n \in \mathbb{R}$.

Demonstração: Seja y uma solução de (3.1). Queremos mostrar que existem constantes $\alpha_1, \alpha_2, \dots, \alpha_n$ tais que

$$y(\tau) = \alpha_1x_1(\tau) + \alpha_2x_2(\tau) + \dots + \alpha_nx_n(\tau), \quad \forall \tau \in \mathbb{R}.$$

Consideremos $y(\tau_0)$ e o sistema linear nas incógnitas $\alpha_1, \dots, \alpha_n$:

$$\alpha_1x_1(\tau_0) + \alpha_2x_2(\tau_0) + \dots + \alpha_nx_n(\tau_0) = y(\tau_0).$$

Como $W(\tau_0) \neq 0$ (pois as soluções são linearmente independentes), tal sistema tem solução única. Como as soluções $\alpha_1x_1(\tau) + \dots + \alpha_nx_n(\tau)$ e $y(\tau)$ de (3.1) coincidem no ponto τ_0 , então pelo Teorema 27 elas são idênticas em \mathbb{R} . \square

Observação 24.

1. Essa Proposição justifica a denominação de solução geral para a expressão $\alpha_1x_1 + \dots + \alpha_nx_n$, se $\{x_1, \dots, x_n\}$ for uma base de soluções de (3.1).
2. Seja $X(\tau)$ uma matriz fundamental de (3.1). Podemos escrever a solução geral desse sistema em termos matriciais: definindo o vetor $y = (\alpha_1 \dots \alpha_n)^t \in \mathbb{R}^n$, temos

$$x(\tau) = X(\tau)y. \tag{3.7}$$

Em particular, a resolução de um problema de valor inicial pode ser descrita de maneira bastante simples por meio da matriz fundamental: se $x(\tau_0) = x_0 \in \mathbb{R}^n$, então devemos resolver

$$x(\tau_0) = X(\tau_0)y \implies X(\tau_0)y = x_0.$$

Como as colunas de $X(\tau)$ são linearmente independentes, $X(\tau_0)$ é invertível, donde obtemos $y = X^{-1}(\tau_0)x_0$ e, portanto, substituindo em (3.7), temos

$$x = X(\tau)[X^{-1}(\tau_0)x_0].$$

Salientamos, entretanto, que a obtenção direta de $y = (\alpha_1 \dots \alpha_n)^t$ quase sempre é preferível à obtenção de $X^{-1}(\tau_0)x_0$.

Notamos que vale a equação matricial

$$X' = AX$$

(a derivada da matriz $X(\tau)$ é a derivação de cada uma de suas componentes), pois cada coluna de $X(\tau)$ é solução de $x' = Ax$. Considerado o isomorfismo entre $M_{n \times n}(\mathbb{R})$ e \mathbb{R}^{n^2} , o Teorema 27 é aplicável à equação matricial (3.1). Em virtude da ligação entre os sistemas $X' = AX$ e $x' = Ax$, a matriz fundamental desse último sistema também é chamada solução fundamental.

Proposição 13. Sejam x_1, \dots, x_n soluções do sistema (3.1). Então

- (i) ou $W(x_1, \dots, x_n)(\tau) \equiv 0$ (e as soluções são linearmente dependentes);
- (ii) ou $W(x_1, \dots, x_n)(\tau) \neq 0$ para todo $t \in \mathbb{R}$ (e as soluções são linearmente independentes).

Demonstração: Suponhamos que $W(x_1, \dots, x_n)(\tau_1) = 0$ para $\tau_1 \in \mathbb{R}$. Então, existem constantes nem todas nulas $\gamma_1, \dots, \gamma_n$, tais que $\gamma_1 x_1(\tau_1) + \dots + \gamma_n x_n(\tau_1) = 0$.

Defina $w(\tau) = \gamma_1 x_1(\tau) + \dots + \gamma_n x_n(\tau)$. Então $w(\tau)$ satisfaz (3.1) e $w(\tau_1) = 0$. Pelo Teorema 27, temos que $w(\tau) \equiv 0$. Isto mostra que $x_1(\tau), \dots, x_n(\tau)$ são linearmente dependentes em todos os pontos de \mathbb{R} . □

Definição 41. Sejam x_1, \dots, x_n soluções de (3.1) satisfazendo as condições iniciais $x_i(0) = e_i$, em que $\{e_1, \dots, e_n\}$ é a base canônica do \mathbb{R}^n . Então a matriz $X(\tau) = (x_1(\tau) \dots x_n(\tau))$ é chamada fluxo linear do sistema (3.1) e denotada por

$$e^{A\tau} \quad \text{ou} \quad \exp(A\tau).$$

Temos que $X(\tau) = (x_1(\tau) \dots x_n(\tau))$, assim $X(0) = (x_1(0) \dots x_n(0))$. Com as condições iniciais: $x_1(0) = e_1, x_2(0) = e_2, \dots, x_n(0) = e_n$. Colocando em colunas, obtemos:

$$[e_1, e_2, \dots, e_n] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = I_n$$

Se $e^{A\tau}$ for o fluxo do sistema $x' = Ax$, temos que $W(x_1, \dots, x_n)(0) \neq 0$. Tendo em vista a Proposição 13, vale $W(x_1, \dots, x_n)(\tau) \neq 0$ para todo $\tau \in \mathbb{R}$, de modo que $e^{A\tau}$ é uma solução fundamental do sistema (3.1). Assim, uma maneira de encontrar n soluções linearmente independentes de $x' = Ax$ consiste em obter o fluxo $e^{A\tau}$. A existência do fluxo $e^{A\tau}$ está provada na Seção 2.3. Como as colunas de $e^{A\tau}$ formam uma base do espaço de soluções de (3.1), podemos enunciar o seguinte resultado:

Teorema 28. O problema de valor inicial (3.1)-(3.2) possui uma única solução para todo vetor $x_0 \in \mathbb{R}^n$.

Com o intuito de ilustrar a utilização do Teorema 27 na demonstração de resultados, vamos apresentar algumas propriedades do fluxo $e^{A\tau}$.

Proposição 14. Sejam $A, B \in M_n(\mathbb{R})$. Se $AB = BA$, então

1. $e^{A\tau}B = Be^{A\tau}$.
2. $e^{A\tau}e^{B\tau} = e^{(A+B)\tau}$ se, e somente se, $AB = BA$.

Além disso, se v for um autovetor de A associado ao autovalor λ , então, para todo $\tau \in \mathbb{R}$, v é um autovetor de $e^{A\tau}$ associado ao autovalor $e^{A\tau}$.

Demonstração: 1. Pela definição de cálculo funcional apresentada no Capítulo 2, $e^{A\tau} = r(A)$, onde r é um polinômio que interpola $f_\tau(z) = e^{z\tau}$ nos autovalores de A . Seja $r(z) = \alpha_0(\tau) + \alpha_1(\tau)z + \dots + \alpha_{m-1}(\tau)z^{m-1}$, onde m é o grau do polinômio mínimo de A . Então

$$e^{A\tau} = \alpha_0(\tau)I + \alpha_1(\tau)A + \dots + \alpha_{m-1}(\tau)A^{m-1}.$$

Multiplicando à direita por B , obtemos

$$\begin{aligned} e^{A\tau}B &= \alpha_0(\tau)B + \alpha_1(\tau)AB + \dots + \alpha_{m-1}(\tau)A^{m-1}B \\ &= \alpha_0(\tau)B + \alpha_1(\tau)BA + \dots + \alpha_{m-1}(\tau)BA^{m-1} \\ &= B(\alpha_0(\tau)I + \alpha_1(\tau)A + \dots + \alpha_{m-1}(\tau)A^{m-1}) = Be^{A\tau} \end{aligned}$$

onde, na segunda igualdade, usamos que $AB = BA$ (e por consequência $A^k B = BA^k$ para todo $k \in \mathbb{N}$).

2. (\Rightarrow) Defina $Z(\tau) = e^{A\tau} e^{B\tau}$. Pela Proposição 11 (item 2),

$$Z'(\tau) = \frac{d}{d\tau}[e^{A\tau} e^{B\tau}] = Ae^{A\tau} e^{B\tau} + e^{A\tau} Be^{B\tau}.$$

Pelo item 1), $e^{A\tau} B = Be^{A\tau}$, já que $AB = BA$, assim

$$Z'(\tau) = Ae^{A\tau} e^{B\tau} + Be^{A\tau} e^{B\tau} = (A + B)e^{A\tau} e^{B\tau} = (A + B)Z(\tau).$$

Da condição inicial:

$$Z(0) = e^{A0} e^{B0} = I \cdot I = I.$$

Portanto, $Z(\tau)$ satisfaz

$$Z'(\tau) = (A + B)Z(\tau), \quad Z(0) = I.$$

Mas $e^{(A+B)\tau}$ também satisfaz essa mesma EDO matricial. Pelo Teorema 27, temos:

$$Z(\tau) = e^{(A+B)\tau},$$

ou seja,

$$e^{A\tau} e^{B\tau} = e^{(A+B)\tau}.$$

(\Leftarrow) Reciprocamente, derivando $e^{(A+B)\tau} = e^{A\tau} e^{B\tau}$, encontramos

$$(A + B)e^{(A+B)\tau} = Ae^{A\tau} e^{B\tau} + e^{A\tau} Be^{B\tau}.$$

Nova derivação produz

$$(A + B)^2 e^{(A+B)\tau} = A^2 e^{A\tau} e^{B\tau} + 2Ae^{A\tau} Be^{B\tau} + e^{A\tau} B^2 e^{B\tau}. \quad (3.8)$$

Pela hipótese, $e^{A\tau} e^{B\tau} = e^{(A+B)\tau}$, assim $\frac{d}{d\tau^2}[e^{A\tau} e^{B\tau}] = \frac{d}{d\tau^2}[e^{(A+B)\tau}]$. Tomando $\tau = 0$ em (3.8), obtemos

$$(A + B)^2 = A^2 + 2AB + B^2 \implies AB = BA.$$

E, finalmente, pelo Teorema do mapeamento Espectral (Teorema 25), se f é euclidiana com respeito a A , então $f(A)v = f(\lambda)v$. Tome $f(z) = e^{z\tau}$, que é uma função analítica em todo \mathbb{C} e portanto euclidiana. Então, $e^{A\tau}v = e^{\lambda\tau}v$. Ou seja, v é autovetor de $e^{A\tau}$ com autovalor $e^{\lambda\tau}$. \square

Observação 25. Suponhamos que $A = C^{-1}BC$ para alguma matriz invertível C . Então, é claro que as matrizes A e B têm os mesmos autovalores com as mesmas multiplicidades. Isso quer dizer que o polinômio interpolador $p(\tau) = \alpha_{n-1}(\tau)\tau^{n-1} + \dots + \alpha_1(\tau)\tau + \alpha_0(\tau)$

que produz $e^{A\tau}$ também produz $e^{B\tau}$. Assim,

$$\begin{aligned} e^{A\tau} &= e^{(C^{-1}BC)\tau} = \alpha_{n-1}(\tau)C^{-1}B^{n-1}C + \dots + \alpha_1(\tau)C^{-1}BC + \alpha_0(\tau)I \\ &= C^{-1}[\alpha_{n-1}(\tau)B^{n-1} + \dots + \alpha_1(\tau)B + \alpha_0(\tau)I]C \\ &= C^{-1}e^{B\tau}C. \end{aligned}$$

Definição 42. Dizemos que os sistemas

$$x' = Ax \quad \text{e} \quad x' = Bx$$

(ou os fluxos que lhes são associados) são **linearmente conjugados**, se existir uma matriz invertível C tal que

$$C(e^{A\tau}x) = e^{B\tau}Cx.$$

Teorema 29. A aplicação linear $h(x) = Cx$ é uma conjugação linear entre os sistemas

$$x' = Ax \quad \text{e} \quad x' = Bx$$

se, e somente se, C for invertível e $CA = BC$. Em particular, esses sistemas são linearmente conjugados se, e somente se, as matrizes A e B forem semelhantes.

Demonstração: (\Rightarrow) Pela observação 25, se $A = C^{-1}BC$, então $e^{A\tau} = C^{-1}e^{B\tau}C$. Multiplicando à esquerda por C :

$$Ce^{A\tau} = e^{B\tau}C.$$

Aplicando um vetor x qualquer, obtemos:

$$Ce^{A\tau}x = e^{B\tau}Cx,$$

ou seja, $h(x) = Cx$ é uma conjugação linear entre os sistemas $x' = Ax$ e $x' = Bx$.

(\Leftarrow) Reciprocamente, derivando a igualdade $Ce^{A\tau}x = e^{B\tau}Cx$ com relação a τ obtemos

$$CAe^{A\tau}x = Be^{B\tau}Cx.$$

Tomando $\tau = 0$, vem $CAx = BCx$ para todo x . Como C é invertível, podemos multiplicar à direita por C^{-1} , assim

$$CAC^{-1} = B,$$

ou equivalentemente,

$$A = C^{-1}BC.$$

Ou seja, A, B são semelhantes. □

Referências Bibliográficas

- [1] J. L. Boldrini, S. I. Costa, V. Figueredo, and H. G. Wetzler. *Álgebra linear*. Harper & Row, 1980.
- [2] H. P. Bueno. *Álgebra linear Um Segundo Curso*. SBM - Rio de Janeiro, 2006.
- [3] H. P. Bueno. *Funções de Matrizes*. UFMG, 2015.
- [4] L. B. Cardoso and F. P. de Souza. O cálculo funcional: Abordagens do fluxo.
- [5] D. G. de Figueiredo. *Análise I*. Rio de Janeiro: LTC, 2 edition, 1996.
- [6] N. J. Higham. *Function of Matrices, Theory and Computation*. Philadelphia: Society for Industrial and Applied Mathematics, 2008.
- [7] K. Hoffman and R. Kunze. *Álgebra linear, por Kenneth Hoffman e Ray Kunze; traduzido por Adalberto P. Bergamasco*. Polígono S.A, 1970.
- [8] G. Iezzi and S. Hazzan. *Fundamentos de matemática elementar, 4: sequências, matrizes, determinantes, sistemas*. Atual, 2013.
- [9] E. L. Lima. *Análise Real Volume 2*. Rio de Janeiro, 2004.
- [10] E. L. Lima. *Curso de análise vol.1*. Rio de Janeiro: IMPA, 1 edition, 2014.
- [11] M. G. Soares. *Cálculo em uma variável complexa*. Rio de Janeiro: IMPA, 1 edition, 2014.
- [12] J. Stewart. *Cálculo*, volume 2. Cengage Learning, São Paulo, 7 edition, 2013. Tradução da 7^a edição norte-americana.
- [13] G. Ávila. *Análise Matemática Para Licenciatura*. Edgard Blücher LTDA, São Paulo, 1 edition, 2001.
- [14] G. Ávila. *Variáveis complexas e aplicações*. Rio de Janeiro: LTC, 3 edition, 2008.