Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática
Programa de Mestrado Profissional
em Matemática em Rede Nacional

Quase todas as matrizes complexas são diagonalizáveis - uma abordagem transdisciplinar

Rubens Ricardo Miranda Cardoso

Brasília

2019

Rubens Ricardo Miranda Cardoso

Quase todas as matrizes complexas são diagonalizáveis - uma abordagem transdisciplinar

Dissertação apresentada ao Departamento de Matemática da Universidade de Brasília, como parte dos requisitos do Programa de Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, para obtenção do grau de Mestre.

Universidade de Brasília - UnB
Instituto de Ciências Exatas
Departamento de Matemática - MAT
PROFMAT - SBM

Orientador: Prof. Dr. Theo Allan Darn Zapata

Brasília
2019

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

CC268q

Cardoso, Rubens Ricardo Miranda
Quase todas as matrizes complexas são diagonalizáveis - uma abordagem transdisciplinar / Rubens Ricardo Miranda Cardoso; orientador Theo Allan Darn Zapata. -- Brasília, 2019.
69 p.

Dissertação (Mestrado - Mestrado Profissional em Matemática) -- Universidade de Brasília, 2019.

1. Matrizes diagonalizáveis. 2. Densidade. 3. Conjunto de medida-nula. 4. Topologia de Zariski. I. Zapata, Theo Allan Darn, orient. II. Título.

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Quase todas as matrizes complexas são diagonalizáveis —uma abordagem transdisciplinar

por

RUBENS RICARDO MIRANDA CARDOSO

Dissertação apresentada ao Departamento de Matemática da Universidade de Brasília, como parte dos requisitos "Programa" de Mestrado Profissional em Matemática em Rede Nacional – PROFMAT, para obtenção do grau de

MESTRE EM MATEMÁTICA

Brasília, 19 de julho de 2019.

Comissão Examinadora:

Theo Allan Darn Zapata

Prof. Theo Allan Darn Zapata (Orientador)

Leandro Martins Cioletti

Prof. Leandro Marns Cioletti – MAT/UnB

Igor dos Santos Lima

Prof. Igor dos Santos Lima - MAT/UnB

Mauro Ribeiro de Oliveira Júnior

Prof. Mauro Ribeiro de Oliveira Júnior – CEF/Unieuro

Aos meus pais, à minha esposa e ao meu filho Heitor.

Agradecimentos

Agradeço a Deus pelas suas bençãos em minha vida e pelo seu eterno amor.

Aos meus pais pelo carinho, zelo, amor incondicional e pelos profícuos ensinamentos os quais cultivo e pratico.

À minha amada esposa, que sempre me apoiou nas lides da vida e, nessa jornada, contribuiu de maneira substancial.

Aos professores e alunos do PROFMAT que ao longo do curso contribuíram para o meu aprimoramento pessoal e, em especial, ao orientador deste trabalho o qual trouxe oportunas e importantes considerações.

"Uma grande descoberta envolve a solução de um grande problema, mas há uma semente de descoberta na solução de qualquer problema. Seu problema pode ser modesto; porém, se ele desafiar a sua curiosidade e fizer funcionar a sua capacidade inventiva, e caso você o resolva sozinho, então poderá experimentar a tensão e o prazer do triunfo da descoberta. "
George Polya

"A unidade orgânica da matemática é inerente à sua natureza, uma vez que a matemática se constitui no fundamento de todo conhecimento exato dos fenômenos naturais."
David Hilbert

Resumo

Nos primeiros cursos de Álgebra Linear, são apresentados aos estudantes exemplos de matrizes complexas que não são diagonalizáveis. Nesse contexto, é razoável indagar se a maioria das matrizes complexas são diagonalizáveis. A presente dissertação se debruça nessa questão. Por meio da Topologia, faz-se uma prova da densidade do conjunto das matrizes complexas diagonalizáveis, utilizando o Teorema da Decomposição de Schur. No contexto da medida e da integral de Lebesgue, prova-se que o conjunto das matrizes complexas não diagonalizáveis tem medida nula. Na perspectiva da Álgebra, por meio da topologia de Zariski, dá-se uma demonstração da densidade do conjunto das matrizes complexas diagonalizáveis, usando somente polinômios. Discutem-se as interdependências entre os resultados obtidos por meio da Topologia, da Medida e da Álgebra.

Palavras-chaves: Matrizes diagonalizáveis. Densidade. Conjunto de medida-nula. Topologia de Zariski.

Abstract

In the first courses of Linear Algebra, examples of complex matrices that are not diagonalizable are presented to students. In this context, it is reasonable to ask whether most of the complex matrices are diagonalizable. The present dissertation deals with this issue. By means of the topology, the density of the set of the diagonalizable complex matrices is proved using the Schur Decomposition Theorem. In the context of the Lebesgue measure and integral, it is proved that the set of non-diagonalizable complex matrices has null measure. In the Algebra perspective, through the Zariski topology, we demonstrate the density of the set of complex diagonalizable matrices using only polynomials. We discuss the interdependencies among the results obtained through Topology, Measure and Algebra.

Key-words: Diagonalizable matrices. Density. Null measure set. Topology of Zariski.

Lista de símbolos

$\mathbb{C}$	Conjunto dos números complexos
$\mathbb{N}$	Conjunto dos números naturais
$\mathbb{C}[X]$	Conjunto dos polinômios sobre $\mathbb{C}$ em uma indeterminada X
P'	Primeira derivada do polinômio P
$res(P,Q)$	Resultante dos polinômios $P(X)$ e $Q(X)$
$\Delta(P)$	Discriminante do polinômio $P(X)$
$\mathbb{C}[X_1,\ldots,X_n]$	Conjunto dos polinômios sobre $\mathbb{C}$ nas n indeterminadas $X_1,\ldots,X_n$
$Z(P)$	Conjunto dos zeros do polinômio P
$M_n(\mathbb{C})$	Conjunto das matrizes quadradas de ordem n, com entradas em $\mathbb{C}$
I	Matriz identidade
$GL_n(\mathbb{C})$	Conjunto das matrizes quadradas invertíveis de ordem n, com entradas em $\mathbb{C}$
det	Determinante
$\vec{v}$	Vetor
P_M	Polinômio característico da matriz M
$\bar{z}$	Conjugado do número complexo z
$\langle \vec{u}, \vec{v} \rangle$	Produto interno dos vetores $\vec{u}$ e $\vec{v}$
τ	Topologia
(X,τ)	Espaço topológico
$A \times B$	Produto cartesiano de A por B
$\prod$	Produtório
$\mathcal{D}$	Conjunto das matrizes diagonalizáveis
μ_n	Medida de Lebesgue em $\mathbb{R}^n$
$\chi_{Z(P)}$	Função característica do conjunto $Z(P)$

$\int_X f d\mu$	Integral de Lebesgue da função f no conjunto X
$\mathbb{K}$	Corpo
$\mathbb{K}[X_1, \ldots, X_n]$	Anel dos polinômios sobre $\mathbb{K}$ nas indeterminadas $X_1, \ldots, X_n$
$\langle S \rangle$	Ideal gerado por um subconjunto S
$Z(S)$	Conjunto dos zeros comuns a todos os polinômios de S

Sumário

Introdução

"Muitas das atuais teorias matemáticas surgiram da Ciência Aplicada, e só depois adquiriram aquele aspecto axiomático e abstrato que tanto dificulta o seu aprendizado".

V.I.Arnold.

Pensemos, inicialmente, no seguinte problema:

Dentre as matrizes quadradas reais de ordem n, há "mais" matrizes invertíveis ou não invertíveis?

Consideremos, por simplicidade e didática, uma matriz quadrada de ordem 2

$$\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix},$$

em que x_1, x_2, x_3 e x_4 são números reais.

Em Álgebra Linear, ensina-se que uma matriz quadrada será invertível se, e somente se, o seu determinante for diferente de zero. Assim, existe uma bijeção entre o conjunto das matrizes de ordem 2, que não possuem inversa, e o conjunto das soluções da equação polinomial

$$x_1x_4 - x_2x_3 = 0.$$

Mas, desde do ensino fundamental, sabe-se que as equações polinomiais em uma variável, com coeficientes reais, têm pouquíssimas soluções no conjunto universo dos números reais. Cada equação do 1º grau possui apenas uma solução; a do 2º grau, ora uma, ora duas, às vezes nenhuma; e a de n-ésimo grau, no máximo n soluções.

Nada muito diferente acontece com as equações polinomiais de mais de uma variável. Por exemplo, o conjunto das soluções reais da equação $x_1^2 + x_2^2 - 4 = 0$ representa, em termos geométricos, uma circunferência de centro na origem $(0, 0)$ e de raio 2, no plano cartesiano. Ou seja, o conjunto-solução dessa equação é uma curva no plano. Ora, o que é uma curva em face do plano bidimensional, senão um conjunto "pequeno".

Se a equação fosse $x_1 + x_2 + x_3 = 0$, o seu conjunto de soluções reais seria um plano em $\mathbb{R}^3$. Novamente indagaríamos, o que seria um plano ante o espaço tridimensional?

Do exposto, podemos inferir que o conjunto-solução da equação $x_1x_4 - x_2x_3 = 0$ é também "pequeníssimo", comparado ao seu conjunto universo, neste caso $\mathbb{R}^4$.

Percebemos que a ocorrência de matrizes não invertíveis de ordem 2 é tão rara quanto os zeros de polinômios. Para matrizes de ordem superior a 2, o raciocínio é idêntico. Logo, não somente há "mais" matrizes invertíveis que não invertíveis, mas também "quase todas" são invertíveis.

Na literatura, o sentido assumido pela expressão "quase todas" é determinado pela teoria desenvolvida e pelos seus propósitos. Nesta dissertação, não é diferente. Na Topologia, a expressão "quase todas" refere-se ao conceito de "densidade" e, na Teoria da Medida, às propriedades que valem em quase todo ponto de um conjunto X, *i.e.*, existe um conjunto de medida nula $A \subset X$ tal que a propriedade vale para todo ponto x do complementar de A.

Por exemplo, consideram-se o conjunto dos números reais e o subconjunto dos números irracionais. O conjunto dos números irracionais é "denso" no dos números reais, pois todo intervalo aberto $I = (a, b)$ de números reais, com $a < b$, contém sempre um número irracional. Intuitivamente, a noção de densidade de um subconjunto A de um conjunto X remete à ideia de que aquele está suficientemente "pulverizado" a ponto de qualquer "amostra" de X, por menor que esta seja, conterá algum elemento de A. Ou seja, do ponto de vista da Topologia, o subconjunto A não é "muito pequeno".

Por outro lado, do ponto de vista da Teoria da Medida, o conjunto dos números racionais é um "conjunto de medida nula". Isso significa que, para qualquer número real $\varepsilon > 0$, existe uma correspondente coleção enumerável $n = 1, 2, 3, \cdots$, de intervalos abertos $I_n = (a_n, b_n)$ de números reais tais que a união desses intervalos cobrem o conjunto dos números racionais ($i.e.$, $\mathbb{Q} \subset \bigcup_{n\geq 1} I_n$) e a soma dos comprimentos desses intervalos é menor que um número real $\varepsilon > 0$ dado ($i.e.$, $\sum_{n\geq 1}(b_n - a_n) < \varepsilon$). Intuitivamente, o conjunto dos racionais é "pequeno" na perspectiva da Teoria da Medida e, por conseguinte, o dos irracionais não é "pequeno".

Outra terminologia também empregada neste trabalho, é a de matriz diagonalizável. Uma matriz quadrada A é dita diagonalizável se existir uma matriz quadrada C invertível, cuja inversa é C^{-1}, tal que o produto $C^{-1}AC$ é uma matriz diagonal.

Agora, passemos a considerar o problema desta pesquisa:

"Quase todas" as matrizes quadradas complexas são diagonalizáveis? Nota-se que o problema das matrizes invertíveis foi reformulado em termos de polinômios e essa abordagem foi promissora. Assim, é natural indagar-se: seria possível reformularmos o problema das matrizes diagonalizáveis, também, em termos de polinômios? Esta dissertação não só responde positivamente a essa pergunta, como mostra como isso pode ser feito.

As dissertações relacionadas ao tema matrizes diagonalizáveis, produzidas no âmbito do Programa de Mestrado Profissional em Matemática em Rede Nacional-PROFMAT, investigaram o emprego da diagonalização de matrizes no ensino de formas quadráticas e de reconhecimento de cônicas, com o uso de aplicativos educacionais, nas mais variadas sequências didáticas, vide (NETO, 2013) e (SILVA et al., 2016). O público-alvo destes trabalhos foram os estudantes do ensino médio e a diagonalização de matrizes foi estudada, num contexto predominantemente pragmático.

Diferentemente das abordagens predecessoras, porém as complementa, esta dissertação tem como público-alvo os professores do ensino básico e os estudantes dos cursos de graduação em matemática, que buscam entender por que quase todas as matrizes complexas são diagonalizáveis, num texto de linguagem clara e simples. O contexto desta exposição é sobretudo teórico e transdisciplinar, permeando a Álgebra Linear, a Topologia, a Teoria da Medida e a Álgebra. As demonstrações dos Teoremas 2.2.4, 2.2.6, 3.2.5 e 4.2.4 não são difíceis e podem ser encontradas na literatura.

Para conveniência do leitor, introduzimos referências em língua portuguesa para Álgebra Linear, Topologia, Teoria da Medida e Álgebra: (LIMA, 2014), (LIMA, 2009), (ISNARD, 2013) e (GARCIA; LEQUAIN, 2013).

A presente dissertação está organizada em quatro capítulos.

O Capítulo 1, Polinômios e Matrizes, tem o objetivo de apresentar as definições, as notações e os teoremas que serão utilizados nos capítulos seguintes. Este está dividido nas Seções 1.1 e 1.2. Na Seção 1.1, são definidos os polinômios, a resultante e o discriminante. Ademais, o Teorema Fundamental da Álgebra (Teorema 1.1.4) e a Proposição 1.1.3 são mencionados para conveniência do leitor. O Lema 1.1.8 e o Corolário 1.1.9 são os resultados mais importantes desta seção. Na Seção 1.2, são introduzidas as matrizes, o seu polinômio característico, os autovalores e autovetores. A diagonalização de matrizes também é tratada. O Lema 1.2.3 e o Teorema da Decomposição de Schur (Teorema 1.2.5) são os fatos principais desta seção.

O Capítulo 2 é o de Topologia. As definições de topologia, espaço topológico, função contínua e de densidade são expostas na Seção 2.1. Na Seção 2.2, são dadas duas demonstrações da densidade das matrizes diagonalizáveis (Teoremas 2.2.4 e 2.2.6).

O Capítulo 3 de Medida possui também duas seções. Na Seção 3.1, faz-se uma breve apresentação da medida e da integral de Lebesgue. O Terorema 3.2.2 afirma que o conjunto dos zeros de polinômios tem medida nula e o Teorema 3.2.5 assegura que o conjunto das matrizes diagonalizáveis possui medida de Lebesgue total, o seu complementar tem medida nula, esses são objetos da Seção 3.2.

O Capítulo 4 de Álgebra foi organizado nas Seções 4.1 e 4.2. A primeira trata dos conceitos de corpos, anéis, ideais e da topologia de Zariski. O Teorema da Base de Hilbert (Teorema 4.1.2) é demonstrado nessa seção. A segunda apresenta o Lema 4.2.3, o qual afirma que todo aberto não-vazio de Zariski é denso na topologia usual de $\mathbb{R}^n$. Encerra-se a presente seção com o Teorema 4.2.4 – que prova que o conjunto das matrizes diagonalizáveis é denso, relativamente tanto a topologia de Zariski quanto a topologia usual em $\mathbb{R}^n$ – e a Proposição 4.2.5 – que demonstra as principais propriedades da topologia de Zariski.

Por fim, elaboram-se os Apêndices A e B. No primeiro, são apresentados, com um pouco mais de detalhes, os principais resultados para a construção da medida e da integral de Lebesgue em $\mathbb{R}^n$. Há, ainda, no Apêndice A, um diagrama que ilustra a interdependência lógica desses resultados. No segundo, faz-se uma exposição de exemplos concretos de polinômios em duas e três indeterminadas, que ilustram o uso do Teorema de Fubini, na demonstração do Lema 3.2.4.

1 Polinômios e Matrizes

"Enquanto a Álgebra e a Geometria estiveram separadas, seus progressos foram lentos e suas aplicações limitadas. No entanto, quando estas duas ciências foram unidas, deram uma a outra renovada vitalidade e seguiram rapidamente rumo à perfeição".

Lagrange

O objetivo deste capítulo é apresentar as definições, as notações e os teoremas que serão utilizados nos capítulos seguintes. Alguns teoremas e proposições deste capítulo são enunciados sem demonstração, por serem resultados secundários desta dissertação e amplamente estudados na literatura. Adotamos, como referência para este capítulo, as obras (KOSTRIKIN; MANIN, 1989), (HIRSCH; SMALE, 1974) e (ATIYAH, 1969).

1.1 Polinômios

Nesta seção, são definidos polinômios, resultante e discriminante. Ademais, o Teorema Fundamental da Álgebra (Teorema 1.1.4) e a Proposição 1.1.3 são apresentados. O Lema 1.1.8 e Corolário 1.1.9 são os resultados mais importantes desta seção.

Chama-se de ***polinômio*** sobre $\mathbb{C}$ em uma indeterminada X a uma expressão formal $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$, em que $a_i \in \mathbb{C}$ e $n \in \mathbb{N}$. Os números $a_0, a_1, a_2, \cdots, a_n$ são os ***coeficientes***; as parcelas $a_0, a_1X, a_2X^2, \cdots, a_nX^n$, os ***termos***; o número complexo a_n, seu ***coeficiente líder*** e, se o polinômio $P(X)$ não for identicamente nulo, o número natural n será o ***grau*** do polinômio $P(X)$.

Denotamos por $\mathbb{C}[X]$ o conjunto de todos os polinômios sobre $\mathbb{C}$ em uma indeterminada X.

As operações de adição $+$ e multiplicação $\cdot$ em $\mathbb{C}[X]$ são definidas de forma usual, a saber: dados $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ e $Q(X) = b_nX^n + b_{n-1}X^{n-1} + \ldots + b_1X + b_0$ elementos de $\mathbb{C}[X]$, definimos

$$P(X) + Q(X) = \sum_i (a_i + b_i)x^i;$$

$$P(X) \cdot Q(X) = \sum_k c_kx^k, \text{ em que } c_k = \sum_{i+j=k} a_ib_j.$$

Se P e Q são polinômios não-nulos, então $\text{grau}(PQ) = \text{grau}(P) + \text{grau}(Q)$.

Proposição 1.1.1 (Divisão Euclidiana) *Sejam $P(X)$ e $S(X)$ polinômios na indeterminada X, com coeficientes complexos, e $S(X) \neq 0$. Então, existe um único polinômio $Q(X)$ tal que $P(X) - Q(X)S(X)$ é 0 ou de grau menor do que o grau do polinômio $S(X)$.*

Demonstração: Se $P(X)$ for identicamente nulo, a proposição é evidentemente válida. Caso contrário, procedemos por indução sobre o grau de $P(X)$. O que conclui a prova. ■

Da Divisão Euclidiana, resultam as Proposições 1.1.2 e 1.1.3. Para uma demonstração dessas proposições, o leitor poderá consultar (GARCIA; LEQUAIN, 2013, Seção I.3) e (ZAPATA, 2019, p. 26).

Proposição 1.1.2 *Seja $\mathcal{I}$ um conjunto não-vazio de polinômios sobre $\mathbb{C}$, em uma indeterminada. Se $\mathcal{I}$ é fechado com respeito à subtração e satisfaz a condição " se um polinômio P pertence a $\mathcal{I}$, então todos os múltiplos de P pertencem a $\mathcal{I}$", então $\mathcal{I}$ consiste de todos os múltiplos de algum polinômio Q, unicamente determinado salvo de uma multiplicação por uma constante.*

Sejam $P(X)$ e $Q(X)$ polinômios na indeterminada X, com coeficientes complexos. Diz-se que $Q(X)$ ***divide*** $P(X)$ se, e somente se, existe um polinômio $S(X)$ tal que $P(X) = Q(X) \cdot S(X)$. Os polinômios $P(X)$ e $Q(X)$ serão denominados ***coprimos***, se não existir um polinômio não constante $R(X)$ que os divida simultaneamente.

O Lema 1.1.8 fornece uma condição suficiente e necessária para que dois polinômios sejam coprimos. A próxima Proposição 1.1.3 será útil para demonstrar esse Lema.

Proposição 1.1.3 *Sejam $P(X)$ e $Q(X)$ em $\mathbb{C}[X]$ dois polinômios coprimos. Então, existem polinômios $U(X)$ e $V(X) \in \mathbb{C}[X]$ tais que*

$$P(X) \cdot U(X) + Q(X) \cdot V(X) = 1.$$

Seja $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0$ um polinômio com coeficientes complexos. Definimos a ***derivada*** $P'(X)$ de $P(X)$ por $P'(X) = 0$, quando $n = 0$ e, em geral, por

$$P'(X) = n a_n X^{n-1} + (n-1) a_{n-1} X^{n-2} + \ldots + 2 a_2 X + a_1.$$

Um número complexo r é dito ***raiz*** de um polinômio $P(X) \in \mathbb{C}[X]$ se, somente se, $P(r) = 0$. Diz-se que r é uma ***raiz de multiplicidade*** $m \geq 1$ do polinômio $P \in \mathbb{C}[X]$

se $P(X) = (X - r)^m \cdot Q(X)$ e $Q(r) \neq 0$, em que $Q(X) \in \mathbb{C}[X]$. Uma raiz será ***simples*** quando sua multiplicidade for igual a 1.

Em matemática, a existência de certos objetos é garantida por alguns teoremas, chamados teoremas de existência, que não necessariamente exibem ou mostram como se constroem tais objetos matemáticos. Esse é o caso do Teorema Fundamental da Álgebra que assegura a existência de raiz complexa de polinômios em uma indeterminada cujos coeficientes são números complexos.

Teorema 1.1.4 (Teorema Fundamental da Álgebra) *Todo polinômio de grau $n \geq 1$ em uma indeterminada, com coeficientes complexos, possui pelo menos uma raiz complexa.*

Uma consequência do Teorema Fundamental da Álgebra, no contexto das matrizes diagonalizáveis, é que toda matriz quadrada com entradas complexas possui ao menos um autovalor.

Sejam $P(X)$ e $Q(X)$ em $\mathbb{C}[X]$ dois polinômios não-nulos. A **resultante** $res(P, Q)$ é dada pelo produto

$$p^{\text{grau de Q}} \cdot q^{\text{grau de P}} \cdot \prod_{x,\, y \,\in\, \mathbb{C},\;\; P(x)=Q(y)=0} (x - y)$$

em que p e q são os coeficientes líderes, respectivamente, dos polinômios $P(X)$ e $Q(X)$. Havendo raízes múltiplas, os fatores serão repetidos de acordo com suas multiplicidades.

Exemplo 1.1.5 *Consideremos os polinômios $P(X) = aX^2 + bX + c$ e $Q(X) = 2aX + b$. O coeficiente líder de $P(X)$ é a, seu grau é 2 e as suas raízes são $\frac{-b+\sqrt{b^2-4ac}}{2a}$ e $\frac{-b-\sqrt{b^2-4ac}}{2a}$. Já o polinômio $Q(X)$ tem coeficiente líder $2a$, grau 1 e raiz $\frac{-b}{2a}$. Daí, a resultante*

$$\begin{aligned} res(P, Q) &= a^1 \cdot (2a)^2 \cdot \left(\frac{-b + \sqrt{b^2 - 4ac}}{2a} - \frac{-b}{2a} \right) \left(\frac{-b - \sqrt{b^2 - 4ac}}{2a} - \frac{-b}{2a} \right) \\ &= -a \cdot (b^2 - 4ac). \end{aligned}$$

Definição 1.1.6 *O **discriminante** $\Delta(P)$ de um polinômio não-nulo $P(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0$ é dado por*

$$\Delta(P) = \frac{(-1)^{\frac{n(n-1)}{2}}}{a_n} \cdot res(P, P'),$$

em que $res(P, P')$ é a resultante do polinômio P e da sua derivada P'.

Considere, por exemplo, o polinômio $P(X) = X^n - 1$, com coeficiente líder 1 e de grau $n > 1$. As suas raízes são as n-ésimas raízes da unidade

$e^{\frac{2k\pi i}{n}} = \cos\left(\frac{2k\pi}{n}\right) + i\,\mathrm{sen}\left(\frac{2k\pi}{n}\right)$, em que $k = 0, 1, \ldots, n-1$ e i é a unidade imaginária. A sua derivada $P'(X) = nX^{n-1}$ é um polinômio de grau $n-1$, com coeficiente líder n. A raiz de $P'(X) = nX^{n-1}$ é 0. Assim, o discriminante do polinômio é

$$\begin{aligned}\Delta(P) &= (-1)^{\frac{n(n-1)}{2}} \cdot 1^{n-1} \cdot n^n \cdot \prod_{k=0}^{n-1} e^{\frac{2k\pi i}{n}} \\ &= n^n \cdot (-1)^{\frac{n(n-1)}{2}} e^{\pi(n-1)i} = \begin{cases} n^n, \text{ se } n-1 \text{ ou } n-2 \text{ é divisível por } 4; \\ -n^n, \text{caso contrário.} \end{cases}\end{aligned}$$

Observação 1.1.7 *Note que o discriminante $\Delta(P)$ é um polinômio cujas indeterminadas são os coeficientes do polinômio P. De fato, o discriminante $\Delta(P)$ é dado por somas e produtos das raízes do polinômio P e de sua derivada P'. Por sua vez, tanto as raízes do polinômio P quanto as de sua derivada P' estão relacionadas com os coeficientes de P, também por meio de somas e produtos. O Exemplo 1.1.5 nos mostra claramente que o discriminante do polinômio $P(X) = aX^2 + bX + c$ é o polinômio*

$$\Delta(P) = Q(A, B, C) = \frac{(-1)^{\frac{2(2-1)}{2}}}{A} \cdot (-A)(B^2 - 4AC) = B^2 - 4AC,$$

cujas indeterminadas A, B e C são os coeficientes do polinômio P.

Lema 1.1.8 *Sejam $P(X)$ e $Q(X)$ em $\mathbb{C}[X]$ dois polinômios não constantes. Então, a resultante $res(P, Q)$ é diferente de 0 se, e somente se, $P(X)$ e $Q(X)$ são coprimos.*

Demonstração: Se os polinômios $P(X)$ e $Q(X)$ tiverem um divisor comum; então, eles possuiriam uma raiz comum em $\mathbb{C}$ e, por conseguinte, $res(P, Q) = 0$. Reciprocamente, suponha que os polinômios $P(X)$ e $Q(X)$ sejam coprimos. Da Proposição 1.1.3, há dois polinômios $U(X)$ e $V(X) \in \mathbb{C}[X]$ de modo que $P(X) \cdot U(X) + Q(X) \cdot V(X) = 1$. Dessa identidade, infere-se que os polinômios $P(X), Q(X) \in \mathbb{C}[X]$ não apresentam nenhuma raiz em comum e, portanto, $res(P, Q) \neq 0$. O que conclui a demonstração do lema. ■

Uma consequência natural do Lema 1.1.8 é o seguinte corolário.

Corolário 1.1.9 *Um polinômio $P(X) \in \mathbb{C}[X]$ possui só raízes simples se, somente se, $\Delta(P) \neq 0$.*

Demonstração: O polinômio $P(X)$ terá somente raízes simples se, e somente se, for coprimo com a sua derivada $P'(X)$. Aplicando o Lema 1.1.8, chega-se ao resultado desejado. ■

Chamamos de ***polinômio sobre*** $\mathbb{C}$ ***em*** n ***indeterminadas*** $X_1, \ldots, X_n$ a uma expressão formal

$$P(X_1, \ldots, X_n) = \sum_{\substack{0 \le i_1 \le r_1 \\ \cdots \\ 0 \le i_n \le r_n}} a_{i_1 \ldots i_n} X_1^{i_1} \ldots X_n^{i_n},$$

em que $a_{i_1 \ldots i_n} \in \mathbb{C}$ e os $r_1, \ldots r_n$ são inteiros não negativos. Cada termo $X_1^{i_1} \ldots X_2^{i_n}$ é denominado ***monômio*** e o seu ***grau*** é dado por $i_1 + \ldots + i_n$. Define-se o grau de um polinômio em n indeterminadas como o maior dos graus de seus monômios não-nulos. Um polinômio é dito ***homogêneo*** de grau m se todos os seus monômios têm grau m.

Dado um polinômio $P(X_1, \ldots, X_n)$ sobre $\mathbb{C}$ nas n indeterminadas $X_1, \ldots, X_n$, a soma dos seus monômios de grau m é um polinômio homogêneo de grau m nomeado ***componente homogênea de grau m*** do polinômio. Logo, todo polinômio é a soma de polinômios homogêneos de graus dois a dois distintos, porque ele é a soma das suas componentes homogêneas.

Por exemplo, consideremos o polinômio

$$P(X_1, X_2, X_3) = 6 - 5X_1 - 2X_2 + X_3 - 3X_3^3 + 5X_2X_3^2 + 8X_1^3X_2 - 7X_3^2X_2X_1.$$

Esse polinômio é de grau 4 e suas componentes homogêneas são: de grau zero, 6; de grau 1, $-5X_1 - 2X_2 + X_3$; de grau 3, $5X_2X_3^2$ e de grau 4, $8X_1^3X_2 - 7X_1^2X_2X_3$.

Denotamos por $\mathbb{C}[X_1, \ldots, X_n]$ o conjunto de todos os polinômios sobre $\mathbb{C}$ nas n indeterminadas $X_1, \ldots, X_n$.

Seja $P(X_1, \ldots, X_n)$ um polinômio sobre $\mathbb{C}$ nas n indeterminadas $X_1, \ldots, X_n$. O ***conjunto dos zeros do polinômio*** P é definido por

$$Z(P) = \{(a_1, \ldots, a_n) \in \mathbb{C}^n : P(a_1, \ldots, a_n) = 0\}.$$

1.2 Matrizes

Nesta seção, são introduzidas as matrizes, o seu polinômio característico, os autovalores e autovetores. A diagonalização de matrizes também é tratada. O Lema 1.2.3 e o Teorema da Decomposição de Schur (Teorema 1.2.5) são os fatos principais desta seção.

Consideram-se os subconjuntos de números naturais $I = \{1, 2, \ldots, m\}$ e $J = \{1, 2, \ldots, n\}$. Uma ***matriz complexa*** m-por-n é uma função $f : I \times J \longrightarrow \mathbb{C}$. As imagens de f são chamadas de ***entradas*** da matriz e são dispostas em um arranjo retangular de m linhas e n colunas, de modo que a imagem $f(i, j) = a_{ij}$ se localize na interseção da

i-ésima linha com a j-ésima coluna

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

O vetor $(a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbb{C}^n$ é o i-ésimo ***vetor-linha*** da matriz e o vetor $(a_{1j}, a_{2j}, \dots, a_{mj}) \in \mathbb{C}^m$ é o j-ésimo ***vetor-coluna*** da matriz.

Uma matriz

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

com n linhas e n colunas é chamada **matriz quadrada de ordem** n e as entradas $a_{11}, a_{22}, \dots, a_{nn}$ constituem a sua ***diagonal principal***. As matrizes quadradas, cujas entradas fora da diagonal principal são todas zero, são chamadas de **matrizes diagonais**.

Outro tipo especial de matrizes de ordem n são as matrizes triangulares. Uma matriz quadrada é ***matriz triangular superior*** se todas as suas entradas abaixo da diagonal principal são zero e, ***matriz triangular inferior***, quando todas as suas entradas acima da diagonal principal são zero.

Nosso estudo se focará nas matrizes quadradas. Denotamos por $M_n(\mathbb{C})$ o conjunto das matrizes quadradas de ordem n, cujas entradas são números complexos.

Com o intuito de tornar o texto autocontido, apresentaremos, a partir desse ponto, definições de matrizes invertível, inversa, transposta, diagonalizável e conjugada.

Uma matriz $A \in M_n(\mathbb{C})$ é ***invertível*** se existir uma matriz A^{-1}, chamada a ***inversa*** de A, tal que $A \cdot A^{-1} = I$, em que I representa a matriz identidade (matriz diagonal em que todas as entradas da diagonal principal são iguais a 1). É bem conhecido que uma matriz será invertível se, somente se, o seu determinante for não-nulo ou, equivalentemente, os seus vetores colunas formam um conjunto linearmente independente.

Denotaremos por $GL_n(\mathbb{C})$ o conjunto das matrizes invertíveis de ordem n, cujas entradas são números complexos. O determinante de uma matriz A será indicado por

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

ou, abreviadamente, por $det(A)$.

Uma matriz $A \in M_n(\mathbb{C})$ é dita ***diagonalizável*** se existe uma matriz invertível $C \in GL_n(\mathbb{C})$ tal que $D = C^{-1}AC$, em que D é uma matriz diagonal. Dizemos, nesse caso, que C ***diagonaliza*** A.

Seja $A = [a_{ij}]_{m\times n}$ uma matriz. Chamamos de ***matriz transposta*** de A a matriz $A^t = [a_{ji}]_{n\times m}$, na qual as linhas dessa são as colunas daquela e vice-versa. O ***conjugado*** de um número complexo $z = x + yi$ é $\bar{z} = x - yi$, e seu ***valor absoluto*** é $|z| = \sqrt{x^2 + y^2}$, em que $x, y \in \mathbb{R}$ e i é a unidade imaginária. A ***matriz conjugada*** $\bar{A} = [\bar{a}_{ij}]_{m\times n}$ de A é obtida pela substituição dos números complexos, das entradas da matriz A, pelos respectivos conjugados.

Uma matriz quadrada $A \in M_n(\mathbb{C})$ que verifica $A^{-1} = \bar{A}^t$ é intitulada ***unitária***. No final desta seção, demonstraremos o Teorema da Decomposição de Schur (Teorema 1.2.5). Esse teorema afirma que toda matriz $M \in M_n(\mathbb{C})$ pode ser escrita como $M = QTQ^{-1}$, em que Q é uma matriz unitária e T é uma matriz triangular superior.

Exporemos, agora, alguns conceitos básicos de Álgebra Linear, pois esses permitirão um melhor entendimento dos resultados desta seção.

Um ***espaço vetorial*** V sobre o corpo $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ é um conjunto em que seus elementos (***vetores***) podem ser somados e multiplicados por elementos do corpo $\mathbb{K}$ (***escalares***). Essas operações devem gozar das propriedades, para adição $+$, a saber, fechamento ($\vec{u} + \vec{v} \in V$), associatividade ($(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$), comutatividade ($\vec{u} + \vec{v} = \vec{v} + \vec{u}$), elemento neutro (existe $\vec{0} \in V$ tal que $\vec{u} + \vec{0} = \vec{u}$), inverso aditivo (existe $-\vec{u} \in V$ tal que $\vec{u} + (-\vec{u}) = \vec{0}$); para multiplicação $\cdot$, fechamento ($a_1\vec{u} \in V$), associatividade ($a_1(a_2\vec{u}) = (a_1a_2)\vec{u}$), distributividade ($a_1(\vec{u} + \vec{v}) = a_1\vec{u} + a_1\vec{v}$ e $(a_1 + a_2)\vec{u} = a_1\vec{u} + a_2\vec{u}$), em que $\vec{u}$, $\vec{v}$, $\vec{w} \in V$ e a_1 e $a_2 \in \mathbb{K}$.

Seja W um subconjunto não-vazio de um espaço vetorial V. Dizemos que W é um ***subespaço vetorial*** de V quando W, com as operações de adição em V e de multiplicação de vetores de V por escalares, é um espaço vetorial.

Um subconjunto $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ de um espaço vetorial V é dito ***linearmente independente***, se a equação $a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_n\vec{v}_n = \vec{0}$, em que $a_1, \ldots, a_n$ são escalares, é verificada somente quando $a_1 = a_2 = \cdots = a_n = 0$.

Denomina-se de ***base*** um subconjunto linearmente independente $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ de um espaço vetorial V tal que todo vetor $\vec{u} \in V$ pode ser escrito como $\vec{u} = a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_n\vec{v}_n$. Nesse contexto, diz-se que n é a ***dimensão*** de V, já que cada base de V possui o mesmo número cardinal de elementos.

Seja $A \in M_n(\mathbb{C})$ uma matriz. Dizemos que um vetor não-nulo $\vec{v}$ em $\mathbb{C}^n$ é ***autovetor*** de A se $A\vec{v} = \lambda\vec{v}$, para algum escalar λ. O escalar λ é chamado de ***autovalor*** e o vetor $\vec{v}$, de ***autovetor associado a*** λ. Uma condição necessária e suficiente para que λ seja um autovalor de A é λ satisfazer a ***equação característica*** $det(\lambda I - A) = 0$. Nomeia-se de ***autoespaço*** de A associado a λ o espaço-solução do sistema homogêneo $(\lambda I - A)\vec{v} = \vec{0}$.

Ao desenvolver o determinante $det(\lambda I - A)$, obtemos um polinômio da forma:

$$P_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \ldots + c_1\lambda + c_0$$

denominado ***polinômio característico*** de A. Prova-se que o polinômio característico é mônico (seu coeficiente lider é igual a 1), com grau numericamente igual à ordem da matriz A.

Por exemplo, o polinômio característico da matriz

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

é

$$P_A(\lambda) = \lambda^3 - (a_{11}+a_{22}+a_{33})\lambda^2 + (a_{22}a_{33} - a_{23}a_{32} + a_{11}a_{33} - a_{13}a_{31} + a_{11}a_{22} - a_{12}a_{21})\lambda - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Mais geralmente:

Proposição 1.2.1 *Seja $A \in M_n(\mathbb{C})$. Então, seu polinômio característico é*

$$P_A = \lambda^n - S_1\lambda^{n-1} + S_2\lambda^{n-2} - \ldots + (-1)^n S_n,$$

em que S_k são polinômios nas entradas da matriz.

Proposição 1.2.2 *Uma matriz $A \in M_n(\mathbb{C})$ é diagonalizável se, e somente se, possui n autovetores linearmente independentes.*

Demonstração: Suponhamos que A é diagonalizável, ou seja, existe uma matriz invertível $C \in GL_n(\mathbb{C})$ tal que $D = C^{-1}AC$ e D é uma matriz diagonal, equivalentemente,

$$AC = CD. \tag{1.1}$$

Denotando os vetores colunas de C por $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n$ e supondo que as entradas da diagonal de D sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$, os membros da equação (1.1) podem ser expressos por

$$\begin{aligned} AC &= A\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \ldots & \vec{c}_n \end{bmatrix} = \begin{bmatrix} A\vec{c}_1 & A\vec{c}_2 & \ldots & A\vec{c}_n \end{bmatrix} = \\ &= CD = \begin{bmatrix} \lambda_1\vec{c}_1 & \lambda_2\vec{c}_2 & \ldots & \lambda_n\vec{c}_n \end{bmatrix} \end{aligned}$$

A partir dessas últimas igualdades, temos

$$A\vec{c}_1 = \lambda_1\vec{c}_1,\ A\vec{c}_2 = \lambda_2\vec{c}_2,\ \ldots,\ A\vec{c}_n = \lambda_n\vec{c}_n. \tag{1.2}$$

Como a matriz C é invertível, temos que os seus vetores colunas $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n$ são linearmente independentes, logo não-nulos. Das igualdades de (1.2), conclui-se que os n vetores colunas $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n$ de C são os autovetores de A.

Reciprocamente, suponha que A possua n autovetores $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n$ linearmente independentes, com os autovalores associados $\lambda_1, \lambda_2, \ldots, \lambda_n$. Fazendo

$$C = \begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \end{bmatrix}$$

e denotando a matriz diagonal D, cujas entradas da sua diagonal são $\lambda_1, \lambda_2, \ldots, \lambda_n$, obtemos

$$\begin{aligned} AC &= A\begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \end{bmatrix} = \begin{bmatrix} A\vec{c}_1 & A\vec{c}_2 & \cdots & A\vec{c}_n \end{bmatrix} = \\ &= \begin{bmatrix} \lambda_1\vec{c}_1 & \lambda_2\vec{c}_2 & \cdots & \lambda_n\vec{c}_n \end{bmatrix} = CD. \end{aligned}$$

Como os vetores colunas de C são linearmente independentes, concluímos que a matriz C é invertível e, por essa última equação, que $D = C^{-1}AC$.

Logo, a matriz A é diagonalizável. ■

A demonstração desse teorema nos ensina como diagonalizar uma matriz $A = [a_{ij}]_{n\times n}$. Os escalares que formam a diagonal principal da matriz diagonal $D = [d_{ij}]_{n\times n}$ são os autovalores da matriz a ser diagonalizada $A = [a_{ij}]_{n\times n}$; as colunas da matriz $C = [c_{ij}]_{n\times n}$ que diagonaliza A (*i.e.*, $D = C^{-1}AC$) são os autovetores das bases dos autoespaços de $A = [a_{ij}]_{n\times n}$. Mais do que isso, ela diz que uma matriz é diagonalizável se, e somente se, a multiplicidade de cada um de seus autovalores como raiz do polinômio característico coincide com a dimensão do autoespaço correspondente (*i.e.*, as multiplicidades algébrica e geométrica de cada autovalor coincidem).

Por exemplo, considera-se a matriz

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

Após calcular o determinante

$$\begin{vmatrix} \lambda & 0 & -2 \\ 1 & \lambda - 2 & 1 \\ 1 & 0 & \lambda - 3 \end{vmatrix} = 0,$$

fatorar o polinômio característico e igualá-lo a zero, tem-se a equação característica

$$(\lambda - 2)^2(\lambda - 1) = 0.$$

Portanto, obtêm-se os autovalores $\lambda = 2$ e $\lambda = 1$ da matriz A. Logo, a matriz A possui dois autoespaços.

Por definição,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

é um autovetor de A se, e somente se, $\vec{x}$ é um solução não trivial de $(\lambda I - A)\vec{x} = \vec{0}$ ou, em termos de matrizes,

$$\begin{bmatrix} \lambda & 0 & -2 \\ 1 & \lambda - 2 & 1 \\ 1 & 0 & \lambda - 3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \vec{0}. \tag{1.3}$$

Resolvendo o sistema (1.3) para o caso $\lambda = 2$, resulta $x_1 = -a$, $x_2 = b$ e $x_3 = a$. Os autovetores associados a $\lambda = 2$ são da forma

$$\begin{bmatrix} -a \\ b \\ a \end{bmatrix} = \begin{bmatrix} -a \\ 0 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix} = a \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Observa-se que os autovetores

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ e } \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

são linearmente independentes e, então, formam uma base para o autoespaço associado a $\lambda = 2$.

Resolvendo o sistema (1.3) para o caso $\lambda = 1$, chega-se a $x_1 = -2a$, $x_2 = a$ e $x_3 = a$. O autovetor associado a $\lambda = 1$ é da forma

$$\begin{bmatrix} -2a \\ a \\ a \end{bmatrix} = a \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}.$$

Donde, o autovetor

$$\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

é a base do autoespaço associado a $\lambda = 1$.

Assim, as colunas da matriz C, que diagonaliza a matriz A, são os autovetores das bases dos autoespaços da matriz A

$$C = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Verifica-se que

$$C^{-1}AC = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = D.$$

A diagonal da matriz diagonal D é formada pelos autovalores da matriz A. Não importa a ordem das colunas da matriz C, essa afeta somente a ordem dos autovalores da matriz diagonal D.

Os procedimentos realizados acima foram descritos, em detalhes, com o intuito de que o leitor pudesse compreender o procedimento de diagonalização de uma matriz. O Lema 1.2.3 a seguir exibe uma condição suficiente para uma matriz ser diagonalizável.

Lema 1.2.3 *Seja $M \in M_n(\mathbb{C})$ uma matriz. Se o seu polinômio característico P_M possui somente raízes simples, então M é diagonalizável.*

Demonstração: Pela hipótese, pelo Teorema Fundamental da Álgebra (Teorema 1.1.4) e pela Divisão Euclidiana (Proposição 1.1.1), o polinômio característico P_M possui n raízes distintas $\lambda_1, \ldots, \lambda_n$.

À luz da Proposição 1.2.2, basta mostrar que se $M\vec{v}_1 = \lambda_1\vec{v}_1, \ldots, M\vec{v}_n = \lambda_n\vec{v}_n$ com $\vec{v}_1, \ldots, \vec{v}_n$ não-nulos e $\lambda_1, \ldots, \lambda_n$ distintos, então $\{\vec{v}_1, \ldots, \vec{v}_n\}$ é linearmente independente. Provamos isso por indução sobre n. Se $n = 1$, o resultado é evidente. Reordenando os vetores, se necessário, podemos supor que $\{\vec{v}_1, \ldots, \vec{v}_{n-1}\}$ são linearmente independentes e consideremos o caso de n vetores.

Multiplicando por M ambos os membros da equação

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n = \vec{0}, \tag{1.4}$$

obtemos $a_1M\vec{v}_1 + a_2M\vec{v}_2 + \ldots + a_nM\vec{v}_n = \vec{0}$. Ora, $M\vec{v}_i = \lambda_i\vec{v}_i$. Assim,

$$a_1\lambda_1\vec{v}_1 + a_2\lambda_2\vec{v}_2 + \ldots + a_n\lambda_n\vec{v}_n = \vec{0}.$$

Por outro lado, multiplicando por λ_n os membros de (1.4), resulta

$$a_1\lambda_n\vec{v}_1 + a_2\lambda_n\vec{v}_2 + \ldots + a_n\lambda_n\vec{v}_n = \vec{0}.$$

Subtraindo essas duas últimas equações, concluímos que

$$a_1(\lambda_1 - \lambda_n)\vec{v}_1 + a_2(\lambda_2 - \lambda_n)\vec{v}_2 + \ldots + a_{n-1}(\lambda_{n-1} - \lambda_n)\vec{v}_{n-1} = \vec{0}.$$

Do fato de $\lambda_i - \lambda_n \neq 0$ para todo $1 \leq i \leq n-1$, a hipótese de indução assegura que $a_i = 0$ para $1 \leq i \leq n-1$. Substituindo em (1.4), obtém-se $a_n\vec{v}_n = \vec{0}$. Logo, $a_n = 0$, porque $\vec{v}_n \neq \vec{0}$. Portanto, os autovetores $\vec{v}_1, \ldots, \vec{v}_n$ são linearmente independentes. ■

Neste momento, serão introduzidos os conceitos de produto interno e ortonormalidade em $\mathbb{C}^n$. Esses conceitos serão empregados em apenas na Proposição 1.2.4 e no Teorema da Decomposição de Schur a seguir.

Sejam os vetores $\vec{u} = (u_1, u_2, \ldots, u_n)$ e $\vec{v} = (v_1, v_2, \ldots, v_n)$ de $\mathbb{C}^n$. Definimos o ***produto interno (canônico)*** em $\mathbb{C}^n$ por $\langle \vec{u}, \vec{v} \rangle = \sum_{j=1}^{n} u_j \bar{v}_j$, em que os $\bar{v}_j = x_j - iy_j$ são os conjugados dos números complexos $v_j = x_j + iy_j$ e i é a unidade imaginária. Um conjunto de vetores $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ é dito ***ortonormal*** quando o produto interno $\langle \vec{v}_i, \vec{v}_j \rangle$ de dois quaisquer de seus vetores for igual a 1, caso $i = j$, ou a 0, se $i \neq j$.

Proposição 1.2.4 (Processo de Ortonormalização de Gram-Schmidt) *Dada uma base $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ de um subespaço W de $\mathbb{C}^n$. Então, existe uma base $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_k\}$ de W que é ortonormal e tal que o subespaço gerado por $\{\vec{V}_1, \vec{V}_2, \cdots, \vec{V}_k\}$ é igual ao subespaço gerado por $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$.*

Demonstração: Procederemos por indução. Tomando inicialmente $\vec{v}_1 = \vec{w}_1$ e $\vec{V}_1 = \dfrac{\vec{v}_1}{\sqrt{\langle \vec{v}_1, \vec{v}_1 \rangle}}$. É imediato que $\vec{w}_1$ e $\vec{V}_1$ geram um mesmo subespaço. Suponha que os subespaços gerados por $\{\vec{w}_1, \ldots, \vec{w}_{k-1}\}$ e por $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_{k-1}\}$ coincidem e $\langle \vec{V}_i, \vec{V}_j \rangle = 0$, para cada i e $1 \leq i < j < k$. Definimos

$$\vec{v}_k = \vec{w}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{\langle \vec{v}_j, \vec{v}_j \rangle} \vec{v}_j.$$

Daí, $\langle v_k, V_j \rangle = 0$, para$1 \leq j \leq k-1$. Além disso, $\vec{v}_k \neq 0$, porque o vetor $\vec{v}_k$ não pertence ao subespaço gerado por $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_{k-1}\}$. Fazendo $\vec{V}_k = \dfrac{\vec{v}_k}{\sqrt{\langle \vec{v}_k, \vec{v}_k \rangle}}$ e observando que $\vec{v}_k$ pertence ao subespaço gerado por $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_{k-1}, \vec{w}_k\}$, o qual é o mesmo gerado por $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_{k-1}, \vec{V}_k\}$, concluímos que $\{\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_k\}$ é uma base ortonormal de W e gera o mesmo subespaço que $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$. ■

Teorema 1.2.5 (Teorema da Decomposição de Schur) *Toda matriz $M \in M_n(\mathbb{C})$ pode ser escrita como $M = QTQ^{-1}$, em que Q é uma matriz unitária e T é uma matriz triangular superior.*

Demonstração: O resultado é evidente se $n = 1$. Suporemos que o resultado seja verdadeiro para matrizes $(n-1)$-por-$(n-1)$ e provaremos que ele é verdadeiro para matrizes n por n.

Pelo Teorema Fundamental da Álgebra (Teorema 1.1.4) a matriz M possui um autovalor $\lambda_1 \in \mathbb{C}$. Isto significa que existem autovetores associados a λ_1. Seja $\vec{V}_1$ um autovetor de norma igual a 1 associado a λ_1. Sejam $\vec{V}_2, \cdots, \vec{V}_n$ vetores tais que

$\{\vec{V}_1, \ldots, \vec{V}_n\}$ é uma base ortonormal de $\mathbb{C}^n$ (isto pode ser conseguido aplicando-se o processo de ortonormalização de Gram-Schmidt a uma base de $\mathbb{C}^n$ que contenha $\vec{V}_1$, conforme a Proposição 1.2.4). Considere $Q_1 = \begin{bmatrix} \vec{V}_1 & \cdots & \vec{V}_n \end{bmatrix}$. Como $M\vec{V}_1 = \lambda_1 \vec{V}_1$ e $M\vec{V}_2, \ldots, M\vec{V}_n$ são combinações lineares de $\vec{V}_1, \ldots, \vec{V}_n$, obtemos

$$MQ_1 = \begin{bmatrix} M\vec{V}_1 & \cdots & M\vec{V}_n \end{bmatrix} = \begin{bmatrix} \vec{V}_1 & \cdots & \vec{V}_n \end{bmatrix} P = Q_1 P, \tag{1.5}$$

em que

$$P = \left[\begin{array}{c|c} \lambda_1 & P_2 \\ \hline 0 & \\ \vdots & N \\ 0 & \end{array}\right].$$

Em virtude de estarmos supondo o resultado verdadeiro para matrizes $(n-1)$ por $(n-1)$, existe matriz unitária Q_2 e uma matriz triangular superior T_2, ambas $(n-1)$ por $(n-1)$, tais que $N = Q_2 T_2 \bar{Q}_2^{\,t}$.

Considere-se

$$Q_3 = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right].$$

Seja $Q = Q_1 Q_3$. A matriz Q é unitária e pela equação (1.5), temos:

$$MQ = (MQ_1)Q_3 = Q_1 P Q_3 = Q_1 \left[\begin{array}{c|c} \lambda_1 & P_2 Q_2 \\ \hline 0 & \\ \vdots & NQ_2 \\ 0 & \end{array}\right].$$

Ora, $NQ_2 = Q_2 T_2$ e, pois,

$$MQ = Q_1 \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q_2 & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|c} \lambda_1 & P_2 Q_2 \\ \hline 0 & \\ \vdots & NQ_2 \\ 0 & \end{array}\right] = Q_1 Q_3 \left[\begin{array}{c|c} \lambda_1 & P_2 Q_2 \\ \hline 0 & \\ \vdots & NQ_2 \\ 0 & \end{array}\right] = QT, \tag{1.6}$$

em que

$$T = \left[\begin{array}{c|c} \lambda_1 & P_2 Q_2 \\ \hline 0 & \\ \vdots & NQ_2 \\ 0 & \end{array}\right].$$

Multiplicando-se os membros das igualdades (1.6) por $\bar{Q}^t$, concluímos que $M = QT\bar{Q}^t$. ■

2 Topologia

"A Topologia surgiu no século vinte como um tema que unifica quase toda a matemática, um tanto como a filosofia procura coordenar todo o conhecimento. Por causa de seu primitivismo, a topologia está na base de uma parte muito grande da matemática, conferindo-lhe uma inesperada coesão".

Carl B. Boyer

Neste capítulo, o problema das matrizes diagonalizáveis é estudado sob o enfoque da Topologia. Para atingir esse propósito, são desenvolvidas inicialmente as noções de espaço topológico, funções contínuas e densidade, na Seção 2.1. Posteriormente, na Seção 2.2, são expostas duas demonstrações para a densidade das matrizes diagonalizáveis, Teoremas 2.2.4 e 2.2.6.

2.1 Espaços Topológicos e Funções Contínuas

Nesta seção, são introduzidos os conceitos de espaço topológico, de função contínua e de densidade. A topologia usual e a métrica do xadrez em $\mathbb{R}^n$ são também estudadas. Essa métrica foi escolhida; em virtude de as "bolas", normalmente estudadas nos cursos de Cálculo, possuirem a forma de blocos retangulares. Utilizou-se, como referência para essa seção, a obra (MUNKRES, 2000).

Uma ***topologia*** num conjunto X é uma família τ de subconjuntos de X, denominados de ***abertos*** (concernente à topologia τ), que satisfaz as condições:

1. a união de uma família arbitrária de abertos é um aberto;

2. a interseção de uma família finita de abertos é um aberto.

Em particular, os conjuntos $\emptyset$ e X são membros de τ. O par (X, τ) é chamado de ***espaço topológico***.

Definiremos agora a topologia usual no espaço euclidiano $\mathbb{R}^n$.

Sejam $S_1, \ldots, S_n$ conjuntos e n um número natural. Definimos o ***produto cartesiano*** $\prod_{i=1}^{n} S_i$ como o conjunto de todas ênuplas ordenadas $(x_1, \ldots, x_n)$ tais que

$x_i \in S_i$, *i.e.*,

$$\prod_{i=1}^{n} S_i = S_1 \times \ldots \times S_n = \{(x_1, \ldots, x_n) : x_i \in S_i\}.$$

Um ***bloco retangular n-dimensional*** é um subconjunto B de $\mathbb{R}^n$ $(n \geq 1)$ que é ou vazio, ou da forma:

$$B = \prod_{i=1}^{n}]a_i, b_i[\;\; = \;\;]a_1, b_1[\times \ldots \times]a_n, b_n[,$$

em que a_i, $b_i \in \mathbb{R}$, $a_i \leq b_i$ para $i = 1, 2, \ldots, n$.

Consideremos $X = \mathbb{R}^n$. Definimos a seguinte topologia τ: um subconjunto A de $\mathbb{R}^n$ será um aberto se for vazio ou se, para todo ponto $x = (x_1, \ldots, x_n)$ de A, existir um bloco retangular B n-dimensional que possua o ponto x e esteja contido em A. Essa topologia τ é denominada **topologia usual ou euclidiana** em $\mathbb{R}^n$.

Verifiquemos que a família τ satisfaz as condições de topologia.

É imediato que o conjunto vazio e o próprio $\mathbb{R}^n$ pertencem à família τ.

1. Com efeito, tomando um ponto $a \in A_{j'}$ e $A_{j'} \in \tau$, para algum índice j' de um conjunto de índices arbitrário J, existe um bloco retangular n-dimensional $B_{j'}$, com $a \in B_{j'} \subset A_{j'} \subset \bigcup_{j \in J} A_j$. Assim, $\bigcup_{j \in J} A_j \in \tau$.

2. Dado um ponto $a \in (A_1 \cap \ldots \cap A_k)$, em que k é um inteiro positivo. Como $A_1, \ldots, A_k \in \tau$, existem blocos retangulares n-dimensional $B_1, \ldots, B_k$ tais que $a \in B_1 \subset A_1, \ldots, a \in B_k \subset A_k$. Tomemos o bloco n-dimensional $B = B_1 \cap \ldots \cap B_k$, porque a interseção finita de blocos não disjuntos é sempre um bloco. Assim, temos $a \in B \subset (A_1 \cap \ldots \cap A_k)$. Logo, $(A_1 \cap \ldots \cap A_k) \in \tau$.

Acontece que a topologia usual em $\mathbb{R}^n$ pode ser descrita em termos da norma do máximo. De fato, seja $x = (x_1, \ldots, x_n)$ um ponto de $\mathbb{R}^n$, chamamos ***norma do máximo*** a função $\| \cdot \| : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

$$\|x\| = \max\{|x_1|, \ldots, |x_n|\}.$$

O bloco n-dimensional $B \;\; =]y_1 - r, y_1 + r[\times \ldots \times]y_n - r, y_n - r[$, em que $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ e r é um número real positivo, pode ser expresso por $B = \{x \in \mathbb{R}^n : \|y - x\| < r\}$.

O próximo exemplo será utilizado na demonstração do Lema 3.2.2.

Exemplo 2.1.1 *Seja x um ponto de $\mathbb{R}^n$. O conjunto $\mathbb{R}^n - \{x\}$ é um aberto com respeito a topologia usual em $\mathbb{R}^n$. Em particular, $\mathbb{R}^n - \{0\}$ é um aberto.*

Provaremos por indução sobre n. Para $n = 1$, dado um ponto $y \in \mathbb{R} - \{x\}$, toma-se o intervalo $]a, b[\subset \mathbb{R}$ tal que $y \in \;]a, b[$ e $a < b \leq x$ (da mesma forma para $x < a \leq b$). Logo, por construção, $y \in \;]a, b[\subset \mathbb{R} - \{x\}$ e, portanto, $\mathbb{R} - \{x\}$ é aberto.

Supanhamos, por hipótese de indução, que seja válido para $n-1$ e demonstremos que o conjunto $\mathbb{R}^n - \{x\}$ é um aberto, em que $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Dado um ponto $y = (y_1, \ldots, y_{n-1}, y_n) \in \mathbb{R}^n - \{x\}$. Pela hipótese de indução, existe um bloco retangular $n-1$-dimensional $B' =]a_1, b_1[\times \ldots \times]a_{n-1}, b_{n-1}[$ que contém o ponto $y' = (y_1, \ldots, y_{n-1})$ e $x_1 \notin]a_1, b_1[, \ldots, x_{n-1} \notin]a_{n-1}, b_{n-1}[$. Por um argumento análogo ao caso $n = 1$, existe um intervalo $]a_n, b_n[$ tal que $y_n \in]a_n, b_n[\subset \mathbb{R} - \{x_n\}$. Portanto, o bloco n-dimensional $B =]a_1, b_1[\times \ldots \times]a_n b_n[$ contém o ponto $y = (y_1, \ldots, y_{n-1}, y_n)$ e, por construção, está contido em $\mathbb{R}^n - \{x\}$. Logo, pelo Princípio da Indução Finita, o conjunto $\mathbb{R}^n - \{x\}$ é aberto com respeito a topologia usual em $\mathbb{R}^n$, para todo inteiro positivo n.

Considera-se uma função $f : X \longrightarrow Y$. A ***imagem inversa*** de um subconjunto S de Y é definida por

$$f^{-1}(S) = \{x \in X : f(x) \in S\}.$$

Sejam (X, τ_X) e (Y, τ_Y) espaços topológicos. Uma função $f : (X, \tau_X) \longrightarrow (Y, \tau_Y)$ é denominada ***contínua***, quando a imagem inversa de todo aberto de Y, com respeito a topologia τ_Y, for um aberto de X, concernente a topologia τ_X. Essa definição abstrai a ideia intuitiva de que para cada ponto x_0 em X fixado: a função $f(x)$ pode assumir valores tão próximos de $f(x_0)$ quanto se queira, desde que o valor de x seja escolhido suficientemente próximo de x_0. No caso em que $X = \mathbb{R}^n$ e $Y = \mathbb{R}$, isso se traduz em: para cada número real $\varepsilon > 0$, existe um correspondente $\delta > 0$ tal que, para qualquer x em X satisfazendo $0 < \|x - x_0\| < \delta$ vale que $|f(x) - f(x_0)| < \varepsilon$. Para mais detalhes, o leitor poderá consultar (LANG, 2012, p. 25) e (ZAPATA, 2014, aula 5).

Proposição 2.1.2 *Toda função polinomial $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é contínua, com respeito a topologia usual de $\mathbb{R}^n$ e, respectivamente, de $\mathbb{R}$.*

Demonstração: A ideia principal da demonstração é mostrar que a função polinomial $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, digamos

$$f(x_1, \cdots, x_n) = \sum_{\substack{0 \le i_1 \le r_1 \\ \cdots \\ 0 \le i_n \le r_n}} a_{i_1 \cdots i_n} x_1^{i_1} \cdots x_n^{i_n},$$

em que $a_{i_1 \ldots i_n} \in \mathbb{R}$ e os $r_1, \ldots, r_n$ são inteiros não negativos, é contínua em quatro passos. O primeiro passo é demonstrar que as funções polinomiais mais simples da forma $f(x_1, \ldots, x_n) = b$ (funções constantes), com $b \in \mathbb{R}$, são contínuas. O segundo passo é provar a Proposição 2.1.2 para funções polinomiais da forma $f_i(x_1, \ldots, x_n) = x_i$, para $i = 1, \cdots, n$. O terceiro passo é demonstrar que, se $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ e $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ são funções contínuas, então a soma $f(x) + g(x)$ é uma função contínua. O quarto passo é mostrar que, se $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ e $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ são funções contínuas, então o produto $f(x)g(x)$ é também uma função contínua. Dos passos de números dois e quatro, teríamos que as funções da forma $x_1^{i_1}\cdots x_n^{i_n}$ seriam contínuas, para $i_1,\ldots,i_n$ inteiros não negativos. Dos passos de números um, três e quatro, concluiríamos que a função polinomial $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ é contínua, porque é uma soma de funções contínuas da forma $a_{i_1\cdots i_n}x_1^{i_1}\cdots x_n^{i_n}$, em que $a_{i_1\ldots i_n}\in\mathbb{R}$ e $i_1,\ldots,i_n$ são inteiros não negativos.

Primeiro Passo: *A função constante $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ dada por $f(x_1,\ldots,x_n)=b$, com $b\in\mathbb{R}$, é contínua em cada ponto $a=(a_1,\ldots,a_n)\in\mathbb{R}^n$.*

A função constante é contínua, porque a desigualdade $|f(x)-f(a)|=|b-b|=0<\varepsilon$ é válida, para dois pontos quaisquer x e a de $\mathbb{R}^n$ e cada número real $\varepsilon>0$.

Segundo Passo: *As funções $f_i:\mathbb{R}^n\longrightarrow\mathbb{R}$ dadas por $f_i(x_1,\ldots,x_n)=x_i$ são contínuas em cada ponto $a=(a_1,\ldots,a_n)\in\mathbb{R}^n$, para $i=1,\ldots,n$.*

Com efeito, notemos que $|f_i(x)-f_i(a)|\leq\|x_i-a_i\|$, para $i=1,\ldots,n$. Daí, para cada número real $\varepsilon>0$, existe um correspondente $0<\delta\leq\varepsilon$ tal que, para todo x em $\mathbb{R}^n$ satisfazendo $0<\|x-a\|<\delta$ vale que $|f_i(x)-f_i(a)|\leq\|x_i-a_i\|<\varepsilon$.

Terceiro Passo: *Provemos que, se $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ e $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ são funções contínuas, então a função $f+g:\mathbb{R}^n\longrightarrow\mathbb{R}$ definida por $(f+g)(x)=f(x)+g(x)$ é contínua em todo ponto $a=(a_1,\ldots,a_n)\in\mathbb{R}^n$.*

De fato, como f e g são funções contínuas, existem $\delta_1>0$ e $\delta_2>0$ tais que $|f(x)-f(a)|<\frac{\varepsilon}{2}$, para cada x em $\mathbb{R}^n$ satisfazendo $0<\|x-a\|<\delta_1$, e $|g(x)-g(a)|<\frac{\varepsilon}{2}$, para cada x em $\mathbb{R}^n$ satisfazendo $0<\|x-a\|<\delta_2$. Portanto, se $\delta=min\{\delta_1,\delta_2\}$, então $\delta>0$ e a concomitância das condições x em $\mathbb{R}^n$ e $0<\|x-a\|<\delta$ implica

$$|(f+g)(x)-(f+g)(a))|=|(f(x)-f(a))+(g(x)-g(a))|$$

$$\leq|f(x)-f(a)|+|g(x)-g(a)|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Quarto Passo: *Demonstremos que, se $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ e $g:\mathbb{R}^n\longrightarrow\mathbb{R}$ são funções contínuas, então a função $fg:\mathbb{R}^n\longrightarrow\mathbb{R}$ definida por $fg(x)=f(x)g(x)$, é contínua em todo ponto $a=(a_1,\ldots,a_n)\in\mathbb{R}^n$.*

Realmente, da desigualdade triangular, têm-se

$$|f(x)g(x)-f(a)g(a)|=|(f(x)-f(a))\cdot(g(x)-g(a))+f(a)\cdot(g(x)-g(a))+g(a)\cdot(f(x)-f(a))|$$

$$\leq|f(x)-f(a)|\cdot|g(x)-g(a)|+|f(a)|\cdot|g(x)-g(a)|+|g(a)|\cdot|f(x)-f(a)|.$$

Logo, para cada número real $\varepsilon>0$, a fim de que seja $|f(x)g(x)-f(a)g(a)|<\varepsilon$ para todo x em $\mathbb{R}^n$ próximo ao ponto $a=(a_1,\cdots a_n)\in\mathbb{R}^n$, é suficiente que tenhamos cada uma das parcelas $|f(x)-f(a)|\cdot|g(x)-g(a)|$, $|f(a)|\cdot|g(x)-g(a)|$, $|g(a)|\cdot|f(x)-f(a)|$ menor que $\frac{\varepsilon}{3}$. Para tanto, basta que tenhamos $|f(x)-f(a)|<\sqrt{\frac{\varepsilon}{3}}$, $|g(x)-g(a)|<\sqrt{\frac{\varepsilon}{3}}$, $|f(x)-f(a)|<\frac{\varepsilon}{3|f(a)|+1}$ e $|g(x)-g(a)|<\frac{\varepsilon}{3|g(a)|+1}$. Em síntese, é suficiente que se tenha $|f(x)-f(a)|<min\{\sqrt{\frac{\varepsilon}{3}},\frac{\varepsilon}{3|f(a)|+1}\}$ e $|g(x)-g(a)|<min\{\sqrt{\frac{\varepsilon}{3}},\frac{\varepsilon}{3|g(a)|+1}\}$. Fazendo $\varepsilon_1=min\{\sqrt{\frac{\varepsilon}{3}},\frac{\varepsilon}{3|f(a)|+1}\}$ e $\varepsilon_2=min\{\sqrt{\frac{\varepsilon}{3}},\frac{\varepsilon}{3|g(a)|+1}\}$,

temos $\varepsilon_1 > 0$ e $\varepsilon_2 > 0$, e, pela continuidade das funções f e g provada no primeiro parágrafo, existem $\delta_1 > 0$ e $\delta_2 > 0$ tais que $|f(x) - f(a)| < \varepsilon_1$, para cada x em $\mathbb{R}^n$ satisfazendo $0 < \|x - a\| < \delta_1$, e $|g(x) - g(a)| < \varepsilon_2$, para cada x em $\mathbb{R}^n$ satisfazendo $0 < \|x - a\| < \delta_2$. Portanto, se $\delta = min\{\delta_1, \delta_2\}$, então $\delta > 0$ e a concomitância das condições x em $\mathbb{R}^n$ e $0 < \|x - a\| < \delta$ implica, ao mesmo tempo, $|f(x) - f(a)| < \varepsilon_1$ e $|g(x) - g(a)| < \varepsilon_2$, como necessário. ■

Seja (X, τ) um espaço topológico. Dizemos que um conjunto D é ***denso*** em X se todo conjunto aberto tiver algum elemento de D. Em particular, na topologia usual em $\mathbb{R}^n$, um conjunto D será ***denso*** se todo bloco retangular n-dimensional possuir algum elemento de D.

O conjunto $\mathbb{Q}$ dos números racionais é denso em $\mathbb{R}$, com a topologia usual. Ora, todo intervalo aberto da reta contém um número racional e os blocos retangulares 1-dimensional em $\mathbb{R}$ são exatamente os intervalos abertos.

2.2 O Conjunto das Matrizes Diagonalizáveis é Denso

Nesta seção, apresentaremos duas demonstrações da densidade do conjunto das matrizes diagonalizáveis. A primeira se assenta numa perspectiva matricial, na qual são empregados o Lema 2.2.2 e o Teorema 1.2.5 (Teorema da Decomposição de Schur). A segunda se fundamenta numa perspectiva algébrica, mais especificamente, no estudo de polinômios. Essas diferentes abordagens ilustram que a densidade das matrizes diagonalizáveis não depende das métricas adotadas, mas tão somente da topologia.

Podemos facilmente introduzir uma topologia τ_1 no conjunto das matrizes $M_n(\mathbb{C})$, de modo que o espaço topológico $(M_n(\mathbb{C}), \tau_1)$ terá a mesma estrutura de $\mathbb{R}^{2n^2}$ com a topologia usual.

Observação 2.2.1 *Consideremos a bijeção $\phi_1 : M_n(\mathbb{C}) \longrightarrow \mathbb{R}^{2n^2}$ que associa a cada matriz*

$$\begin{bmatrix} x_{11} + iy_{11} & \dots & x_{1n} + iy_{1n} \\ x_{21} + iy_{21} & \dots & x_{2n} + iy_{2n} \\ \vdots & \vdots & \vdots \\ x_{n1} + iy_{n1} & \dots & x_{nn} + iy_{nn} \end{bmatrix}_{n \times n}$$

o elemento

$$(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn}, y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn}).$$

Um subconjunto de matrizes será aberto se, e somente se, a sua imagem pela função ϕ_1 o for. Em símbolos,

$$\tau_1 = \{A \subset M_n(\mathbb{C}) : \phi_1(A) \in \tau\}.$$

Analogamente, insere-se em $\mathbb{C}^n$ a topologia $\tau_2 = \{A \subset \mathbb{C}^n : \phi_2(A) \in \tau\}$, em que a função

$$\phi_2 : \mathbb{C}^n \longrightarrow \mathbb{R}^{2n}$$

é dada por

$$(x_1 + iy_1, x_2 + iy_2, \ldots, x_n + iy_n) \mapsto (x_1, y_1, x_2, y_2, \ldots, x_n, y_n),$$

em que i é a unidade imaginária. Um subconjunto de $\mathbb{C}^n$ será aberto se, e somente se, a sua imagem pela função ϕ_2 o for.

Nesta seção, serão usadas as topologias τ_1 e τ_2, respectivamente, em $M_n(\mathbb{C})$ e $\mathbb{C}^n$.

Considere o espaço vetorial $M_n(\mathbb{C})$ das matrizes de ordem n sobre o corpo $\mathbb{C}$, com as operações usuais de adição de matrizes e produto por escalar. Seja $\|\cdot\| : M_n(\mathbb{C}) \longrightarrow [0, +\infty)$ a função definida por

$$\|A\| = \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{ij}|,$$

para cada matriz $A = [a_{ij}] \in M_n(\mathbb{C})$. Essa função é denominada ***norma do máximo*** e possui a seguinte propriedade.

Lema 2.2.2 *Para quaisquer matrizes $A, B \in M_n(\mathbb{C})$, vale:*

$$\|AB\| \leq n \cdot \|A\| \cdot \|B\|.$$

Em particular, $\|P^{-1}AP - P^{-1}A_0P\| \leq n^2 \cdot \|P\|^2 \cdot \|A - A_0\|$, em que P é uma matriz unitária ($P^{-1} = \bar{P}^t$) e $A, A_0 \in M_n(\mathbb{C})$.

Demonstração: Pelas definições de multiplicação de matrizes e da norma do máximo, e da desigualdade triangular do módulo de um número complexo, temos

$$\begin{aligned} \|AB\| &= \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \cdots + a_{in} \cdot b_{nj}| \\ &\leq \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{i1} \cdot b_{1j}| + \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{i2} \cdot b_{2j}| + \cdots + \max_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} |a_{in} \cdot b_{nj}| \\ &\leq n \cdot \|A\| \cdot \|B\|, \end{aligned}$$

em que na última desigualdade usamos que $\|A\| \geq |a_{ij}|$ e $\|B\| \geq |b_{ij}|$ para todo $1 \leq i,\ j \leq n$.

Passemos a provar o caso particular da propriedade da norma do máximo. Da propriedade distributiva à direita e à esquerda das matrizes, temos que

$$\|P^{-1}AP - P^{-1}A_0P\| = \|P^{-1}(A - A_0)P\|.$$

Aplicando duas vezes a propriedade da norma do máximo, no segundo membro dessa última igualdade, obtém-se

$$\|P^{-1}AP - P^{-1}A_0P\| \leq n^2 \cdot \|P^{-1}\| \cdot \|P\| \cdot \|A - A_0\|.$$

Como $\|\bar{P}^t\| = \|P\|$ e a matriz P é unitária, ou seja, $P^{-1} = \bar{P}^t$, concluímos que

$$\|P^{-1}AP - P^{-1}A_0P\| \leq n^2 \cdot \|P\|^2 \cdot \|A - A_0\|.$$

O que termina a demonstração do lema. ∎

Observação 2.2.3 *Dados quaisquer matriz $M = [m_{ij}] \in M_n(\mathbb{C})$ e número real $\varepsilon > 0$, desejamos exibir, na demonstração do Teorema 2.2.4, uma matriz $N = [n_{ij}] \in M_n(\mathbb{C})$ diagonalizável tal que $max_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}|m_{ij} - n_{ij}| < \varepsilon$. Para esse propósito, será introduzida a topologia induzida pela norma do máximo em $M_n(\mathbb{C})$.*

*Sejam $(M_n(\mathbb{C}), \|\cdot\|)$ o espaço das matrizes, munido com a norma do máximo, uma matriz M e um número real $r > 0$. Define-se uma **bola aberta** de centro em M e raio r como o conjunto $B(M, r)$ de todas as matrizes N de $M_n(\mathbb{C})$ tais que $\|N - M\| < r$. Ou seja,*

$$B(M, r) = \{N \in M_n(\mathbb{C}) : \|N - M\| < r\}.$$

*Diz-se que um subconjunto S de $M_n(\mathbb{C})$ é **aberto** se, para toda matriz $M \in S$, existir uma bola $B(M, r)$ contida em S. A topologia assim definida em $M_n(\mathbb{C})$ é denominada de **topologia induzida pela norma do máximo**.*

Nesse contexto, demonstremos que essa topologia induzida pela norma do máximo em $M_n(\mathbb{C})$ coincide com a topologia usual de $\mathbb{R}^{2n^2}$ em $M_n(\mathbb{C})$, (Observação 2.2.1).

Sejam $m_j = a_j + ib_j$ as entradas da matriz $M \in M_n(\mathbb{C})$ e $n_j = c_j + id_j$ as entradas da matriz $N \in M_n(\mathbb{C})$, em que a_j, b_j, c_j e d_j são números reais, i a unidade imaginária e $j = 1, 2, \ldots, n^2$. Das desigualdades $|x_1| = \sqrt{x_1^2} \leq \sqrt{x_1^2 + \ldots + x_k^2} \leq \sqrt{k}\, max\{|x_1|, \ldots, |x_k|\}$, válida para todo $x_1, \ldots, x_k$ números reais e para todo inteiro positivo k, temos

$$d(M, N) = max\,\{|a_1 - c_1|, \ldots, |a_{n^2} - c_{n^2}|, |b_1 - d_1|, \ldots, |b_{n^2} - d_{n^2}|\} \leq$$
$$\leq max\left\{\sqrt{(a_1 - c_1)^2 + (b_1 - d_1)^2}, \ldots, \sqrt{(a_{n^2} - c_{n^2})^2 + (b_{n^2} - d_{n^2})^2}\right\} = \|M - N\| = d'(M, N)$$
$$\leq n\ max\,\{|a_1 - c_1|, \ldots, |a_{n^2} - c_{n^2}|, |b_1 - d_1|, \ldots, |b_{n^2} - d_{n^2}|\} = n\ d(M, N),$$

em que d (com um abuso de notação, que é irrelevante neste contexto) é a métrica de $\mathbb{R}^{2n^2}$ induzida em $M_n(\mathbb{C})$ (por meio da bijeção da Observação 2.2.1) e $d'(M, N) = \|M - N\|$ é a métrica induzida pela norma do máximo em $M_n(\mathbb{C})$. Assim, valem as sequintes desigualdades

$$d(M, N) \leq d'(M, N) \leq n\ d(M, N). \tag{2.1}$$

Das desigualdades (2.1), concluímos que toda bola aberta $B(M, r_1)$ segundo a métrica $d(M, N)$ contém uma bola aberta de mesmo centro $B'(M, \frac{r_1}{2})$ segundo a métrica $d'(M, N)$. Reciprocamente, cada bola aberta $B'(M, r_2)$ segundo a métrica $d'(M, N)$ contém uma bola aberta de mesmo centro $B(M, \frac{r_2}{n})$ segundo a métrica $d(M, N)$. Portanto, essas topologias coincidem.

Frisa-se que, para efeitos somente da demonstração do Teorema 2.2.4 a seguir, será adotada a topologia induzida pela norma do máximo em $M_n(\mathbb{C})$.

Teorema 2.2.4 *Seja $(M_n(\mathbb{C}), \|\cdot\|)$ o espaço vetorial das matrizes, munido com a norma do máximo. Então, o conjunto das matrizes diagonalizáveis $\mathcal{D}$ é denso em $(M_n(\mathbb{C}), \|\cdot\|)$.*

Demonstração: Dados $\varepsilon > 0$ e $M \in M_n(\mathbb{C})$, vamos encontrar uma matriz M_ε com n autovalores distintos (e portanto diagonalizável) tal que $\|M - M_\varepsilon\| < \varepsilon$.

Pelo Teorema da Decomposição de Schur (Teorema 1.2.5), existe uma matriz unitária Q (isto é, $Q^{-1} = \bar{Q}^t$) tal que

$$QMQ^{-1} = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ 0 & 0 & \ddots & t_{in} \\ 0 & 0 & 0 & t_{nn} \end{bmatrix} = T. \tag{2.2}$$

Sem perda de generalidade, podemos supor que dentre $t_{11}, \cdots, t_{nn}$ não hajam n números distintos. Pois, caso contrário, a matriz triangular superior T e, portanto, M seriam diagonalizáveis, conforme a Proposição 1.2.3.

Sejam $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, distintos dois a dois, de modo que $|\lambda_i - t_i| < \dfrac{\varepsilon}{n^2 \cdot \|Q\|^2}$, para $i = 1, 2, \ldots, n$. Considera-se a seguinte matriz T_ε

$$QM_\varepsilon Q^{-1} = \begin{bmatrix} \lambda_1 & t_{12} & \cdots & t_{1n} \\ 0 & \lambda_2 & \cdots & t_{2n} \\ 0 & 0 & \ddots & t_{in} \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} = T_\varepsilon, \tag{2.3}$$

obtida de T, substituindo os termos da diagonal principal por λ_i, com $i = 1, 2, \ldots, n$. Observa-se que as matrizes triangulares T_ε e T se distinguem somente pelas suas entradas da diagonal principal e as diferenças dessas entradas não superam o número real $\dfrac{\varepsilon}{n^2 \cdot \|Q\|^2}$. Daí, temos que

$$\|T - T_\varepsilon\| < \frac{\varepsilon}{n^2 \cdot \|Q\|^2}. \tag{2.4}$$

A partir das igualdades (2.2) e (2.3), temos que $M = Q^{-1}TQ$ e $M_\varepsilon = Q^{-1}T_\varepsilon Q$. Pelo Lema 2.2.2, obtem-se que

$$\begin{aligned} \|M - M_\varepsilon\| &= \|Q^{-1}TQ - Q^{-1}T_\varepsilon Q\| \leq n^2 \cdot \|Q\|^2 \cdot \|T - T_\varepsilon\| \\ &\leq n^2 \cdot \|Q\|^2 \cdot \|T - T_\varepsilon\| < n^2 \cdot \|Q\|^2 \cdot \frac{\epsilon}{n^2 \cdot \|Q\|^2} \\ &= \varepsilon. \end{aligned}$$

Notemos que a matriz triangular T_ε, por construção, tem somente autovalores simples. Logo, pela Proposição 1.2.3, a matriz T_ε e, por semelhança, a matriz M_ε é diagonalizável. ∎

Apresentaremos, no Teorema 2.2.6, uma segunda prova da densidade do conjunto $\mathcal{D}$ das matrizes diagonalizáveis. A demonstração desse teorema nos permitirá constatar as analogias entre os argumentos que serão empregados nos Capítulos 3 e 4, e os deste capítulo.

Lema 2.2.5 *Seja $P \in \mathbb{R}[X_1, \ldots, X_n]$ um polinômio não identicamente nulo. Então, o complementar do conjunto dos seus zeros $Z(P)$ é um subconjunto denso em $\mathbb{R}^n$.*

Demonstração: Sejam $a = (a_1, \ldots, a_n)$ um ponto de $\mathbb{R}^n$ e $r > 0$ um número real. Desejamos mostrar que existe algum ponto x do bloco n-dimensional

$$B = \,]a_1 - r, a_1 + r[\; \times \;]a_2 - r, a_2 + r[\; \times \ldots \times \;]a_n - r, a_n + r[$$

tal que $P(x) \neq 0$. Demonstraremos por contraposição. Suponhamos que $P(x) = 0$ para todo ponto x do bloco n-dimensional B. Para cada vetor unitário $\vec{u}$ de $\mathbb{R}^n$, a função polinomial em uma variável de $t \in \mathbb{R}$ definida por $f_{\vec{u}}(t) = P(a + t\vec{u})$ é 0 para todo $t \in]-r, r[$. Então, cada $f_{\vec{u}}(t)$ é constante igual a 0. Isso é o mesmo que dizer que $P(x) = 0$ em cada reta passando pelo ponto a. Isso implica que o polinômio P se anula em todo $\mathbb{R}^n$ e, portanto, o polinômio P é identicamente nulo. Logo, existe algum ponto do complementar $Z(P)$ que pertence ao bloco B. ∎

Teorema 2.2.6 *O conjunto $\mathcal{D} \subset M_n(\mathbb{C})$ de matrizes diagonalizáveis é denso em $M_n(\mathbb{C})$.*

Demonstração: Seja B um bloco $2n^2$-dimensional não-vazio arbitrário. Queremos mostrar que existe uma matriz diagonalizável pertencente a tal bloco.

Considere a função $\varphi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ que associa a cada matriz o discriminante de seu polinômio característico. Como o discriminante do polinômio característico de uma matriz também é um polinômio nas entradas dessa matriz, a função $\varphi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ é polinomial (Observação 1.1.7 e Proposição 1.2.1).

Pelo Corolário 1.1.9, o conjunto $\varphi^{-1}(\{0\})$ consiste de matrizes cujos autovalores não são todos simples. Além disso, conforme o Lema 2.2.5, o conjunto $\varphi^{-1}(\{0\})$ não contém nenhum bloco $2n^2$-dimensional B. Logo, existe uma matriz M cujos autovalores são simples – raízes simples do polinômio característico – que pertence ao bloco $2n^2$-dimensional B e, de acordo com o Lema 1.2.3, diagonalizável. Portanto, o conjunto $\mathcal{D}$ das matrizes diagonalizáveis é denso em $M_n(\mathbb{C})$. ■

3 Medida

"É a Análise Matemática ... apenas um jogo da mente? Ao físico ela só pode dar uma linguagem conveniente; não é este um auxílio medíocre, e, estritamente falando, dispensável? E não é de se temer que essa linguagem artificial seja um véu interposto entre a realidade e a visão do físico? Longe disso; sem essa linguagem, a maior parte das analogias íntimas das coisas teria ficado para sempre desconhecidas por nós; e teríamos ignorado eternamente a harmonia interna do mundo, que é... a única realidade objetiva verdadeira".

Henri Poincaré

O presente capítulo foi organizado em duas seções. Na Seção 3.1, são apresentados conceitos e resultados da medida e integral de Lebesgue em $\mathbb{R}^n$. Na seção 3.2, esses conceitos e resultados são empregados para demonstrar que o conjunto das matrizes não diagonalizáveis tem medida nula. A partir da obra (LANG, 2012), foi elaborado o Apêndice A, o qual exibe os resultados centrais da medida e da integral de Lebesgue em $\mathbb{R}^n$.

3.1 Medida de Lebesgue e Integral de Lebesgue

Nesta seção, faz-se uma brevíssima exposição dos conceitos de medida, sigma-álgebra, funções mensuráveis e conjunto de medida nula. Apresentam-se, ainda, o Teorema de Fubini e a medida de Lebesgue. Maiores detalhes poderão ser obtidos no Apêndice A e na referência (LANG, 2012).

Um conjunto é dito ***contável*** se for finito ou estiver em correspondência biunívoca (*i.e.*, bijeção) com o conjunto dos números naturais. Uma família $\mathscr{M}$ de subconjuntos de um conjunto X é chamada de ***sigma-álgebra*** se contiver qualquer diferença de dois conjuntos da família e a interseção e a união de qualquer coleção contável da família. Em particular, os conjuntos $\emptyset$ e X são membros de $\mathscr{M}$.

Uma função $\mu(A)$, a valores ≥ 0 e $\leq +\infty$, definida sobre uma sigma-álgebra $\mathscr{M}$ de subconjuntos A de um conjunto X, é dita ***contavelmente aditiva*** se, sempre que um conjunto A em $\mathscr{M}$ é uma união de uma família contável de conjuntos A_n de $\mathscr{M}$, dois a dois sem ponto em comum, temos $\mu(A) = \sum_n \mu(A_n)$. Uma tal função será chamada de ***medida***, e os conjuntos de $\mathscr{M}$ serão ditos ***mensuráveis***. A medida μ é ***completa*** se cada subconjunto de cada conjunto mensurável de medida nula é mensurável; caso ela não seja assim, podemos estendê-la de uma única maneira à menor sigma-álgebra que contém $\mathscr{M}$ e todos os subconjuntos de conjuntos de $\mathscr{M}$ de medida nula.

Uma função f definida sobre X, tomando valores reais ≥ 0, é ***mensurável*** se, qualquer que seja o número real a, o conjunto de pontos onde $f(x) \geq a$ é mensurável. Então, as somas de Lebesgue tornam possível atribuir a f uma integral, finita ou não: se esta for finita, $f(x)$ será dita ***integrável*** e a integral será denotada por $\int_X f(x)d\mu$ (onde x designa um ponto genérico de X). Qualquer função formada a partir de funções mensuráveis, pelos procedimentos usuais de operações algébricas e limite de uma sequência contável é mensurável.

Ademais, sejam X e Y dois espaços mensuráveis (que podem coincidir), $\lambda(A)$ uma medida definida em uma sigma-álgebra de subconjuntos A de X, e $\mu(B)$ uma medida definida em uma sigma-álgebra de subconjuntos B de Y. Vamos denotar por x um ponto genérico de X, por y um ponto genérico de Y, e por $A \times B$ o conjunto de pontos (x, y) de $X \times Y$ que satisfazem $x \in A$ e $y \in B$. É possível definir, em $X \times Y$, uma medida ν, chamada ***produto das medidas*** de λ e μ, tal que $\nu(A \times B) = \lambda(A)\mu(B)$ para todo conjunto A mensurável, relativamente a λ, e todo conjunto B mensurável, relativamente a μ; a integral formada por meio de ν possui então todas as propriedades clássicas da integral dupla, e principalmente aquilo que é expresso pelo ***Teorema de Fubini***:

$$\begin{aligned}\int_{X\times Y} f(x,y)d\nu &= \int_X \left(\int_Y f(x,y)\,d\mu\right) d\lambda \\ &= \int_Y \left(\int_X f(x,y)\,d\lambda\right) d\mu\,,\end{aligned}$$

sempre que $f(x, y)$ é integrável em relação a ν.

Para encerrar esta seção, vamos apresentar a construção da medida de Lebesgue e o conceito de conjunto de medida Lebesgue nula em $\mathbb{R}^n$.

Considere a menor sigma-álgebra $\mathscr{B}$ que contém os subconjuntos abertos de $\mathbb{R}^n$, com respeito a topologia usual. Essa sigma-álgebra é denominada a ***sigma-álgebra de Borel*** em $\mathbb{R}^n$ e os conjuntos de $\mathscr{B}$ são chamados ***borelianos***. São portanto borelianos todos os subconjuntos abertos e também todos os fechados de $\mathbb{R}^n$. Se $B = \prod\limits_{i=1}^{n}]a_i, b_i[\; = \;]a_1, b_1[\times \ldots \times]a_n, b_n[$ é um bloco n-dimensional de $\mathbb{R}^n$, em que a_i, $b_i \in \mathbb{R}$, $a_i \leq b_i$ para $i = 1, 2, \ldots, n$, definimos $\mu(B) = \prod\limits_{i=1}^{n}(b_i - a_i)$. Como todo subconjunto A de $\mathbb{R}^n$ pode ser coberto por uma família enumerável de blocos n-dimensional $B_1, B_2, B_3, \cdots$, dois a dois disjuntos (*i.e.*, $A \subseteq \bigcup\limits_{j=1}^{\infty} B_j$), define-se a medida $\mu : \mathscr{B} \longrightarrow [0, \infty]$ por $\mu(A) = \inf \left\{ \sum\limits_{j=1}^{\infty} \mu(B_j) : A \subseteq \bigcup\limits_{j=1}^{\infty} B_j \right\}$.

Em $\mathbb{R}$, a sigma-álgebra de Borel $\mathscr{B}'$ é a menor sigma-álgebra que contém os intervalos abertos. Seja $\mathscr{L}'$ a sigma-álgebra de Lebesgue em $\mathbb{R}$, ou seja, o completamento da sigma-álgebra de Borel em $\mathbb{R}$. Temos que a ***sigma-álgebra de Lebesgue*** em $\mathbb{R}^n$ é o completamento da sigma-álgebra produto $\mathscr{B}' \otimes \cdots \otimes \mathscr{B}'$ de n fatores iguais a sigma-álgebra de Borel; ou equivalentemente, o completamento da sigma-álgebra produto $\mathscr{L}' \otimes \cdots \otimes \mathscr{L}'$ de n fatores iguais a sigma-álgebra de Lebesgue em $\mathbb{R}$. O completamento da medida $\mu : \mathscr{B} \longrightarrow [0, \infty]$ é a ***medida de Lebesgue*** em $\mathbb{R}^n$. Em consequência, todo boreliano de $\mathbb{R}^n$ é mensurável à Lebesgue.

Por fim, um subconjunto Z de $\mathbb{R}^n$ é dito ***conjunto de medida de Lebesgue nula*** se, somente se, dado $\varepsilon > 0$, existir uma família contável $B_1, B_2, B_3, \cdots$ de blocos n-dimensional tais que $Z \subseteq \bigcup_{i \in \mathbb{N}} B_i$ e $\sum_{i \in \mathbb{N}} \mu(B_i) < \varepsilon$. Conjuntos finitos e a união contável de conjuntos de medida nula são exemplos de conjuntos de medida de Lebesgue nula em $\mathbb{R}^n$.

3.2 O Conjunto das Matrizes Diagonalizáveis Tem Medida Total

Na teoria da medida e integração, os conjuntos de medida nula são, para certos propósitos, “negligíveis”. Por exemplo, quando uma função é contínua, a menos de um conjunto de medida nula, diz-se que essa função é contínua em quase todo ponto. Nesta seção, será demonstrado que o conjunto das matrizes não diagonalizáveis é um conjunto de medida nula.

Antes de expor os resultados centrais deste capítulo, vejamos como munir o conjunto das matrizes com uma estrutura de espaço de medida. A ideia é a mesma da Observação 2.2.1.

Observação 3.2.1 *Consideremos a bijeção* $\phi_1 : M_n(\mathbb{C}) \longrightarrow \mathbb{R}^{2n^2}$ *que associa a cada matriz*

$$\begin{bmatrix} x_{11} + iy_{11} & \dots & x_{1n} + iy_{1n} \\ x_{21} + iy_{21} & \dots & x_{2n} + iy_{2n} \\ \vdots & \vdots & \vdots \\ x_{n1} + iy_{n1} & \dots & x_{nn} + iy_{nn} \end{bmatrix}$$

o elemento

$$(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn}, y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn}).$$

Um subconjunto de matrizes será mensurável se, e somente se, sua imagem pela função ϕ_1 *o for. A sua medida será igual à medida de Lebesgue da sua imagem pela função* ϕ_1 *em* $\mathbb{R}^{2n^2}$.

Neste capítulo, a medida de Lebesgue em $\mathbb{R}^n$ será representada por μ_n.

Lema 3.2.2 *Seja $P \in \mathbb{R}[X_1, \ldots, X_n]$ um polinômio não-nulo. Então, o conjunto dos seus zeros $Z(P)$ é mensurável a Lebesgue em $\mathbb{R}^n$.*

Demonstração: Como o polinômio P é contínuo na topologia usual de $\mathbb{R}^n$ e $\{0\}$ é fechado, de acordo com Exemplo 2.1.1, temos que $Z(P)$ é fechado. Logo, $Z(P)$ pertence a sigma-álgebra de Borel, a sigma-álgebra gerada pelos abertos da topologia usual de $\mathbb{R}^n$, pois uma sigma-álgebra contém o complementar de cada um de seus conjuntos e o complementar de um conjunto aberto é fechado. Finalmente, a sigma-álgebra de Lebesgue contém a sigma-álgebra de Borel, isto é, os borelianos são mensuráveis a Lebesgue. Em particular, $Z(P)$ é mensurável a Lebesgue. O que conclui a demonstração do presente lema. ■

Para ilustrar o uso do Teorema de Fubini na demonstração do Lema 3.2.4 a seguir, elaborou-se o Apêndice B.

Observação 3.2.3 *Uma função $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ é dita simples quando assume um número finito de valores. Ou seja, essa função pode ser representada $f(x) = \sum_{i=1}^{n} f(A_i)\chi_{A_i}(x)$, em que $f(A_i)$ são números reais, $\chi_{A_i} : \mathbb{R}^n \longrightarrow \mathbb{R}$ é uma função que assume valor 1 nos pontos de A_i e 0 nos demais pontos e cada A_i é um subconjunto de $\mathbb{R}^n$ mensurável a Lebesgue. A integral de Lebesgue de uma função simples é definida por*

$$\int_{\mathbb{R}^n} f \; d\mu_n = \sum_{i=1}^{n} f(A_i)\mu_n(A_i).$$

Em particular, a medida de um subconjunto mensurável $Z(P)$ de $\mathbb{R}^n$ é dada por

$$\int_{\mathbb{R}^n} \chi_{Z(P)} \; d\mu_n = \mu_n(Z(P)).$$

Lema 3.2.4 *Seja $P \in \mathbb{R}[X_1, \ldots, X_n]$ um polinômio não-nulo. Então, $Z(P)$ é um conjunto de medida de Lebesgue nula em $\mathbb{R}^n$.*

Demonstração: Provaremos esse lema por indução sobre n. Para $n = 1$, o polinômio $P \in \mathbb{R}[X_1]$ possui apenas uma indeterminada. Como P é não-nulo, tem-se que $Z(P)$ é finito e, portanto, $\mu(Z(P)) = 0$.

Suponhamos, por hipótese de indução, que o lema é válido para todo polinômio não-nulo em $n-1$ indeterminadas. O polinômio P pode ser escrito como

$$P(X_1, \ldots, X_n) = Q_0(X_1, \ldots, X_{n-1}) + Q_1(X_1, \ldots, X_{n-1})X_n + \ldots + Q_m(X_1, \ldots, X_{n-1})X_n^m,$$

em que $Q_0(X_1, \ldots, X_{n-1}), \ldots, Q_m(X_1, \ldots, X_{n-1})$ são polinômios em $n-1$ indeterminadas e m um número natural. Como o polinômio P é não-nulo, vamos supor que $Q_m(X_1, \ldots, X_{n-1})$ não é identicamente nulo.

Conforme a Observação 3.2.3, tem-se

$$\mu_n(Z(P)) = \int_{\mathbb{R}^n} \chi_{Z(P)}(x_1, \ldots, x_n)dx_1 \ldots dx_n,$$

em que $\chi_{Z(P)} : \mathbb{R}^n \longrightarrow \mathbb{R}$ é a função característica de $Z(P)$, que associa o valor 1 aos pontos do conjunto $Z(P)$ e o valor 0 aos demais pontos de $\mathbb{R}^n$, e μ_n denota a medida de Lebesgue em $\mathbb{R}^n$.

A função $\chi_{Z(P)} : \mathbb{R}^n \longrightarrow \mathbb{R}$ é mensurável, porque o conjunto $Z(P)$ é mensurável a Lebesgue, conforme o Lema 3.2.2. Então, pelo Teorema de Fubini, obtemos

$$\int_{\mathbb{R}^n} \chi_{Z(P)}(x_1, \ldots, x_n)dx_1 \cdots dx_n = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x_1, \ldots, x_n)dx_n \right) dx_1 \cdots dx_{n-1}.$$

A função seção $f(x_1, \ldots, x_{n-1}) = \int_{\mathbb{R}} \chi_{Z(P)}(x_1, \cdots, x_n)dx_n$ está definida em quase todo ponto de $\mathbb{R}^{n-1}$. De fato, o conjunto $A = \{(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} : P(x_1, \ldots x_n) = 0$, para todo número real $x_n\}$ possui medida nula. Um ponto de $\mathbb{R}^{n-1}$ será elemento de A se for zero dos polinômios $Q_0(X_1, \ldots, X_{n-1}), \ldots, Q_m(X_1, \ldots, X_{n-1})$, em particular, do polinômio $Q_m(X_1, \ldots, X_{n-1})$. Por hipótese de indução, o conjunto dos zeros de $Q_m(X_1, \ldots, X_{n-1})$ tem medida nula e, consequentemente, o conjunto A tem medida nula.

Ademais, temos que $f(x_1, \ldots, x_{n-1}) = \int_{\mathbb{R}} \chi_{Z(P)}(x_1, \ldots, x_n)dx_n = 0$ em quase todo ponto de $\mathbb{R}^{n-1}$. Realmente, para cada ponto fixado $(x_1, \cdots, x_{n-1})$ do complementar de A, o polinômio $P(x_1, \ldots, x_{n-1}, X_n)$ em uma indeterminada X_n não é identicamente nulo. Donde, o conjunto de suas raízes é finito e, consequentemente, a função $\chi_{Z(P)}(x_1, \ldots, x_{n-1}, \cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ assume valor 1, num conjunto de medida nula, e 0 nos demais pontos de $\mathbb{R}$.

Logo,

$$\begin{aligned} \mu_n(Z(P)) &= \int_{\mathbb{R}^n} \chi_{Z(P)}(x_1, \ldots, x_n)dx_1 \ldots dx_n \\ &= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x_1, \ldots, x_n)dx_n \right) dx_1 \ldots dx_{n-1} \\ &= \int_{\mathbb{R}^{n-1}} f(x_1 \ldots, x_{n-1})dx_1 \ldots dx_{n-1} = 0. \end{aligned}$$

Portanto, pelo Princípio de Indução Finita, $\mu_n(Z(P)) = 0$ para todo número natural n. O que conclui a demonstração do lema. ∎

Teorema 3.2.5 *O conjunto $\mathcal{D} \subset M_n(\mathbb{C})$ das matrizes diagonalizáveis é um conjunto de medida de Lebesgue total (isto é, o seu complementar é um conjunto de medida nula).*

Demonstração: Seja $\varphi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ a função que faz corresponder a cada matriz o discriminante do seu polinômio característico. Do fato do discriminante do polinômio característico de uma matriz ser um polinômio nas entradas dessa matriz (Observação 1.1.7 e Proposição 1.2.1), temos que a função $\varphi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ é polinomial.

Pelo Lema 3.2.4, o conjunto $\varphi^{-1}(\{0\})$ tem medida de Lebesgue nula e, pelo Corolário 1.1.9, $\varphi^{-1}(\{0\})$ consiste de matrizes cujos autovalores não são todos simples. Então, o seu complementar é um subconjunto das matrizes diagonalizáveis, conforme o Lema 1.2.3, e possui medida total. Então, o conjunto $\mathcal{D}$ das matrizes diagonalizáveis tem medida de Lebesgue total. ■

4 Álgebra

"A Álgebra é generosa: frequentemente ela dá mais do que se lhe pediu".

D'Alembert

Os resultados obtidos, nos capítulos de topologia e de medida, se fundamentaram na análise do conjunto dos zeros de polinômios. Assim, torna-se razoável estudar abstratamente conjuntos dos zeros de polinômios em espaços $\mathbb{K}^n$, desprovidos de estruturas especiais (topologias, medidas e espaço vetorial). Esse capítulo se assenta nessas ideias.

4.1 Topologia de Zariski

Nesta seção, são introduzidos os conceitos de anel, ideal e corpo. Ademais, a topologia de Zariski é definida e demonstra-se o Teorema da Base de Hilbert (Teorema 4.1.2), por meio do qual, conclui-se que o estudo de subconjuntos do anel de polinômios pode ser resumido àqueles que são constituídos por um número finito de polinômios. Adotou-se como referência para esta seção (ATIYAH, 1969).

Chamamos de ***operação binária sobre um conjunto*** $\mathcal{A}$ uma função $* : \mathcal{A} \times \mathcal{A} \longrightarrow \mathcal{A}$ que toma um par (a, b) de elementos de $\mathcal{A}$ e o faz corresponder a um elemento $a * b$ também de $\mathcal{A}$. Um conjunto não-vazio $\mathcal{A}$ em que estão definidas duas operações $+$ (adição) e $\cdot$ (multiplicação) é dito ***anel*** se essas operações possuem as propriedades a seguir, para quaisquer a, b e $c \in \mathcal{A}$:

1. associatividade da adição: $(a + b) + c = a + (b + c)$;

2. existência de elemento neutro para adição: existe $0 \in A$ tal que $a + 0 = a$;

3. existência de inverso aditivo: para todo $a \in \mathcal{A}$ existe um único $-a \in \mathcal{A}$ tal que $a + (-a) = 0$.

4. comutatividade da adição: $a + b = b + a$;

5. distributividade à esquerda e à direita: $a\cdot(b+c) = a\cdot b+a\cdot c$ e $(a+b)\cdot c = a\cdot c+b\cdot c$.

O conjunto dos números inteiros $\mathbb{Z}$ e o conjunto dos números reais $\mathbb{R}$, munidos com as operações de adição e multiplicação usuais, são exemplos de anéis.

Denominamos ***ideal*** de um anel $\mathcal{A}$ um subconjunto não-vazio $\mathcal{I}$ de $\mathcal{A}$ que goza das propriedades:

1. se $x, y \in \mathcal{I}$, então $x + (-y) \in \mathcal{I}$;

2. se $a \in \mathcal{A}$ e $x \in \mathcal{I}$, então $ax \in \mathcal{I}$ e $xa \in \mathcal{I}$.

Por exemplo, o subconjunto $\{0\}$ e o próprio $\mathbb{R}$ são ideais do anel $(\mathbb{R}, +, \cdot)$. Fixado um número inteiro m, o subconjunto $m\mathbb{Z} = \{mx : x \in \mathbb{Z}\}$ é um ideal no anel dos inteiros $(\mathbb{Z}, +, \cdot)$. Nesses exemplos de ideais, as operações de adição e multiplicação são as usuais.

Seja S um subconjunto não-vazio de um anel $\mathcal{A}$. Dizemos que um ***ideal é gerado por S***, representado por $\langle S \rangle$, se ele é o conjunto de todas as somas finitas da forma $a_1 s_1 b_1 + \ldots + a_n s_n b_n$, com $a_i,\ b_i \in \mathcal{A}$ e $s_i \in S$ para todo $1 \leq i \leq n$. É possível demonstrar que o ideal gerado por S é o menor ideal do anel $\mathcal{A}$ que contém S.

Por exemplo, o conjunto de todos os polinômios em uma indeterminada, com coeficientes complexos, múltiplos de $X + i$ (em que i é a unidade imaginária) é um ideal finitamente gerado em $\mathbb{C}[X]$. De fato, o conjunto gerador desse ideal é finito, especificamente, o conjunto unitário $\{X + i\}$. Os elementos desse ideal são da forma $\sum_{j}^{n} P_j(X)(X + i)Q_j(X)$, em que n é um natural, cada $P_j(X)$ e cada $Q_j(X)$ pertence ao anel dos polinômios $\mathbb{C}[X]$.

Um ***corpo*** $\mathbb{K}$ é um anel com as operações adição $+$ e multiplicação $\cdot$, que dispõem das propriedades:

1. associatividade da multiplicação: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;

2. existência de elemento neutro para multiplicação: existe $1 \in \mathbb{K}$ tal que $a \cdot 1 = 1 \cdot a = a$

3. comutatividade da multiplicação: $a \cdot b = b \cdot a$;

4. existência de inverso multiplicativo: para todo $a \in \mathbb{K}$ não-nulo, existe um único $a^{-1} \in \mathbb{K}$ tal que $a \cdot a^{-1} = 1$,

para quaisquer a, b e c em $\mathbb{K}$.

O conjunto dos números reais $\mathbb{R}$ e o conjunto dos números complexos $\mathbb{C}$ são exemplos de corpos. O anel $\mathbb{Z}$ dos números inteiros não é um corpo.

Sejam $\mathbb{K}$ um corpo e S um subconjunto do anel $\mathbb{K}[X_1, \ldots, X_n]$ dos polinômios sobre $\mathbb{K}$ nas indeterminadas $X_1, \cdots, X_n$. Definimos o seguinte conjunto

$$Z(S) = \{x \in \mathbb{K}^n : P(x) = 0, \forall P \in S\}.$$

Assim, $Z(S)$ é o conjunto dos pontos $x \in \mathbb{K}^n$ que são zeros de todos os polinômios $P \in S$. Quando um subconjunto S possui somente o polinômio $P \in \mathbb{K}[X_1, \ldots, X_n]$, denotaremos simplesmente por $Z(P)$ o conjunto dos pontos $x \in \mathbb{K}^n$ que são zeros de P.

Proposição 4.1.1 *Seja $\mathbb{K}$ um corpo. Valem as propriedades em $\mathbb{K}^n$:*

1. *$Z(\mathbb{K}[X_1, \ldots, X_n]) = \emptyset$, $Z(\emptyset) = \mathbb{K}^n$;*
2. *se S e S' são subconjuntos de $\mathbb{K}[X_1, \ldots, X_n]$, então $Z(S) \cup Z(S') = Z(SS')$, onde SS' é o conjunto formado pelos produtos P_1P_2, com $P_1 \in S$ e $P_2 \in S'$;*
3. *se S_i são subconjuntos de $\mathbb{K}[X_1, \ldots, X_n]$, para algum conjunto J (finito ou infinito) de índices, então $\bigcap\limits_{i \in J} Z(S_i) = Z(\bigcup\limits_{i \in J} S_i)$.*

Demonstração: 1. Não existe ponto de $\mathbb{K}^n$ que é zero de todos os polinômios do anel $\mathbb{K}[X_1, \ldots, X_n]$. De fato, o polinômio constante $P(X_1, \ldots, X_n) = a$, em que a é um ponto não-nulo de $\mathbb{K}^n$, não possui zeros. Logo, $Z(\mathbb{K}[X_1, \ldots, X_n]) = \emptyset$. Por vacuidade, temos que $Z(\emptyset) = \mathbb{K}^n$.

2. Observemos que $P_1P_2(x) = 0$ se, somente se, $P_1(x) = 0$ ou $P_2(x) = 0$, em que x é um ponto de $\mathbb{K}^n$. Assim, os zeros comuns dos polinômios P_1P_2 de SS' são também comuns aos polinômios P_1 de S ou aos polinômios P_2 de S', e reciprocamente. Portanto, $Z(S) \cup Z(S') = Z(SS')$.

3. Seja S_i uma família de subconjuntos de $\mathbb{K}[X_1, \ldots, X_n]$, em que i pertence a algum conjunto índice J (finito ou infinito). Como $S_i \subseteq \bigcup\limits_{i \in J} S_i$, tem-se que $Z(\bigcup\limits_{i \in J} S_i)) \subseteq Z(S_i)$, para todo $i \in J$; porque, ao aumentar a quantidade de polinômios, diminue a quantidade de zeros comuns aos polinômios. Daí, obtém-se que $Z(\bigcup\limits_{i \in J} S_i)) \subseteq \bigcap\limits_{i \in J} Z(S_i)$.

Por outro lado, a validade de $\bigcap\limits_{i \in J} Z(S_i) \subset Z(\bigcup\limits_{i \in J} S_i))$ decorre de $\bigcap\limits_{i \in J} Z(S_i) \subseteq Z(S_i)$, para todo $i \in J$. Portanto, $\bigcap\limits_{i \in J} Z(S_i) = Z(\bigcup\limits_{i \in J} S_i)$. ■

As propriedades 1), 2) e 3) da Proposição 4.1.1 mostram que os conjuntos $Z(S)$ satisfazem os axiomas para conjuntos fechados em um espaço topológico, ou seja, os conjuntos abertos em $\mathbb{K}^n$ são complementares dos conjuntos $Z(S)$. A topologia assim definida em $\mathbb{K}^n$ é chamada de ***topologia de Zariski***.

Em Álgebra, muitos problemas consistem em estudar os ideias de um anel. No contexto do anel de polinômios $\mathbb{K}[X_1, \ldots, X_n]$, esse estudo é simplificado, pois cada ideal é gerado por um número finito de polinômios. Esse é o conteúdo do Teorema da Base de Hilbert que será demonstrado a seguir.

Teorema 4.1.2 (Teorema da Base de Hilbert) *Seja $\mathbb{K}$ um corpo. Então, todo ideal $\mathcal{I}$ do anel de polinômios $\mathbb{K}[X_1, \ldots, X_n]$ pode ser gerado por um conjunto finito de polinômios.*

Demonstração: Provaremos por indução sobre n. Se $n = 1$, então cada ideal $\mathcal{I}$ do anel de polinômios $\mathbb{K}[X_1]$ é gerado por um único polinômio, de acordo com a Proposição 1.1.2.

Suponhamos, por hipótese de indução, que todo ideal do anel de polinômios $\mathbb{K}[X_1, \ldots, X_{n-1}]$ pode ser gerado por um conjunto finito de polinômios. Se $\mathcal{I} = \langle 0 \rangle$, então o ideal $\mathcal{I}$ é gerado pelo único polinômio $P(X_1, \ldots, X_n) = 0$. Se $\mathcal{I} \neq \langle 0 \rangle$, existe P_1 em $\mathcal{I}$, não-nulo e de grau d_1 mínimo. Seja $Q_1 \in \mathbb{K}[X_1, \ldots, X_{n-1}]$ o coeficiente líder de P_1, isto é, $P_1 = Q_1 X_n^{d_1} +$ termos de grau menor e termos nulos.

Se $\mathcal{I} = \langle P_1 \rangle$, a prova termina. Se $\mathcal{I} \neq \langle P_1 \rangle$, existe P_2 em $\mathcal{I}$, distinto de P_1 e de grau d_2 mínimo com essa propriedade; temos $d_2 \geq d_1$, caso contrário teríamos escolhido incorretamente P_1. Seja $Q_2 \in \mathbb{K}[X_1, \ldots, X_{n-1}]$ o coeficiente líder de P_2. Continua-se enquanto $\mathcal{I} \neq \langle P_1, P_2, \ldots, P_m \rangle$, produzindo $d_1 \leq d_2 \leq \cdots \leq d_{m+1}$ e coeficientes líderes $Q_1, Q_2, \ldots, Q_{m+1}$. Seja m tal que $\langle Q_1, Q_2, \ldots, Q_m \rangle = \langle Q_1, Q_2, \ldots, Q_{m+1} \rangle$, pois, por hipótese de indução, todo ideal do anel de polinômios $\mathbb{K}[X_1, \ldots, X_{n-1}]$ pode ser gerado por um conjunto finito de polinômios. Donde, resulta a relação

$$Q_{m+1} = R_1 Q_1 + R_2 Q_2 + \cdots + R_m Q_m, \ R_i \in \mathbb{K}[X_1, \ldots, X_{n-1}].$$

Efetuamos agora o seguinte cálculo, fazendo $u = d_{m+1}$ (grau de P_{m+1}),

$$S = P_{m+1} - (R_1 X_n^{u-d_1} P_1 + R_2 X_n^{u-d_2} P_2 + \cdots + R_m X_n^{u-d_m} P_m).$$

O termo líder de P_{m+1} cancela com o da expressão entre parênteses. Notemos que $S \neq 0$ do contrário P_{m+1} pertenceria ao ideal dos anteriores, contrariando o procedimento. Por construção, S é um elemento de $\mathcal{I}$. Mas o grau de S é estritamente menor que d_{m+1}, contradizendo a escolha de P_{m+1}. Então, o processo encerra com $\mathcal{I} = \langle P_1, P_2, \ldots, P_m \rangle$. ∎

Corolário 4.1.3 *Seja $\langle S \rangle$ o ideal em $\mathbb{K}[X_1, \ldots, X_n]$ gerado por S, então $Z(S) = Z(\langle S \rangle)$.*

Demonstração: Do fato de $S \subset \langle S \rangle$, resulta que $Z(\langle S \rangle) \subseteq Z(S)$. Pelo Teorema da Base de Hilbert, tem-se que todo elemento P de $\langle S \rangle$ é da forma $P_1 Q_1 + \ldots + P_m Q_m$ para $Q_i \in \mathbb{K}[X_1, \ldots, X_n]$ e $P_i \in S$ apropriados. Decorre que, se $x \in Z(S)$, então $P_i(x) = 0$ para todo $P_i \in S$ e, assim, $P(x) = 0$. Logo, $x \in Z(\langle S \rangle)$ e, como x é arbitrário, concluímos que $Z(S) \subseteq Z(\langle S \rangle)$. Portanto, $Z(S) = Z(\langle S \rangle)$. Como queríamos demonstrar. ∎

4.2 O Conjunto das Matrizes Diagonalizáveis é Aberto e Denso

Nesta seção, provaremos que todo aberto não-vazio é denso na topologia de Zariski (Lema 4.2.5) e os abertos não-vazios de Zariski também são densos na topologia usual de $\mathbb{R}^n$ (Lema 4.2.3). Por meio dos Lemas 4.2.5 e 4.2.3, demonstraremos que conjunto das matrizes diagonalizáveis é denso relativamente a essas topologias (Teorema 4.2.4).

Lema 4.2.1 *Seja $\mathbb{K}$ um corpo infinito. Se $P \in \mathbb{K}[X_1, \ldots, X_n]$ é um polinômio tal que $P(a_1, \ldots, a_n) = 0$ para todo ponto $a = (a_1, \ldots, a_n) \in \mathbb{K}^n$, então o polinômio P é identicamente nulo.*

Demonstração: Demonstraremos esse lema por indução sobre o número n de indeterminadas do polinômio P. Se $n = 1$ e $P \in \mathbb{K}[X_1]$ é um polinômio tal que $P(a) = 0$ para todo ponto $a \in \mathbb{K}$, então o polinômio P numa indeterminada tem infinitas raízes, porque $\mathbb{K}$ é um corpo infinito. Porém, todo polinômio não-nulo sobre $\mathbb{K}$ numa indeterminada tem um número finito de raízes em $\mathbb{K}$. Logo, P é identicamente nulo.

Suponhamos, por hipótese de indução, que o lema seja válido para todo polinômio sobre $\mathbb{K}$ nas indeterminadas $X_1, \ldots, X_r$, em que $1 \leq r < n$. Consideremos, agora, o polinômio P sobre $\mathbb{K}$ nas indeterminadas $X_1, \ldots, X_n$ (ou seja, em n indeterminadas) tal que $P(a_1, a_2, \ldots, a_n) = 0$ para todo $a = (a_1, a_2, \ldots, a_n) \in \mathbb{K}^n$. Podemos escrevê-lo como $P = \sum_{i=1}^{m} P_i(X_1, \ldots, X_r)X_n^i$, em que cada P_i é um polinômio sobre $\mathbb{K}$ nas indeterminadas $X_1, \ldots, X_r$, em que $1 \leq r < n$.

Fixemos um ponto $b = (b_1, b_2, \ldots, b_{n-1}) \in \mathbb{K}^{n-1}$. Assim, o polinômio $Q = \sum_{i=1}^{m} P_i(b_1, b_2, \ldots, b_{n-1})X_n^i$ tem infinitas raízes em $\mathbb{K}$.

Por outro lado, o polinômio Q possui uma única indeterminada, a saber, X_n. Como um polinômio não identicamente nulo numa indeterminada sobre um corpo $\mathbb{K}$ tem um número finito de raízes, o polinômio Q é identicamente nulo e, portanto, $P_i(b_1, b_2, \ldots, b_{n-1}) = 0$, para todo $1 \leq i \leq m$. Como o ponto $b = (b_1, b_2, \ldots, b_{n-1}) \in \mathbb{K}^{n-1}$ foi tomado arbitrariamente, resulta da hipótese de indução que os polinômios P_i são identicamente nulos. Daí, concluímos que o polinômio $P = \sum_{i=1}^{m} P_i(X_1, \ldots, X_r)X_n^i$ é identicamente nulo. Pelo Princípio da Indução Finita, o presente lema é válido para todo polinômio P sobre $\mathbb{K}$ nas indeterminadas $X_1, \ldots, X_n$ e todo natural n. ■

Observação 4.2.2 *Notemos que o complementar de um conjunto $Z(S)$ pode ser escrito como a união finita dos complementares dos conjuntos $Z(P_i)$, em que cada P_i pertence ao ideal gerado por $S \subset \mathbb{K}[X_1, \ldots, X_n]$, isso segue do Teorema da Base de Hilbert (Teorema 4.1.2) e da Proposição 4.1.1.*

Lema 4.2.3 *Todo aberto não-vazio da topologia de Zariski em $\mathbb{R}^n$ é um subconjunto aberto e denso da topologia usual em $\mathbb{R}^n$.*

Demonstração: Seja $P \in \mathbb{R}[X_1, \cdots, X_n]$ um polinômio tal que o complementar do conjunto $Z(P)$ é não-vazio. Observemos que, de acordo com a Observação 4.2.2, basta mostrarmos que o complementar de $Z(P)$ é aberto e denso da topologia usual em $\mathbb{R}^n$.

Provamos, primeiro, que ele é aberto. De acordo com a Proposição 2.1.2, todo polinômio é contínuo segundo a topologia usual em $\mathbb{R}^n$ e, em particular, P. Por definição, $Z(P) = P^{-1}(\{0\})$ é o conjunto dos zeros de P. Como o conjunto $\{0\}$ é fechado da topologia usual em $\mathbb{R}^n$ – consequência do Exemplo 2.1.1 – e P é contínuo, concluímos que $Z(P)$ é fechado e, portanto, seu complementar é aberto.

Agora, passemos a provar a densidade do complementar de $Z(P)$. Seja B um bloco n-dimensional não-vazio de $\mathbb{R}^n$. Com fundamento no Lema 2.2.5, $Z(P)$ não contém B, daí a interseção do complementar de $Z(P)$ e o bloco n-dimensional B é não-vazia. Portanto, todo aberto não-vazio da topologia de Zariski é aberto e denso da topologia usual em $\mathbb{R}^n$. Como queríamos demonstrar. ∎

Teorema 4.2.4 *O conjunto $\mathcal{D} \subset M_n(\mathbb{C})$ de matrizes diagonalizáveis é denso com respeito a topologia usual em $M_n(\mathbb{C})$.*

Demonstração: Consideremos a função $\varphi : M_n(\mathbb{C}) \longrightarrow \mathbb{C}$ que associa a cada matriz o discriminante do seu polinômio característico. Essa função φ é polinomial, em virtude do discriminante do polinômio característico de uma matriz ser um polinômio nas entradas dessa matriz (Observação 1.1.7 e Proposição 1.2.1).

O complementar do conjunto $\varphi^{-1}(\{0\})$, conforme o Corolário 1.1.9, consiste de matrizes que só tem autovalores simples e, consequentemente, diagonalizáveis nos termos do Lema 1.2.3. Nota-se que o complementar do conjunto $\varphi^{-1}(\{0\})$ é não-vazio, por causa do Lema 4.2.1 e do fato de o polinômio φ não ser identicamente nulo. Então, pelo Lema 4.2.3, o complementar do conjunto $\varphi^{-1}(\{0\})$ é denso, relativamente a topologia usual em $M_n(\mathbb{C})$. Portanto, o conjunto $\mathcal{D}$ de matrizes diagonalizáveis é denso em $M_n(\mathbb{C})$. ∎

Constata-se que foram usados apenas polinômios, na demonstração do Teorema 4.2.4. Essencialmente, a prova de que o conjunto $\mathcal{D}$ de matrizes diagonalizáveis é denso ou possui medida total se reduziu a apenas estudar zeros de polinômios.

Ao encerrar essa seção, colecionamos algumas propriedades da topologia de Zariski, na Proposição 4.2.5, a seguir.

Proposição 4.2.5 *Seja $\mathbb{K}$ um corpo. Considerando $\mathbb{K}^n$ com a topologia de Zariski, temos:*

1. *se $\mathbb{K}$ é infinito, então todo aberto não-vazio de $\mathbb{K}^n$ é denso;*

2. se $\mathbb{K}$ é finito, então cada subconjunto de $\mathbb{K}^n$ é um fechado.

3. toda função $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ que é polinomial em cada coordenada é contínua.

Demonstração: 1. Dado um polinômio $P \in \mathbb{K}[X_1, \ldots, X_n]$, é suficiente mostrar que se o complementar de $Z(P)$ for não-vazio, então ele será denso na topologia de Zariski em $\mathbb{K}^n$, conforme a Observação 4.2.2.

Tomemos um ponto a de $\mathbb{K}^n$ que pertença ao conjunto $Z(P)$. Suponhamos que o complementar do conjunto $Z(Q)$, com $Q \in \mathbb{K}[X_1, \ldots, X_n]$, é não-vazio e possui também o ponto a. Desejamos provar que existe um outro ponto de $\mathbb{K}^n$ que pertence tanto ao complementar de $Z(P)$ quanto ao de $Z(Q)$.

Com efeito, pela Proposição 4.1.1, resulta que $Z(PQ) = Z(P) \cup Z(Q)$. Como os polinômios P e Q não são identicamente nulos, temos que PQ também é um polinômio não identicamente nulo. Então, pelo Lema 4.2.1, existe um ponto b de $\mathbb{K}^n$ que pertence ao complementar de $Z(PQ)$. Logo, aplicando as Leis de De Morgan em $Z(PQ) = Z(P) \cup Z(Q)$, tem-se que b pertence tanto ao complementar de $Z(P)$ quanto ao de $Z(Q)$. Portanto, o complementar de $Z(P)$ é denso na topologia de Zariski em $\mathbb{K}^n$.

2. Se $\mathbb{K}$ é finito, então cada ponto $(a_1, a_2, \ldots, a_n)$ de $\mathbb{K}^n$ é um fechado, porque $Z(\{X_1 - a_1, X_2 - a_2, \ldots, X_n - a_n\}) = \{(a_1, a_2, \ldots, a_n)\}$. Portanto, todo subconjunto de pontos de $\mathbb{K}^n$ é um fechado, concernente a topologia de Zariski, pois são uniões finitas de fechados. Logo, a topologia de Zariski é a discreta em $\mathbb{K}^n$.

3. Basta verificarmos a continuidade da função $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ para o complementar de $Z(P)$, com $P \in \mathbb{K}[X_1, \ldots, X_n]$, de acordo com a Observação 4.2.2. Como a função $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ é polinomial em cada coordenada, temos que

$$\varphi(a) = (P_1(a_1, \ldots, a_n), \ldots, P_n(a_1, \ldots, a_n)),$$

em que, para cada $1 \leq i \leq n$, P_i é um polinômio nas indeterminadas $X_1, \ldots, X_n$ e $a = (a_1, \ldots, a_n)$ é um ponto de $\mathbb{K}^n$. Denotando $P_i(a) = P_i(a_1, \ldots, a_n)$ e o complementar de $Z(P)$ por $Z(P)^c$, temos

$$\varphi^{-1}(Z(P)^c) = \{a \in \mathbb{K}^n : P(\varphi(a)) = P(P_1(a), \ldots, P_n(a)) \neq 0\}.$$

Observando que $\varphi^{-1}(Z(P)^c) = Z(Q)^c$, em que $Q(X_1, \ldots, X_n) = P(P_1(a), \ldots, P_n(a))$, concluímos que $\varphi^{-1}(Z(P)^c)$ é um aberto e, consequentemente, $\varphi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ é uma função contínua, relativamente a topologia de Zariski. ■

5 Considerações Finais

"A história ensina a continuidade do desenvolvimento da ciência. Sabemos que toda era tem seus próprios problemas, que a idade seguinte resolve ou deixa de lado como sem proveito e substitui por novos."

David Hilbert

Nesta dissertação, foi demonstrado que quase todas as matrizes complexas são diagonalizáveis, por meio de diferentes abordagens, a saber: da Topologia, da Medida e da Álgebra. Nesse contexto, o leitor poderia questionar-se se o problema – quase todas as matrizes complexas são diagonalizáveis – poderia ser abordado por meio da Probabilidade. Para uma demonstração destes resultados probabilísticos acerca de matrizes diagonalizáveis, o leitor pode consultar o artigo (HETZEL; LIEW; MORRISON, 2007).

Nos capítulos de Topologia e Medida, foi demonstrado que o conjunto das matrizes diagonalizáveis, visto como subconjunto de $\mathbb{R}^{2n^2}$, é denso na topologia usual (Teoremas 2.2.4 e 2.2.6) e possui medida de Lebesgue total (Teorema 3.2.5). O Teorema 3.2.5 é mais forte que os Teoremas 2.2.4 e 2.2.6, ou seja, a medida total implica a densidade. De fato, o complementar do conjunto dos zeros de um polinômio, não identicamente nulo, ser denso em $\mathbb{R}^n$ (Lema 2.2.5) decorre do fato de o conjunto dos zeros de um polinômio ter medida de Lebesgue nula (Lema 3.2.4) e os Teoremas 2.2.4 e 2.2.6, do Teorema 3.2.5.

A densidade do conjunto das matrizes complexas diagonalizáveis foi demonstrado nas topologias da norma do máximo em $M_n(\mathbb{C})$, da usual e de Zariski em $\mathbb{R}^{2n^2}$. As topologias induzida pela norma do máximo em $M_n(\mathbb{C})$ e a usual de $\mathbb{R}^{2n^2}$ são equivalentes em $M_n(\mathbb{C})$. Ser denso em norma implica a densidade em Zariski, ou seja, aquela é mais fina do que esta; vale a recíproca parcial no contexto do $\mathbb{R}^n$, conforme o Lema 4.2.3.

Do ponto de vista heurístico, há uma analogia entre as abordagens da Topologia, da Medida e da Álgebra, ao se provar que quase todas as matrizes complexas são diagonalizáveis: todas se fundamentaram no estudo de zeros do discriminante do polinômio característico de cada matriz. Esse é o núcleo do problema desta pesquisa.

O conjunto das matrizes reais invertíveis é aberto em $M_n(\mathbb{R})$, com respeito a topologia usual de $\mathbb{R}^{n^2}$ (Observação 2.2.1). Realmente, consideremos a função $\psi : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ que associa a cada matriz o seu determinante. Como a função determinante de uma matriz é polinomial nas entradas da matriz, a função $\psi : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ é contínua com respeito a topologia usual de $\mathbb{R}^{n^2}$, de acordo com a Proposição 2.1.2. Do Exemplo 2.1.1, temos que o conjunto $\mathbb{R} - \{0\}$ é aberto e, consequentemente, a imagem inversa $\psi^{-1}(\mathbb{R} - \{0\})$ é um subconjunto aberto de $M_n(\mathbb{R})$. Do resultado bastante conhecido de Álgebra Linear – uma matriz será invertível se, somente se, o seu determinante for diferente de zero – concluímos que $\psi^{-1}(\mathbb{R} - \{0\})$ é conjunto das matrizes invertíveis e, portanto, aberto. Notemos que, aplicando o Lema 2.2.5 do Capítulo de Topologia à função $\psi : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$, concluímos também que o conjunto das matrizes invertíveis de $M_n(\mathbb{C})$ é denso, com respeito a topologia usual de $\mathbb{R}^{n^2}$ (Observação 2.2.1).

Sobre o corpo dos números reais, não há resultados análogos aos Teoremas 2.2.4 e 3.2.5. Por exemplo, considere a função $\varphi : M_2(\mathbb{R}) \longrightarrow \mathbb{C}$ que associa a cada matriz o discriminante do seu polinômio característico e adotemos a topolologia usual de $\mathbb{R}^4$ em $M_2(\mathbb{R})$ (Observação 2.2.1). Essa função φ é polinomial (Observação 1.1.7 e Proposição 1.2.1). Logo, a função $\varphi : M_2(\mathbb{R}) \longrightarrow \mathbb{C}$ é contínua, de acordo com a Proposição 2.1.2. Consideremos o intervalo $I =]-\infty, 0[$. Daí, o conjunto $\varphi^{-1}(I)$ é aberto em $M_n(\mathbb{R})$, porque $I =]-\infty, 0[$ é aberto em $\mathbb{R}$. Ora, o conjunto $\varphi^{-1}(I)$ consiste de matrizes não diagonalizáveis em $\mathbb{R}$, pois o discriminante do polinômio característico das matrizes de $\varphi^{-1}(I)$ é negativo. Como o conjunto $\varphi^{-1}(I)$ é aberto, temos que, para cada matriz $M \in \varphi^{-1}(I)$, existe um bloco 4-dimensional B tal que $M \in B \subset \varphi^{-1}(I)$. Donde, o bloco B consiste de somente matrizes não diagonalizáveis. Portanto, o conjunto das matrizes diagonalizáveis de $M_2(\mathbb{R})$ não é denso, com respeito a topologia usual de $\mathbb{R}^4$.

Há, todavia, o seguinte resultado: o fecho do conjunto das matrizes reais diagonalizáveis é o conjunto das matrizes reais triangularizáveis. Em outras palavras, para cada número real $\varepsilon > 0$ e cada matriz triangularizável $M \in M_n(\mathbb{R})$, existe uma matriz diagonalizável M_ε tal que $\|M - M_\varepsilon\| < \varepsilon$, com respeito a topolologia usual de $\mathbb{R}^{2n}$ em $M_n(\mathbb{R})$ (Observação 2.2.1). De fato, como M é triangularizável, existe uma matriz invertível $Q \in GL_n(\mathbb{R})$ tal que $M = Q^{-1}TQ$, em que T é uma matriz triangular. Seja $m = min|\lambda_i - \lambda_j|$ o mínimo sobre todos os autovalores distintos λ_i e λ_j de M. Tomemos $s < min\{\frac{\varepsilon}{n}, \frac{m}{n}\}$ e definimos a matriz diagonal $D = diag(s, 2s, \ldots, ns)$. Note que a matriz $M_\varepsilon = Q^{-1}(T + D)Q$ é diagonalizável, porque tem autovalores distintos para todo número real ε (Lema 1.2.3), e $\|M - M_\varepsilon\| < \varepsilon$. Portanto, a matriz triangularizável M pertence ao fecho das matrizes reais diagonalizáveis. No contexto das matrizes complexas, o Teorema da Decomposição de Schur (Teorema 1.2.5) assegura que toda matriz complexa é triangularizável, o que não ocorre com as matrizes reais.

O interior do conjunto das matrizes complexas diagonalizáveis é formado por matrizes complexas cujos autovalores são todos simples. O conjunto das matrizes não diagonalizáveis é fechado com interior vazio, portanto magro. Logo, “pequeno”, do ponto de vista da Topologia. Os espaços estudados nessa dissertação, com a da norma do máximo em $M_n(\mathbb{C})$ e a topologia usual de $\mathbb{R}^{2n^2}$, são completos e, portanto, são espaços topológicos de Baire. Todo espaço de Banach é de Baire. Em espaços de Baire, os complementares de conjuntos magros são abertos e densos. Portanto, o conjunto das matrizes complexas diagonalizáveis, além de denso, é aberto (*cf.* Lema 4.2.3).

As noções de conjunto magro (segundo a Topologia) e de conjunto de medida nula (segundo a Teoria da Medida) remetem à ideia de "pequeno". Entretanto, essas noções podem não coincidir. Por exemplo, seja $\mathbb{Q} \cap \,]0,1[\, = \{x_1, x_2, \ldots\}$, em que $\mathbb{Q}$ é o conjunto dos números racionais. Consideremos os conjuntos $G_j = \bigcup_{i=1}^{\infty} \left]x_i - \frac{1}{2^{i+j+1}}, x_i + \frac{1}{2^{i+j+1}}\right[$ e $G = \bigcap_{j=1}^{\infty} G_j$. O conjunto G tem medida nula; pois, para cada j, $G \subset G_j$ e, para cada número real $\varepsilon > 0$, existe um natural j tal que $\frac{1}{2^j} < \varepsilon$. Além disso, $\sum_{i=1}^{\infty} \mu\left(\left]x_i - \frac{1}{2^{i+j+1}}, x_i + \frac{1}{2^{i+j+1}}\right[\right) = \frac{1}{2^j}$, em que μ é medida de Lebesgue em $\mathbb{R}$. Os conjuntos G_j são abertos e densos, pois são uniões de intervalos abertos que contém os números racionais $x_1, x_2, \ldots$. Logo, os complementares de G_j são fechados com interior vazio, ou seja, G_j^c são conjuntos magros. Portanto, temos que $[0,1] = F \cup G$ e $\mu(\,[0,1]\,) = 1$, em que $F = \bigcup_{j=1}^{\infty} G_j^c \cup \{0,1\}$ é um conjunto magro e $G = \bigcap_{j=1}^{\infty} G_j$ é um conjunto de medida nula. Para um estudo dessas analogias entre espaços topológicos e de medida, o leitor pode consultar (OXTOBY, 2013).

O surgimento da topologia de Zariski permitiu uma interessante ligação entre a Álgebra, a Geometria e a Topologia, causando forte impacto sobre a Geometria Algébrica, sendo provedora de avanços significativos nesta corrente de pesquisa. A caracterização de conjuntos fechados por meio de equações algébricas é a brilhante contribuição dessa topologia. A topologia de Zariski possui algumas propriedades não muito usuais, porém interessantes: o espaço topológico $\mathbb{C}^n$ é compacto, ele não é Hausdorff se $n > 0$, e não é homeomorfo ao espaço $\mathbb{C}^m \times \mathbb{C}^{n-m}$ com a topologia produto das topologias de Zariski.

Como aplicação da topologia de Zariski, demonstraremos a seguir o clássico Teorema de Cayley-Hamilton da Álgebra Linear.

Teorema 5.0.1 (Teorema de Cayley-Hamilton) *Seja $M \in M_n(\mathbb{C})$ uma matriz, e seja P_M o polinômio característico de M. Então, $P_M(M) = 0$, em que 0 é a matriz nula de $M_n(\mathbb{C})$.*

Demonstração: O teorema é imediatamente válido se M for uma matriz diagonal. De fato, se $m_1, \cdots, m_n$ são as entradas da diagonal da matriz M, logo $P_M(\lambda) = (\lambda - m_1)\ldots(\lambda - m_n)$ e esse polinômio é anulado por M, pois $m_1, \cdots, m_n$ são suas raízes.

Se M é uma matriz diagonalizável, então podemos escrever $M = CDC^{-1}$ para uma matriz invertível $C \in GL_n(\mathbb{C})$ e uma matriz diagonal $D \in M_n(\mathbb{C})$. Como M e

D são matrizes semelhantes, temos que $P_M = P_D$. É necessário, também, verificar que $P_D(M) = 0$. Com efeito,

$$P_D(M) = P_M(CDC^{-1}) = CP_M(D)C^{-1} = 0.$$

A última igualdade decorre do primeiro parágrafo e $P_M(CDC^{-1}) = CP_M(D)C^{-1}$ é consequência da linearidade e da igualdade $(CDC^{-1})^k = CD^kC^{-1}$, válida para todo número inteiro $k \geq 1$.

Para provar o Teorema de Cayley-Hamilton, no caso de M ser uma matriz não diagonalizável, utilizaremos a topologia de Zariski.

Podemos interpretar cada matriz $M \in M_n(\mathbb{C})$ como um ponto de $\mathbb{C}^{n^2}$ e o discriminante de seu polinômio característico $\Delta(P_M)$ como um polinômio em n^2 indeterminadas, porque os coeficientes do polinômio característico P_M são polinômios nas entradas da matriz M e o discriminante $\Delta(P_M)$ é um polinômio nos coeficientes de P_M. Então, o complementar de $Z(\Delta(P_M))$ é aberto e denso na topologia de Zariski de $\mathbb{C}^{n^2}$, de acordo com o item 1 da Proposição 4.2.5.

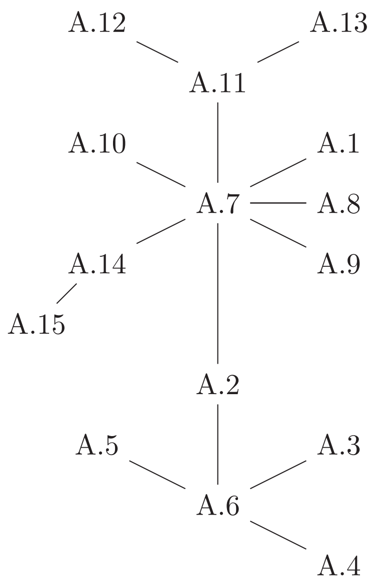
Considera-se a função $\varphi : \mathbb{C}^{n^2} \longrightarrow \mathbb{C}^{n^2}$ que associa a cada matriz M a matriz $P_M(M)$. Essa função φ é polinomial em cada componente do contradomínio. Logo, a função $\varphi : \mathbb{C}^{n^2} \longrightarrow \mathbb{C}^{n^2}$ é contínua, com respeito a Topologia de Zariski, conforme o item 3 da Proposição 4.2.5. Como $\{0\}$ é fechado na topologia de Zariski, de acordo com o item 2 da Proposição 4.2.5, e, pela continuidade da função $\varphi : \mathbb{C}^{n^2} \longrightarrow \mathbb{C}^{n^2}$, temos que $\varphi^{-1}(\{0\})$ é fechado.

Por outro lado, o complementar de $Z(\Delta(P_M))$ é formado somente por matrizes diagonalizáveis, segundo o Corolário 1.1.9 e o Lema 1.2.3. Ora, demonstramos que $M \in \varphi^{-1}(\{0\})$ para toda matriz M diagonalizável. Logo, $\varphi^{-1}(\{0\})$ contém o complementar de $Z(\Delta(P_M))$ que é um denso em $\mathbb{C}^{n^2}$. Portanto, $\varphi^{-1}(\{0\})$ é fechado e denso, o que só pode ocorrer se $\varphi^{-1}(\{0\}) = \mathbb{C}^{n^2}$. Como queríamos demonstrar. $\blacksquare$

Apêndices

APÊNDICE A – Medida e Integral de Lebesgue em $\mathbb{R}^n$

Neste apêndice são enunciados os principais resultados para uma construção da medida e da integral de Lebesgue em $\mathbb{R}^n$. O diagrama abaixo, que exibe as interdependências lógicas dos resultados deste apêndice, foi elaborado pelo autor, com uso do pacote TikZ do Latex. O leitor poderá consultar a obra (LANG, 2012), para uma demonstração destes resultados.



O propósito deste Apêndice é atender os anseios de leitores por uma leitura clara e sucinta da Medida e Integral de Lebesgue. Com esse propósito, apresentamos definições e fixamos notações mínimas para o entendimento dos Teoremas, Lemas, Corolários e Proposições expostos. Elaboramos, ainda, o diagrama acima para auxiliar o leitor no entendimento da sequência lógica e da interdependência desses resultados.

O Diagrama de Ishikawa, também conhecido como Diagrama de Causa e Efeito, serviu-nos de inspiração para a produção do diagrama deste Apêndice. Neste diagrama, os resultados são identificados de $A.1$ a $A.15$. Os resultados $A.6, A.2, A.7, A.11, A.12$ e $A.13$, que estão no centro e nas extremidades superiores do diagrama, são os fundamentais. Os $A.6, A.2$ e $A.7$ nos dão a medida e a integral de Lebesgue; os $A.11, A.12$ e $A.13$, a medida produto e a integral de Lebesgue em espaço produto. As relações entre os resultados são sinalizadas por traços, por exemplo, $A.7 - A.10$ significa que, na demonstração de $A.10$, faz-se uso de $A.7$. Os fatos que se encontram nas extremidades são demonstrados com o emprego dos que estão no centro.

Seja X um conjunto não-vazio. Uma família $\mathscr{M}$ de subconjuntos de X é chamada ***sigma-álgebra*** se possui as seguintes propriedades:

1. $\emptyset \in \mathscr{M}$;
2. se $A \in \mathscr{M}$, então $A^c \in \mathscr{M}$, em que A^c é o complementar de A em X;
3. se cada $A_n \in \mathscr{M}$, então $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{M}$.

Chamamos de ***álgebra*** uma família de subconjuntos de X que contém o conjunto vazio, a união, a interseção e a diferença de dois quaisquer conjuntos da família.

Uma ***medida positiva*** sobre uma sigma-álgebra $\mathscr{M}$ é uma função $\mu : \mathscr{M} \longrightarrow [0,\infty]$ tal que $\mu(\emptyset) = 0$ e, se $\{A_n\}_{n=1}^{\infty}$ é uma coleção enumerável de conjuntos de $\mathscr{M}$, disjuntos dois a dois, então $\mu\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = \sum\limits_{n=1}^{\infty} \mu(A_n)$.

Um terno ordenado $(X, \mathscr{M}, \mu)$, em que X é um conjunto, $\mathscr{M}$ é uma sigma-álgebra e μ é uma medida sobre $\mathscr{M}$, é denominado ***espaço de medida***.

A.1 (La 93, Exercise 7(a), p.173-174) *Sejam* $(X, \mathscr{M}, \mu)$ *um espaço de medida e* $\overline{\mathscr{M}}$ *constituído por todos os subconjuntos* Y *de* X *tais que, para algum conjunto* A *em* $\mathscr{M}$, *o subconjunto* $(Y - A) \cup (A - Y)$ *está contido em um conjunto de medida nula. Então:*

1. $\overline{\mathscr{M}}$ *é uma Sigma-Algebra;*
2. *Nessas condições, define-se* $\bar{\mu}(Y) = \mu(A)$. *Assim,* $\bar{\mu}$ *está bem definida e é uma medida em* $\overline{\mathscr{M}}$.

Dizemos que $(X, \overline{\mathscr{M}}, \bar{\mu})$ *é o* ***espaço de medida completo determinado por*** $(X, \mathscr{M}, \mu)$ *e* $\bar{\mu}$ *é o* ***completamento*** *de* μ.

Seja $(E, \|\cdot\|)$ um espaço vetorial normado (sobre o corpo dos números reais ou dos números complexos). Uma sequência (x_n) de E é chamada ***sequência de Cauchy*** se, para todo número real $\varepsilon > 0$ dado, existe $n_0 \in \mathbb{N}$ tal que $m, n > n_0$ implica $\|x_n - x_m\| < \varepsilon$. Se toda sequência de Cauchy de elementos de E converge para algum elemento de E, então, E é dito ***completo***, também chamado de ***espaço de Banach***.

Seja A um conjunto mensurável. Chama-se ***partição de*** A uma sequência $\{A_i\}$ $(i = 1, 2, \cdots, r)$ finita de conjuntos mensuráveis disjuntos tais que $A = \bigcup\limits_{i=1}^{r} A_i$. Uma função $f : X \longrightarrow E$ é chamada de ***função escada com relação a tal partição*** se f for igual a 0 fora de A e $f(A_i)$ tiver um elemento para cada i (ou seja, f é constante em

A_i). Definimos a ***integral de uma função escada*** por

$$\int_X f d\mu = \sum_{i=1}^{r} \mu(A_i) f(A_i).$$

Denotamos por $Es(\mu, E)$ o espaço vetorial das funções escadas. Definimos a ***L^1-seminorma***, no espaço vetorial das funções escadas, por

$$||f||_1 = \int_X |f| d\mu.$$

Uma sequência de funções escadas é designada ***L^1-Cauchy*** se for Cauchy relativamente a L^1-seminorma. Denominamos de ***L^1-nula*** toda sequência de funções (f_n) tais que, para todo número real $\varepsilon > 0$, existe um número natural n_0 tal que $n > n_0$ implica $||f_n||_1 < \varepsilon$.

Diz-se que uma sequência de funções $f_n : X \longrightarrow \mathbb{E}$ ***converge pontualmente*** para uma função $f : X \longrightarrow E$ se, para cada ponto $x \in X$ e cada número real $\varepsilon > 0$, existe um número natural $n_0 = n_0(x, \varepsilon)$ (que depende do ε dado e do ponto x considerado) tal que $n > n_0(x, \varepsilon)$ implica $||f_n - f||_1 < \varepsilon$.

Diremos que uma sequência de funções $f_n : X \longrightarrow \mathbb{E}$, não necessariamente funções escadas, ***converge uniformemente*** para uma função $f : X \longrightarrow E$ se, para cada número real $\varepsilon > 0$, for possível obter um número natural $n_0 = n_0(\varepsilon)$ (dependendo somente do ε dado) tal que $n > n_0$ implica $||f_n - f||_1 < \varepsilon$.

Uma sequência de funções (f_n) ***converge absoluta e uniformemente*** para uma função f se, e somente se, a sequência dos valores absolutos $(|f_n|)$ converge uniformemente para f.

Quando uma sequência de funções escadas $(f_n(x))$ converge para uma função $f(x)$, salvo nos pontos x de um conjunto de medida nula, dizemos que essa sequência ***converge em quase todo ponto***.

A.2 (La 93, Lemmas 3.1 e 3.2) *Sejam (f_n), (g_n) e (h_n) sequências de Cauchy de funções escadas de um espaço de medida X em um espaço de Banach E.*

1. *existe uma subsequência de (f_n) que converge ponto a ponto em quase toda parte, e satisfaz a propriedade adicional: dado $\varepsilon > 0$ existe um conjunto Z de medida $< \varepsilon$ tal que esta subsequência converge absoluta e uniformemente fora de Z;*

2. *Se (g_n) e (h_n) convergem em quase todo ponto para uma mesma função. Então, os seguintes limites existem e são iguais:*

$$\lim \int_X g_n = \lim \int_X h_n.$$

Além disso, (g_n) e (h_n) são equivalentes, isto é, $(g_n - h_n)$ é uma sequência L^1-nula.

Denotamos por $\mathscr{L}^1(\mu, E)$, ou simplesmente por $\mathscr{L}^1(\mu)$ ou $\mathscr{L}^1$, o espaço das funções $f : X \longrightarrow E$ tal que existe uma sequência de Cauchy de funções escadas (f_n) convergindo em quase todo ponto para $f : X \longrightarrow E$. Nessas condições, diz-se que a função $f : X \longrightarrow E$ é ***aproximada*** pela sequência $\{f_n\}$ e a sua integral é definida por

$$\int_X f d\mu = \lim \int_X f_n d\mu;$$

além disso, como $|f|$ é aproximada por $(|f_n|)$, estendemos a seminorma $\|\cdot\|_1$ a $\mathscr{L}^1(\mu, E)$ por

$$\|f\|_1 = \int_X |f| = \lim \int_X |f_n|.$$

A.3 (La 93, Thm 3.4) *Para cada sequência de Cauchy (f_n) de elementos em $\mathscr{L}^1(\mu, E)$, existe alguma função f em $\mathscr{L}^1(\mu, E)$ tal que dado um número real $\varepsilon > 0$, temos $\|f_n - f\|_1 < \varepsilon$ para n suficientemente grande.*

A.4 (La 93, Thm 4.2) *Sejam E e F espaços de Banach. Então, temos uma bijeção linear contínua cuja inversa também é linear e contínua*

$$\mathscr{L}^1(\mu, E \times F) \longrightarrow \mathscr{L}^1(\mu, E) \times \mathscr{L}^1(\mu, F).$$

Se $f : X \longrightarrow E \times F$ é uma função, com funções coordenadas $f = (g, h)$ em E e F, respectivamente, então $f \in \mathscr{L}^1$ se, e somente se, g e h estão em $\mathscr{L}^1$, e

$$\int_X f = \left(\int_X g, \int_X h\right).$$

Seja $\mathscr{A}$ uma subálgebra de $\mathscr{M}$, consistindo de conjuntos de medida finita. Um conjunto X é dito ***sigma-finito com relação à subálgebra*** $\mathscr{A}$ se X pode ser escrito como uma união contável de elementos de $\mathscr{A}$. Uma ***função escada*** f com relação a $\mathscr{A}$ é uma função que assume valor 0 fora de algum elemento A de $\mathscr{A}$, e existe uma partição $\{A_1, \cdots, A_r\}$ de A, consistindo de elementos de $\mathscr{A}$, tal que f é uma função escada com relação a essa partição. Denotemos por $Es(\mathscr{A})$ o espaço das funções escadas com relação a $\mathscr{A}$.

Seja $\mathscr{A}_A$ a álgebra de todos os elementos de $\mathscr{A}$ contidos em A. Denotamos por $Es(\mathscr{A}_A)$ o espaço das funções escadas com relação à álgebra $\mathscr{A}_A$.

A.5 (La 93, Lemma 6.1) *Seja A um elemento da subálgebra $\mathscr{A}$, consistindo de conjuntos de medida finita. Seja $\mathscr{N}_A$ a coleção de subconjuntos mensuráveis Y de $\mathscr{A}$ cuja função característica χ_Y possui a propriedade de dado $\varepsilon > 0$, existe uma função escada $\varphi \in Es(\mathscr{A}_A, \mathbb{R})$ satisfazendo*

$$\|\chi_Y - \varphi\|_1 < \varepsilon.$$

Então, $\mathscr{N}_A$ é uma sigma-álgebra em A.

A.6 (La 93, Thm 6.3 (Teorema da Aproximação)) *Sejam $\mathscr{A}$ uma subálgebra de $\mathscr{M}$, consistindo de conjuntos de medidas finitas, e $\mathscr{M}$ a menor sigma-álgebra que contém $\mathscr{A}$. Supondo que X é sigma-finito em relação a $\mathscr{A}$. Então, o espaço $Es(\mathscr{A})$ de funções escadas em relação a $\mathscr{A}$ é denso em $\mathscr{L}^1(\mu, E)$. Além disso, se $\{A_n\}$ é uma família contável de $\mathscr{A}$ cuja união é X, então, para todo $Y \in \mathscr{M}$ e todo n, vale o seguinte: para cada número real $\varepsilon > 0$, existe uma função escada $\phi \in Es(\mathscr{A}_{A_n}, \mathbb{R})$ satisfazendo*

$$\|\chi_{Y \cap A_n} - \phi\| < \varepsilon,$$

em que $\chi_{Y \cap A_n}$ é a função característica de $Y \cap A_n$.

Seja $\mathscr{N}$ uma sigma-álgebra num conjunto X. Uma ***medida exterior*** μ em $\mathscr{N}$ é uma função $\mu : \mathscr{N} \longrightarrow [0, \infty]$ que satisfaz as seguintes condições:

1. $\mu(\emptyset) = 0$;
2. se $A, B \in \mathscr{N}$ e $A \subseteq B$, então $\mu(A) \leq \mu(B)$;
3. se $\{A_n\}$ é uma sequência de elementos de $\mathscr{N}$, então

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

A.7 (La 93, Thm 7.1 (Hahn)) *Seja μ uma medida positiva em uma álgebra $\mathscr{A}$ em X, e suponha que X pode ser expresso como uma união enumerável de conjuntos de $\mathscr{A}$. Então, μ pode ser estendida para uma medida positiva sobre a menor sigma-álgebra $\mathscr{M}$ contendo $\mathscr{A}$, de forma que para $Y \in \mathscr{M}$,*

$$\mu(Y) = \inf \sum_{n=1}^{\infty} \mu(A_n),$$

o inf *sendo tomado em todas as seqüências (A_n) em $\mathscr{A}$ cuja união contém Y. Se X puder ser expresso como uma união contável de conjuntos de medida finita em $\mathscr{A}$, então existe uma extensão única de μ para uma medida positiva sobre $\mathscr{M}$.*

A.8 (La 93, Lemma 7.2) *Seja μ uma medida positiva sobre uma álgebra $\mathscr{A}$ em X, e suponha que X pode ser expresso como uma união enumerável de conjuntos em $\mathscr{A}$. Na sigma-álgebra de todos os subconjuntos de X, defina*

$$\mu^*(Y) = \inf \sum_{n=1}^{\infty} \mu(A_n),$$

o inf tomando todas as seqüências $\{A_n\}$ de elementos de $\mathscr{A}$ cuja união contém Y. Então μ^* é uma medida exterior que estende μ.

A.9 (La 93, Lemma 7.3) *Seja μ uma medida exterior definida sobre todos os subconjuntos de X. Seja $\mathscr{L}$ a coleção de todos os subconjuntos A de X tais que, para todo subconjunto Z de X, vale a igualdade*

$$\mu(Z) = \mu(Z \cap A) + \mu(Z \cap A^c),$$

em que A^c representa o complementar do conjunto A. Então, $\mathscr{L}$ é uma sigma-álgebra, e μ é uma medida positiva em $\mathscr{L}$.

A.10 (La 93, Cor 7.4) *Sob as hipóteses do Teorema de Aproximação, um subconjunto Z de X tem μ^*-medida nula se, e somente se, dado um número real $\varepsilon > 0$, existe uma coleção contável $\{A_n\}$ em $\mathscr{A}$ cuja união cobre Z e tal que*

$$\sum_{n=1}^{\infty} \mu(A_n) < \varepsilon.$$

A.11 (La 93, Thm 8.2) *Sejam $(X, \mathscr{M}, \mu)$ e $(Y, \mathscr{N}, \nu)$ espaços de medida sigma-finitos e $\mathscr{M} \otimes \mathscr{N}$ a menor sigma-álgebra que contém a família $\mathscr{M} \times \mathscr{N}$. Existe uma única medida positiva $(\mu \otimes \nu) : \mathscr{M} \otimes \mathscr{N} \longrightarrow [0, \infty]$ tal que para todos os conjuntos, A e B de medida finita em $\mathscr{M}$ e $\mathscr{N}$, respectivamente, temos*

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B).$$

A.12 (La 93, Exercise 7(b), p.173-174) *Sejam $(X_i, \mathscr{M}_i, \mu_i)$, com $i = 1, 2$ e 3, espaços de medida. Então,*

$$(\mathscr{M}_1 \otimes \mathscr{M}_2) \otimes \mathscr{M}_3 = \mathscr{M}_1 \otimes (\mathscr{M}_2 \otimes \mathscr{M}_3)$$

$$(\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3).$$

Se $(X, \mathscr{M}, \mu)$ e $(Y, \mathscr{N}, \nu)$ são espaços de medida; então, em termos das notações da Proposição A.1, valem as igualdades

$$\overline{\overline{\mathscr{M}} \otimes \overline{\mathscr{N}}} = \overline{\mathscr{M} \otimes \mathscr{N}} \;\text{ e }\; \overline{\overline{\mu} \otimes \overline{\nu}} = \overline{\mu \otimes \nu}.$$

A.13 (La 93, Thm 8.4(Teorema de Fubini)) *Seja $f \in \mathscr{L}^1(\mu \otimes \nu)$. Então, para quase todo x, a função f_x, definida sobre Y por $f_x(y) = f(x, y)$, está em $\mathscr{L}^1(\nu)$, a função dada por*

$$x \mapsto \int_Y f_x \, d\nu$$

para quase todo x (e definido arbitrariamente para outro x) está em $\mathscr{L}^1(\mu)$; e temos

$$\int_{X \times Y} f \, d(\mu \otimes \nu) = \int_X \left(\int_Y f_x \, d\nu \right) d\mu(x).$$

A.14 (La 93, Thm 9.1) *Seja* (f_n) *uma sequência de funções* ≥ 0 *em um intervalo fechado e limitado* I*, decrescendo monotonicamente para* 0. *Suponha que cada* f_n *é uma função escada com respeito a intervalos. Então, a sequência de integrais (simples e ordinárias)*

$$\int_I f_n(x)dx$$

decresce para 0.

A.15 (La 93, Cor 9.2) *A função de comprimento dos intervalos estende-se unicamente a uma medida na álgebra que consiste de uniões disjuntas de intervalos limitados.*

APÊNDICE B – Ilustrações da Demonstração do Lema 3.2.4

> Este Apêndice tem como escopo ilustrar o uso do Teorema de Fubini, na demonstração do Lema 3.2.4. Nesse mister, faz-se a exposição de exemplos concretos de polinômios em duas e três indeterminadas. Os gráficos deste Apêndice foram elaborados pelo autor, com o uso do aplicativo Geogebra.

Por questões de didática, estudaremos primeiro um exemplo de polinômio em duas indeterminadas. Depois, examinaremos um exemplo de polinômio em três indeterminadas que, embora um pouco mais complexo do que o de duas indeterminadas, fornece uma melhor interpretação álgebrica e geométrica do Teorema de Fubini.

B.1 *Consideremos o polinômio* $P(X,Y) = X^3 - XY^2$.

Em termos da notação da demonstração do Lema 3.2.4, temos

$$P(X,Y) = Q_0(X) + Q_1(X)Y + Q_2(X)Y^2,$$

em que $Q_0(X) = X^3$, $Q_1(X) = 0$, $Q_2(X) = -X$. Reescrevendo o polinômio, obtemos $P(X,Y) = X(X+Y)(X-Y)$.

Mostremos que o conjunto dos zeros $Z(P)$ do polinômio $P(X,Y) = X^3 - XY^2$ tem medida nula (*i.e.*, $\mu_2(Z(P)) = 0$), conforme a demonstração do Lema 3.2.4.

De acordo com a Observação 3.2.3, tem-se

$$\mu_2(Z(P)) = \int_{\mathbb{R}^2} \chi_{Z(P)}(x,y)\,dx\,dy,$$

em que $\chi_{Z(P)} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ é a função característica de $Z(P)$, que associa o valor 1 aos pontos do conjunto $Z(P)$ e o valor 0 aos demais pontos de $\mathbb{R}^2$, e μ_2 denota a medida de Lebesgue em $\mathbb{R}^2$.

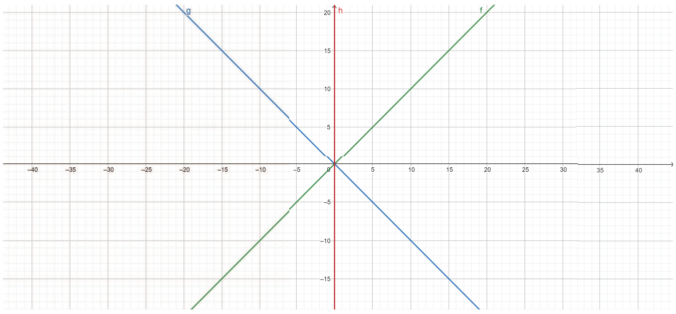


Figura 1 – Os zeros de $P(X, Y) = X^3 - XY^2$

A Figura 1 ilustra que o conjunto dos zeros $Z(P)$ é fechado, com respeito a topologia usual $\mathbb{R}^2$. Logo, o conjunto dos zeros de um polinômio é boreliano e, consequentemente, um conjunto mensurável a Lebesgue de $\mathbb{R}^2$, conforme o Lema 3.2.2. Do exposto, temos que a função $\chi_{Z(P)} : \mathbb{R}^2 \longrightarrow \mathbb{R}$ é mensurável. Então, pelo Teorema de Fubini, obtemos

$$\int_{\mathbb{R}^2} \chi_{Z(P)}(x, y)\, dx\, dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x, y)\, dy \right) dx.$$

A Figura 1 nos mostra que a função seção $f(x) = \int_{\mathbb{R}} \chi_{Z(P)}(x, y)\, dy$ está definida em quase todo ponto de $\mathbb{R}$. De fato, o conjunto $A = \{x \in \mathbb{R} : P(x, y) = x(x + y)(x - y) = 0,\ \text{para todo } y \in \mathbb{R}\}$ é unitário: $\{0\}$. Portanto, o conjunto A tem medida nula.

Notemos que, em cada ponto x de A, a função $f(x) = \int_{\mathbb{R}} \chi_{Z(P)}(x, y)\, dy$ não está definida, pois nesses pontos $\int_{\mathbb{R}} \chi_{Z(P)}(x, y)\, dy = \infty$.

Ademais, para cada ponto fixado x do complementar do conjunto A, por exemplo $x = 2$, o polinômio $P(2, Y) = 8 - 2Y^2$ em uma indeterminada Y não é identicamente nulo. Donde, o conjunto de suas raízes é finito, a saber, $\{-2, 2\}$ e, consequentemente, a função $\chi_{Z(P)}(x, \cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ assume o valor 1, num conjunto de medida nula, e 0 nos demais pontos de $\mathbb{R}$. Então, temos que $f(x) = \int_{\mathbb{R}} \chi_{Z(P)}(x, y)\, dy = 0$ em quase todo ponto de $\mathbb{R}$.

Logo,

$$\begin{aligned} \mu_2(Z(P)) &= \int_{\mathbb{R}^2} \chi_{Z(P)}(x, y)\ dx\ dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x, y)\ dy \right) dx \\ &= \int_{\mathbb{R}} f(x)\ dx = 0 \end{aligned}$$

Constata-se que há três possibilidades para que o polinômio $P(X, Y) = X(X + Y)(X - Y)$ seja zerado. A primeira é de um ponto de $\mathbb{R}^2$ pertencer a reta de equação $Y = X$. A segunda é de algum ponto de $\mathbb{R}^2$ pertencer a reta de equação

$Y = -X$. A terceira é de um ponto de $\mathbb{R}^2$ pertencer a reta de equação $X = 0$ e, nessa possibilidade, Y pode assumir qualquer valor.

A figura 1 corrobora com a conclusão $\mu_2(Z(P)) = 0$; porque, em termos geométricos, o conjunto $Z(P)$ é constituído por três retas, que são claramente conjuntos de medida nula em $\mathbb{R}^2$.

O próximo exemplo, por envolver figuras geométricas espaciais, esclarecerá melhor o uso do Teorema de Fubini na demonstração do Lema 3.2.4.

B.2 *Consideremos o polinômio* $P(X,Y,Z) = X^2Z^3 + Y^2Z^3 - Z^3 - X^2 - Y^2 + 1$.

Em termos da notação da demonstração do Lema 3.2.4, temos

$$P(X,Y,Z) = Q_0(X,Y) + Q_1(X,Y)Z + Q_2(X,Y)Z^2 + Q_3(X,Y)Z^3,$$

em que $Q_0(X,Y) = -X^2-Y^2+1$, $Q_1(X,Y) = 0$, $Q_2(X,Y) = 0$ e $Q_3(X,Y) = X^2+Y^2-1$.

De acordo com a Observação 3.2.3, temos que

$$\mu_3(Z(P)) = \int_{\mathbb{R}^3} \chi_{Z(P)}(x,y,z)\, dx\, dy\, dz,$$

em que $\chi_{Z(P)} : \mathbb{R}^3 \longrightarrow \mathbb{R}$ é a função característica de $Z(P)$, que associa o valor 1 aos pontos do conjunto $Z(P)$ e o valor 0 aos demais pontos de $\mathbb{R}^3$, e μ_3 denota a medida de Lebesgue em $\mathbb{R}^3$.

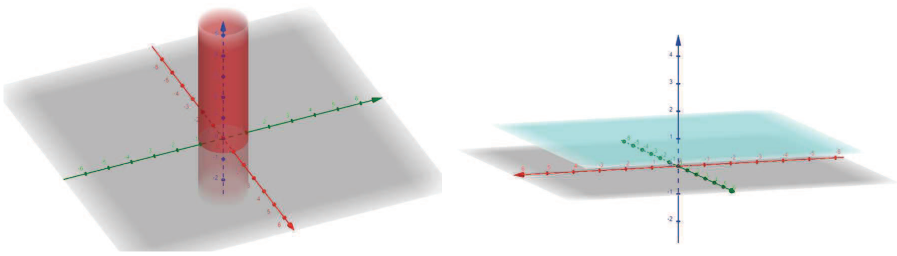


Figura 2 – Os zeros de $P(X,Y,Z) = (-X^2 - Y^2 + 1) + (X^2 + Y^2 - 1)Z^3$

A Figura 2 ilustra que o conjunto dos zeros $Z(P)$ é fechado, com respeito a topologia usual $\mathbb{R}^3$. Logo, o conjunto dos zeros de um polinômio é boreliano e, consequentemente, um conjunto mensurável a Lebesgue de $\mathbb{R}^3$, conforme o Lema 3.2.2. Do exposto, temos que a função $\chi_{Z(P)} : \mathbb{R}^3 \longrightarrow \mathbb{R}$ é mensurável. Então, pelo Teorema de Fubini, obtemos

$$\int_{\mathbb{R}^3} \chi_{Z(P)}(x,y,z)\, dx\, dy\, dz = \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\, dz \right) dx\, dy.$$

A Figura 2 nos mostra que a função seção $f(x,y) = \int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\, dz$ está definida em quase todo ponto de $\mathbb{R}^2$. De fato, o conjunto

$A = \{(x,y) \in \mathbb{R}^2 : P(x,y,z) = (-x^2 - y^2 + 1) + (x^2 + y^2 - 1)z^3 = 0, \text{ para todo } z \in \mathbb{R}\}$ é, em termos geométricos, uma circunferência. Portanto, o conjunto A tem medida nula em $\mathbb{R}^2$.

Notemos que, em cada ponto (x,y) de A, a função $f(x,y) = \int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\,dz$ não está definida, porque nesses pontos $\int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\,dz = \infty$.

Ademais, para cada ponto fixado (x,y) do complementar do conjunto A, por exemplo o ponto $(1,1)$, o polinômio $P(1,1,Z) = -1 + Z^3$ em uma indeterminada Z não é identicamente nulo. Donde, o conjunto de suas raízes é finito, a saber, $\{1\}$ e, consequentemente, a função $\chi_{Z(P)}(x,y,\cdot) : \mathbb{R} \longrightarrow \mathbb{R}$ assume o valor 1, num conjunto de medida nula, e 0 nos demais pontos de $\mathbb{R}$. Então, temos que $f(x,y) = \int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\,dz = 0$ em quase todo ponto de $\mathbb{R}$.

Logo,

$$\begin{aligned}
\mu_3(Z(P)) &= \int_{\mathbb{R}^3} \chi_{Z(P)}(x,y,z)\,dx\,dy\,dz \\
&= \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \chi_{Z(P)}(x,y,z)\,dz \right) dx\,dy \\
&= \int_{\mathbb{R}^2} f(x,y)\,dx\,dy = 0
\end{aligned}$$

Nota-se que há três possibilidades para que o polinômio $P(X,Y,Z) = (-X^2 - Y^2 + 1) + (X^2 + Y^2 - 1)Z^3$ seja zerado. A primeira é de a projeção ortogonal $A' = (x,y)$ sobre o plano de equação $z = 0$, de um ponto $A = (x,y,z) \in \mathbb{R}^3$, pertencer a circunferência de centro na origem $(0,0)$ e de raio 1, ou seja, é solução da equação $X^2 + Y^2 - 1 = 0$. Nessa situação, não depende dos valores assumidos por Z. A segunda é de um ponto $B = (x,y,z) \in \mathbb{R}^3$ petencer ao plano de equação $Z = 1$ e também, nessa possibilidade, X e Y podem assumir qualquer valor. A última possibilidade é, quando fixados X e Y, e Z é a raiz do polinômio resultante. Por exemplo, fixando $X = 1$ e $Y = 2$, Z tem que ser raiz do polinômio resultante $P(Z) = 3Z^3 - 4$, neste caso, $Z = \sqrt[3]{\frac{4}{3}}$.

Observamos, na figura 2, que o conjunto dos zeros do polinômio $P(X,Y,Z) = X^2Z^3 + Y^2Z^3 - Z^3 - X^2 - Y^2 + 1$ é constituído por um cilindro, quando as projeções dos pontos de $\mathbb{R}^3$ sobre o plano de equação $Z = 0$ pertencem a circunferência de equação $X^2 + Y^2 - 1 = 0$ (Z assume qualquer valor, nessa situação); por um plano, na hipótese de os pontos de $\mathbb{R}^3$ serem soluções da equação $Z = 1$ (neste caso, X e Y assumem qualquer valor); e por pontos isolados de $\mathbb{R}^3$, que são raízes de polinômios em uma indeterminada. Isso ratifica que $\mu_3(Z(P)) = 0$.

Referências

ATIYAH, M. *Macdonald., IG, Introduction to commutative algebra*. Universidade of Oxford: Addison-Wesley Publishing Co, 1969.

GARCIA, A.; LEQUAIN, Y. *Elementos de álgebra*. Rio de Janeiro: SBM, 2013.

HETZEL, A. J.; LIEW, J. S.; MORRISON, K. E. *The probability that a matrix of integers is diagonalizable*. Reino Unido: Taylor & Francis, 2007. v. 114. 491–499 p.

HIRSCH, M. W.; SMALE, S. *Differential equations, dynamical systems, and linear algebra*. Cambridge: Academic press, 1974. v. 60.

ISNARD, C. *Introdução à medida e integração*. Rio de Janeiro: SBM, 2013.

KOSTRIKIN, A. I.; MANIN, Y. I. *Linear algebra and geometry*. Amsterdã: CRC Press, 1989.

LANG, S. *Real and functional analysis*. New York: Springer Science & Business Media, 2012. v. 142.

LIMA, E. L. *Elementos de topologia geral*. Rio de Janeiro: SBM, 2009.

LIMA, E. L. *Álgebra linear*. Rio de Janeiro: SBM, 2014.

MUNKRES, J. R. *Topology. 2nd*. Londres: Pearson Education, 2000.

NETO, J. A. B. *Diagonalização de matrizes 2X2 e reconhecimento de cônicas*. Rio de Janeiro: Dissertação de mestrado do PROFMAT, 2013.

OXTOBY, J. C. *Measure and category: A survey of the analogies between topological and measure spaces*. New York: Springer Science & Business Media, 2013. v. 2.

SILVA, R. G. et al. *Reconhecimento de quádricas via diagonalização de matrizes*. Rio de Janeiro: Dissertação de mestrado do PROFMAT, 2016.

ZAPATA, T. A. D. *Várias variáveis reais: em quadrinhos.* 2014. Notas de Aula.

ZAPATA, T. A. D. *Teoria dos números.* 2019. Notas de Aula.