

Universidade Federal de Uberlândia
Faculdade de Matemática

Anais da X Mostra IC

24 a 26 de maio de 2021



Uberlândia
2021

Copyright© 2016 da Faculdade de Matemática (UFU)
Todos os direitos reservados

Sumário

Análise multivariada em relação à empregabilidade da população no Brasil	5
<i>Marcos Vinícius Vieira de Souza, Priscila Neves Faria</i>	
Álgebra Linear aplicada a circuitos elétricos	11
<i>Ariely da Silva Camargo, Ana Paula Tremura Galves</i>	
Álgebra Linear aplicada aos Jogos de estratégia	17
<i>Beatriz Harumi Nishiyama, Lígia Laís Fêmina</i>	
Autovalores e Autovetores de Matrizes Toeplitz Tridiagonais Simétricas	23
<i>João Vitor Caixeta Alcarria, Santos Alberto Enriquez-Remigio</i>	
Polinômios, Raios Laser e Tartarugas	29
<i>Roney Andrade, Germano Abud</i>	
Introdução às bases de Gröbner e uma aplicação	35
<i>Selma Alves de Araújo, Cícero Carvalho</i>	
Identificação de Parâmetros Envolvendo a Equação da Difusão Bidimensional com Termo Fonte via Métodos Adjuntos	41
<i>Thaís Barbosa Caetano Souza, Alessandro Alves Santana</i>	
O ponto e seus grupos de homologia	47
<i>Alef Alves Fidelis, Francielle Rodrigues de Castro Coelho</i>	
Uma representação do número π em Frações Contínuas	53
<i>Aloisio da Silva Teixeira, Lígia Laís Fêmina</i>	
O Teorema da Aplicação Aberta e a essencialidade de suas hipóteses	58
<i>Ana Laura Mendonça Marangoni, Fábio José Bertoloto</i>	
Completamento de Espaços Métricos	64
<i>Ana Lídia da Silva Lício, Elisa Regina dos Santos</i>	
O toro e a esfera são homeomorfos?	70
<i>Anderson Martins Rocha do Prado, Lígia Laís Fêmina</i>	
Diagonalização de Matrizes Simétricas na Identificação de Cônicas	75
<i>Bruno Henrique Viana de Moraes, Catiana Casonatto</i>	
O problema da Trissecção do Ângulo	81

Cinira Aparecida de Oliveira, Adriana Rodrigues da Silva

Controlabilidade em Sistemas Dinâmicos Lineares	87
<i>Dhara Cristina de Freitas Lago Grande, Luciana Aparecida Alves</i>	
O Inverso de Dirichlet e a Fórmula de Inversão de Möbius	92
<i>Fernanda de Andrade Flor, Victor Gonzalo Lopez Neumann</i>	
Núcleos e centro de um quasegrupo	98
<i>Gabriel Simão Mucci, Dylene Agda Souza de Barros</i>	
Uma definição de caos e sua redução	104
<i>Guilherme Daniel Nogueira, Jean Venato Santos</i>	
Polinômios sobre anéis não comutativos	110
<i>Júlia Bernardes Coelho, Dylene Agda Souza de Barros</i>	
Operadores Ortogonais	116
<i>Laura Barbosa Goulart, Adriana Rodrigues da Silva</i>	
Dinâmica Unidimensional	122
<i>Leonardo Henrique Soria, Marcus Augusto Bronzi</i>	
O Sistema de Lorenz: Uma Alusão ao Caos em Dinâmica Contínua	128
<i>Matheus Gonçalves Silva de Castro, Jean Venato Santos</i>	
Construindo a distância de Hausdorff	134
<i>Paulo César da Silva Andreucci, Ana Paula Tremura Galves</i>	
Perturbações de sistemas lineares bidimensionais $x' = Ax$	140
<i>Rodrigo Lima Almeida, Hernán Roberto Montúfar López</i>	
Propriedades notáveis de ℓ_p, $1 \leq p \leq \infty$	146
<i>Stefânia Carvalho de Sousa, Fábio José Bertoloto</i>	
De Grupos até Subgrupos Normais: Qual a Construção?	152
<i>Victor Cruz Borges, Cícero Fernandes de Carvalho</i>	
O determinante de um nó	158
<i>Victor Patrick Sena Barbosa Lima, Taciana Oliveira Souza</i>	
Classificação dos grupos de ordem ≤ 8	164
<i>Victor Rodrigues Silva, Adriana Rodrigues da Silva</i>	
Detecção de danos em vigas de alumínio via modelagem fuzzy	170
<i>Fernando Augusto Freitas, Rosana Sueli da Motta Jafelice, José Waldemar da Silva, Carlos Alberto Gallo, Julio Endress Ramos</i>	
Predição do modelo de crescimento econômico de Solow com parâmetro fuzzy	177
<i>Paulo Victor Santana, Rosana Sueli da Motta Jafelice, Julio Fernando Costa</i>	



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Análise multivariada em relação à empregabilidade da população no Brasil

Marcos Vinícius Vieira de Souza

UFU, Departamento de Matemática, Uberlândia, MG, Brasil

vsmarcos@ufu.br

Prof^a Dr^a Priscila Neves Faria

UFU, Departamento de Matemática, Uberlândia, MG, Brasil

priscilaneves@ufu.br

Resumo

Palavras-chave

Empregabilidade.
Desigualdade.
Análise multivariada.
Agrupamentos.
Correlação cofenética.

A ação do Estado, por meio da mobilização do emprego público e das transferências de renda, constituiu importante elemento para compreensão da desigualdade de renda no Brasil. O mapeamento e a análise da empregabilidade ajudam na compreensão da decisiva atuação do Estado na geração e distribuição de renda dentro do país. Assim, o presente estudo visa analisar com técnicas estatísticas multivariadas dados econômicos no Brasil nos anos de 2015 a 2019. Dentre outros resultados, foi obtido que os grupos formados forneceram indícios que o estado de São Paulo se isola dos demais na maioria dos temas econômicos considerados neste estudo, bem como o destaque do rendimento médio do Distrito Federal em relação aos outros estados.

1 Introdução

Segundo Araújo (1997) economia regional brasileira é bastante heterogênea, ou mesmo fragmentada, existindo áreas de grande dinamismo convivendo com numerosas regiões caracterizadas pela pobreza, estagnação e retrocesso. O desenvolvimento econômico é importante de ser analisado pois é ele, em grande medida, o que define os padrões salariais. Conforme Borjas (1996) e Ehrenberg e Smith (1996), as regiões mais dinâmicas economicamente irão atrair as maiores empresas aumentando a demanda por mão-de-obra qualificada, o que pode resultar em um incremento da média salarial da região e, conseqüentemente, numa melhoria da qualidade da mão-de-obra disponível.

Reis e Barros (1991) estudaram a relação entre a distribuição da educação e o nível de desigualdade salarial no Brasil buscando verificar qual o papel das diferenças regionais de educação na desigualdade salarial. Os autores concluíram que a educação explica cerca de 50% da desigualdade salarial no Brasil metropolitano e que esta desigualdade é explicada não pela distribuição da educação no país mas sim pelas diferenças na inclinação da equação de rendimentos regionais.

Analisando os dados econômicos no período de 2015 a 2019, foi possível perceber um composto de resultados que avaliam financeiramente a capacidade de cada Unidade Federativa (UF) no Brasil. Um estudo a respeito do estado econômico das UF permite compreender a situação do Brasil como um todo e onde se concentra recursos e ferramentas necessários para o desenvolvimento da nação. Assim, o presente estudo teve como objetivo promover a informação referente à capacidade econômica das UF brasileiras no que diz respeito às pessoas ocupadas, pessoas desocupadas e rendimento médio real utilizando a metodologia Cluster Analysis que compõe parte das metodologias da estatística multivariada.

2 Metodologia

2.1 Banco de dados

Todas as informações do presente estudo foram obtidas diretamente do site do IPEA, Instituto de Pesquisa Econômica Aplicada, que é uma fundação pública federal vinculada ao Ministério da Economia. Os dados contemplam o período de 2015 a 2019, dados estes compostos pelos estados federativos do Brasil e as covariáveis os fatores econômicos destes, a saber: número de pessoas ocupadas e desocupadas, bem como o rendimento médio dos trabalhadores.

Foram consideradas como pessoas ocupadas aquelas que, na semana de referência da entrevista da pesquisa, trabalharam pelo menos uma hora completa em trabalho remunerado ou em trabalho sem remuneração direta, em ajuda à atividade econômica de membro do domicílio ou, ainda, as pessoas que tinham trabalho remunerado do qual estavam temporariamente afastadas nessa semana. Para pessoas desocupadas, foram consideradas as pessoas sem trabalho (que geram rendimentos para o domicílio) na semana da entrevista. O rendimento médio foi captado utilizando os dados fornecidos pelo preenchimento de um questionário daqueles que foram classificados como pessoas ocupadas. A

pesquisa é realizada por meio de uma amostra probabilística de domicílios, extraída de uma amostra mestra de setores censitários, de forma a garantir a representatividade dos resultados para os diversos níveis geográficos definidos.

Os dados foram analisados estatisticamente e obtidos os grupos (clusters) em que os estados que compunham um mesmo cluster seriam considerados similares em relação aos fatores analisados. Para definir tal semelhança foram necessárias avaliar medidas de dissimilaridade, métodos hierárquicos de agrupamentos, dentre outras medidas que compõem a análise multivariada de dados. As análises preliminares dos dados foram realizadas no software Excel e as posteriores a esta analisadas por meio dos softwares Past, R e Python.

2.2 Medidas de distância (dissimilaridade)

O termo dissimilaridade surgiu em função da relação da distância entre dois pontos P e Q, definida como $d(P,Q)$, pois, à medida que ela cresce, diz-se que a divergência entre os pontos (unidades amostrais) P e Q aumenta, ou seja, tornam-se cada vez mais dissimilares. Os valores de distâncias são geralmente obtidos a partir de informações de “n” unidades amostrais, mensurados em relação a “p” caracteres (variáveis).

De acordo com Cruz et al. (2011), é necessário especificar um coeficiente de semelhança que indique a proximidade entre os indivíduos sendo importante considerar, em todos os casos semelhantes a este, a natureza da variável (discreta, contínua, binária) e a escala de medida (nominal, ordinal, real ou razão).

A "distância euclidiana" preconiza que a distância entre duas observações (i e j) corresponda à raiz quadrada da soma dos quadrados das diferenças entre os pares de observações (i e j), para todas as p variáveis (FÁVERO et al., 2009), conforme a fórmula $d_{ij} = \sqrt{\sum_j (Y_{ij} - Y'_{ij})^2}$, considerando Y_{ij} a observação no i-ésimo indivíduo para a j-ésima variável.

Na definição das variáveis, espera-se que elas apresentem contribuição equivalente na análise de agrupamento, de modo que a distância entre as amostras não seja alterada com a adoção de características com unidades de medidas distintas, num mesmo conjunto de dados, de forma que as variáveis admitidas apresentem poder discriminatório semelhante, não baseado na amplitude de seus valores. Caso uma determinada característica apresente uma maior amplitude em seus valores, em comparação às demais, ela certamente terá um maior peso na análise.

Conforme descrito por Johnson e Wichern (1998), a Distância Euclidiana é insatisfatória para muitas situações estatísticas. Isso ocorre devido à contribuição de cada coordenada ter o mesmo peso para o cálculo da distância. Segundo Fávero et al. (2009, p. 198), "A maior parte das medidas de distância sofre influência das diferentes escalas ou magnitudes das variáveis de similaridade". Vejamos o exemplo: suponha que uma pesquisa caracteriza os indivíduos utilizando duas variáveis: altura e peso.

A normalização transforma os dados das variáveis originais em escores padrão (também denomina-

dos de escores Z), de maneira a apresentar média 0 (zero) e desvio padrão 1 (um), conforme a fórmula $Z_j = \frac{Y_j - \bar{Y}}{\hat{\sigma}_j}$, em que $\hat{\sigma}_j$ é estimativa do desvio-padrão e \bar{Y} é a estimativa da média, ambos associados à j -ésima variável.

A chama Distância Euclidiana Quadrada consiste exatamente na Distância Euclidiana, porém sem extração da raiz quadrada: "A Distância Euclidiana Quadrada tem a vantagem de que não é necessário calcular a raiz quadrada, o que acelera o tempo de computação, e é a distância recomendada para o método de agrupamento Ward" (HAIR JR. et al., 2005, p. 394).

A Distância de Mahalanobis é muito utilizada uma vez que considera as diferenças de variação e a presença de correlação, ou seja, é uma distância que depende das variâncias e das covariâncias amostrais. Esta distância tem sua expressão definida por $D_i^2 = (x_i - \mu)' \Sigma^{-1} (x_i - \mu)$, em que x_i é um vetor de observações da população avaliada, μ_i é o centroide da população avaliada, Σ^{-1} é a inversa da matriz de covariância, sendo também outra medida de distância muito utilizada no caso em que as variáveis analisadas sejam correlacionadas entre si.

Atualmente, diversas medidas de dissimilaridade são propostas na literatura, principalmente devido ao grande desenvolvimento e utilização das técnicas multivariadas (Khattree e Naik, 2000).

Utilizando todas as variáveis disponíveis e, depois de aplicado algum método hierárquico de agrupamento à matriz de distâncias, os grupos (clusters) são representados por um gráfico bidimensional denominado dendrograma (diagrama bidimensional em forma de árvore). Nele estão dispostos linhas ligadas segundo os níveis de similaridade, que agrupará pares de indivíduos ou de variáveis.

3 Resultados e discussão

Foram utilizadas a Distância Euclidiana com o método UPGMA como medida de distância e método de ligação para se analisar os agrupamentos para pessoas ocupadas e pessoas desocupadas.

Para pessoas ocupadas, utilizando a Estatística Pseudo T^2 , o número ótimo de grupos foi 6, resultando no dendrograma apresentado a seguir:

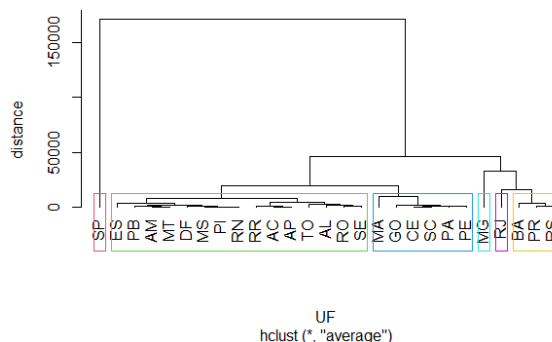


Figura 1: Dendrograma referente à aplicação da distância euclidiana em conjunto ao método UPGMA para as 27 UF do Brasil considerando as pessoas ocupadas no período de 2015 a 2019.

O Coeficiente de Variação obtido foi alto em todos os anos, e é possível observar um crescimento constante dessa medida, começando em 126,89% no ano de 2015 e terminando em 130,42% em 2019.

Ao se analisar os grupos formados para pessoas ocupadas é possível traçar um paralelo entre os grupos e a quantidade de habitantes das UF. De acordo com uma estimativa do IBGE (2020), as 12 UF brasileiras mais populosas são, respectivamente, SP, MG, RJ, BA, PR, RS, PE, CE, PA, SC, GO e MA. É possível notar que os três primeiros formaram, cada um, um grupo unitário. BA, PR e RS, os próximos na lista se agruparam. PE, CE, PA, SC, GO e MA formaram outro grupo, e as UF restantes, que são aquelas que não estão nessa lista, formaram o último grupo.

Os passos seguidos na análise de pessoas ocupadas também foram usados no estudo acerca das pessoas desocupadas. De acordo com as Estatísticas Pseudo F e Pseudo T^2 o número ótimo de grupos para pessoas desocupadas é 4.

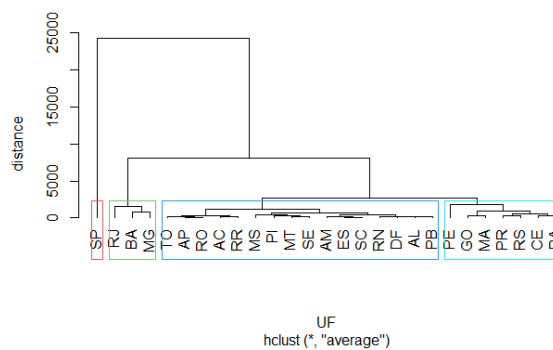


Figura 2: Dendrograma referente à aplicação da distância euclidiana em conjunto ao método UPGMA para as 27 UF do Brasil considerando as pessoas desocupadas no período de 2015 a 2019.

Ao analisar o dendrograma é possível concluir que assim como para pessoas ocupadas, os agrupamentos de pessoas desocupadas segue a tendência da quantidade de habitantes.

Durante o estudo foi possível notar que nos anos de 2015 e 2016 a quantidade de pessoas desocupadas é crescente, e a partir do ano de 2017 isso se inverte e com o passar dos trimestres essa quantidade tende à diminuir. De acordo com o G1 (2018) "O desemprego no Brasil seguiu em queda e ficou em 11,8% no quarto trimestre de 2017, atingindo 12,3 milhões de pessoas, 5% a menos do que nos três meses anteriores. Apesar de as taxas indicarem uma tendência de queda a cada trimestre, a qualidade dos empregos não melhorou, já que a maioria das vagas não oferece carteira assinada".

4 Considerações finais

Para pessoas ocupadas, o coeficiente de variação obtido foi alto em todos os anos, sendo possível observar um crescimento constante dessa medida. Portanto, a variabilidade dos dados em torno do número médio de pessoas ocupadas registrado vem crescendo com o passar dos anos. Para pessoas desocupadas conclui-se que a queda na quantidade de pessoas desocupadas no final dos anos se deu

pelo aumento de oportunidades para se trabalhar no mercado de trabalho informal, ou conta própria.

Referências

- [1] ARAÚJO, T. **Dinâmica regional brasileira e integração competitiva**. Anais do VII Encontro Nacional da ANPUR, Recife, 1997.
- [2] BORJAS, G. **Labor Economics**. McGraw-Hill, 1996, p. 318–357.
- [3] EHRENBER, R SMITH, R. **Modern Labor Economics: theory and public policy**. Addison-Wesley Educational Publishers, 1996.
- [4] REIS, J.G.A. BARROS, R. **Wage inequality and the distribution of education: a study of the evolution of regional differences in inequality in metropolitan Brazil**. Journal of Development Economics, 36, 1991.
- [5] CRUZ, C.D.; FERREIRA, F.M.; PESSONI, L.A. **Biometria aplicada ao estudo da diversidade genética**. Visconde do Rio Branco: Suprema, 2011. 620p.
- [6] FÁVERO, L. P. et al. **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Campus Elsevier, 2009.
- [7] JOHNSON, R.A. and WICHERN, D.W. **Applied Multivariate Statistical Analysis**. New Jersey-USA: Englewood Cliffs, 642p. 1998.
- [8] HAIR JR., J. F. et al. **Análise multivariada de dados**. 5. ed. Tradução Adonai Schlup Sant’Anna e Anselmo Chaves Neto. Porto Alegre: Bookman, 2005.
- [9] KHATTREE, R.; NAIK, D.N. **Multivariate data reduction and discrimination with SAS software**. Cary, SAS Institute, 558p., 2000.
- [10] IBGE. **Estimativa Populacional 2020**. Disponível em <https://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2020/estimativa_dou_2020.pdf>. Acesso em: 05 de novembro de 2020.
- [11] CURY A.; BRITO C. **Desemprego fica em 11,8% no 4º trimestre de 2017**. Disponível em <<https://g1.globo.com/economia/noticia/desemprego-fica-em-118-no-4-trimestre-de-2017.ghtml>>. Acesso em: 24 de outubro de 2020.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Álgebra Linear aplicada a circuitos elétricos Linear Algebra applied to electrical circuits

Ariely da Silva Camargo

UFU, Faculdade de Engenharia Mecânica, Uberlândia, Minas Gerais, Brasil

ariely.camargo@ufu.br

Ana Paula Tremura Galves

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

ana.galves@ufu.br

Palavras-chave

Leis de Kirchhoff.
Circuitos elétricos.
Sistemas lineares.

Resumo

Nos estudos de circuitos elétricos, é importante obter a intensidade da corrente em cada ponto de um circuito. Em circuitos complexos, as intensidades são obtidas através das Leis de Kirchhoff. No entanto, essas leis pressupõem o uso de Sistemas Lineares. Assim, esse trabalho aborda não somente as Leis de Kirchhoff, como também métodos de resolução de Sistemas Lineares.

Keywords

Kirchhoff's Laws.
Electric circuits.
Linear systems.

Abstract

In studies of electrical circuits, it is important to obtain the current intensity at each point in a circuit. In complex circuits, intensities are obtained through Kirchhoff's Laws. However, these laws presuppose the use of Linear Systems. Thus, this work addresses not only Kirchhoff's Laws, but also methods of solving Linear Systems.

1 Introdução

Este trabalho trata de circuitos elétricos básicos, para os quais será mostrado que é possível determinar as suas correntes elétricas por meio da resolução de um sistema linear, conteúdo básico de Álgebra Linear. Os circuitos são ferramentas indispensáveis à sociedade, visto que o entendimento e aperfeiçoamento dessa ferramenta possibilita o avanço da tecnologia. Assim, utilizando as regras de Kirchhoff, o qual baseia o Método do Potencial, e técnicas de resolução de sistemas lineares, será possível calcular as intensidades das correntes em um dado circuito.

2 Conceitos

A seguir alguns conceitos físicos importantes são definidos. Todas essas definições foram baseadas nas referências [1] e [2].

- Corrente elétrica: fenômeno que ocorre quando há o fluxo líquido de cargas através de uma superfície. Associa-se, a este fenômeno, a grandeza intensidade da corrente, a qual a unidade de medida é o ampère, simbolizado por A .
- Tensão: é a diferença de potencial (ddp) entre dois pontos do espaço. Nos estudos de circuitos, percebe-se que ao submeter dois pontos de um fio metálico a uma diferença de potencial, surge uma corrente elétrica. A unidade de medida da tensão é o volt, simbolizado por V . Um tipo de equipamento capaz de estabelecer uma diferença de potencial entre dois pontos de um fio é uma fonte de tensão. Alguns exemplos de fontes são as pilhas e as baterias.
- Resistência: ao submeter diferentes materiais a uma mesma tensão e medir as intensidades da corrente, percebe-se que elas são diferentes. Define-se, portanto, resistência r do material quando submetido a uma tensão u pela razão $\frac{u}{i}$, em que i é a intensidade da corrente. Além disso, quando esse material é metálico, por exemplo, os valores encontrados para r são iguais para um certo intervalo de valores de u . Assim, a expressão $r = \frac{u}{i}$ ficou conhecida como *Primeira Lei de Ohm*. A unidade de medida da resistência é o ohm, simbolizado por Ω .
- Resistores: são condutores elétricos cuja função no circuito é oferecer resistência elétrica. Os resistores obedecem a Primeira Lei de Ohm.

Utilizando a Primeira Lei de Ohm e os conceitos básicos explicados acima, pode-se calcular a intensidade da corrente em um circuito sem precisar utilizar aparelhos de medição. Assim, na próxima seção será apresentado um método para isso, chamado de Método do Potencial, o qual consiste em determinar as intensidades das correntes a partir da queda e aumento do potencial em um circuito. Para utilizar esse método, é necessário o uso das leis de Kirchhoff, as quais serão enunciadas de acordo com a necessidade de utilizá-las. Neste sentido, será estudado, primeiramente, um circuito de uma malha.

3 Circuito de uma malha

O circuito elétrico de uma malha estudado neste trabalho é apresentado na Figura 1.

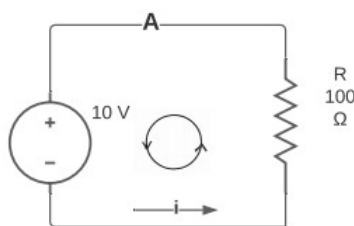


Figura 1: Circuito de uma malha

Nesse circuito, há a presença de uma fonte de tensão, um resistor e fios metálicos com resistência desprezível. Sabendo a resistência do resistor e a tensão que a fonte fornece, pode-se calcular a intensidade da corrente do circuito, utilizando o Método do Potencial.

Para isso, tome um ponto inicial A do circuito cujo potencial é v_A . Após isso, "percorra" o circuito em um sentido adotado qualquer e retorne ao ponto A. Durante esse percurso, ao passar pelo resistor, há uma diminuição de potencial e ao passar pela fonte, há um aumento ou queda de potencial. No entanto, ao retornar ao ponto A, o potencial medido em A continua sendo v_A . Esse princípio embasa a *Lei das tensões de Kirchhoff*.

Lei das tensões de Kirchhoff. A soma das quedas e ganhos de potencial ao longo de uma malha de um circuito é igual a zero.

Quando há queda, a tensão é negativa, mas quando há o aumento de potencial a tensão é positiva. Além disso, vale as seguintes observações:

Considere uma corrente que atravessa um resistor de resistência r , pela primeira Lei de Ohm a tensão u que os terminais do resistor está submetido é $r \cdot i$. No entanto, como ocorre a queda de tensão a variação do potencial é $-r \cdot i$. Agora, considere uma fonte de tensão que fornece uma tensão de u V ao circuito. Assim, se a fonte de tensão é atravessada do terminal negativo para o terminal positivo, ocorre uma variação de potencial de u V. Caso contrário, a variação é $-u$ V.

Assim, no circuito da Figura 1, em que o ponto A está destacado e o sentido adotado da corrente é o anti-horário, obtém-se a seguinte equação: $-r \cdot i - 10 = 0 \Rightarrow i = \frac{-10}{100} = -0,1$. Logo, a intensidade da corrente é $-0,1$ A.

Perceba que a intensidade da corrente é negativa. No entanto, isso apenas significa que o sentido adotado para a corrente na resolução é diferente do seu sentido convencional. Assim, na verdade, o sentido da corrente é horário.

4 Circuito com muitas malhas

Antes de analisar circuitos com mais de uma malha, é preciso entender o conceito de nó de um circuito e a *Lei dos nós de Kirchhoff*.

Um nó de um circuito é um ponto em que a corrente se divide.

Lei dos nós de Kirchhoff. A soma das intensidades da corrente que chega em um nó é igual a soma das intensidades da corrente que sai desse nó.

Usando a ideia de que a corrente é o movimento de partículas, essa lei é o princípio da conservação das cargas, pois uma partícula não pode ser criada nem destruída ao chegar em um nó. Agora, será apresentado como calcular a intensidade da corrente em um circuito com mais de uma malha. Para isso, vale-se o uso do exemplo apresentado na Figura 2.

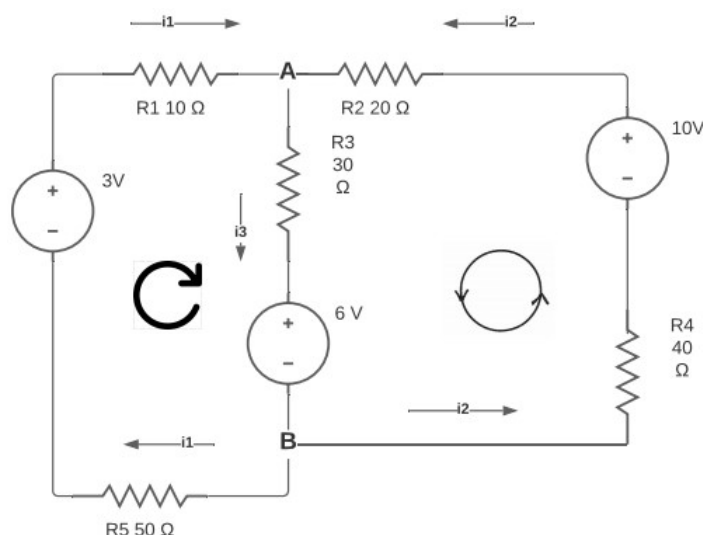


Figura 2: Circuito com duas malhas

Neste circuito, observa-se a existência dos nós A e B, utilizando a Lei dos nós de Kirchhoff e chamando de i_1 e i_2 as correntes que chegam em A, tem-se que $i_3 = i_1 + i_2$. Já no nó B, percebe-se que chega-se na mesma expressão anterior. Com isso, fica-se adotado os sentidos das correntes no circuito. Na malha da direita a corrente tem o sentido anti-horário, já na da esquerda, a corrente tem o sentido horário.

Agora, aplicando-se a Lei das tensões de Kirchhoff nas malhas da direita e esquerda, respectivamente, obtém-se as equações II e III, e juntando-as com a equação $i_3 = i_1 + i_2$, obtém-se o seguinte sistema linear:

$$\begin{cases} -i_1 - i_2 + i_3 = 0 \\ -60i_2 - 30i_3 = -4 \\ -60i_1 - 30i_3 = 3 \end{cases} \quad (1)$$

Assim, percebe-se que o problema resultante consiste em resolver um sistema linear.

Para resolver um sistema linear com poucas incógnitas, pode-se usar o Método de Gauss ou a eliminação gaussiana. Já sistemas com muitas incógnitas podem ser resolvidos por outros métodos, como a Decomposição LU. Para mais informações sobre esses métodos, consulte [3]. Neste trabalho, no entanto, serão detalhadas as três operações que podem ser aplicadas em sistemas lineares, a fim de encontrar sua solução. Essas operações são a base dos métodos citados e de outros.

5 Sistemas Lineares

Definição 5.1. Considere uma equação do tipo $\sum_{i=1}^n a_i x_i = b$, em que $a_i, b \in \mathbb{R}, \forall i \in \{1, \dots, n\}$. Equações desse tipo são chamadas equações lineares nas incógnitas x_1, \dots, x_n . Os reais a_i e b são chamados, respectivamente, de coeficientes e termo independente da equação.

Definição 5.2. Considere uma equação linear. Neste sentido, uma ênupla ordenada $(\alpha_1, \dots, \alpha_n)$ é dita solução da equação se, e somente se, ao substituir respectivamente x_1, \dots, x_n por $(\alpha_1, \dots, \alpha_n)$, a sentença é verdadeira.

Definição 5.3. Um sistema linear nas incógnitas x_1, \dots, x_n é um conjunto de m equações lineares nas incógnitas x_1, \dots, x_n .

Definição 5.4. Seja S um sistema linear de m equações e n incógnitas. A ênupla ordenada $(\alpha_1, \dots, \alpha_n)$ é dita solução do sistema se é solução de cada equação do sistema.

Definição 5.5. Dois sistemas S_1 e S_2 são ditos equivalentes quando todas as soluções de S_1 são soluções de S_2 e todas as soluções de S_2 são soluções de S_1 .

Definição 5.6. Um sistema de m equações e n incógnitas é dito escalonado quando a quantidade de coeficientes nulos antes do primeiro coeficiente não nulo aumenta de uma equação para outra.

5.1 Técnicas de resolução de sistemas lineares

Agora, será mostrado como encontrar as soluções de sistemas escalonados e não escalonados.

Para encontrar as soluções de um sistema escalonado basta fazer uma retrosubstituição. Assim, começando da equação com o maior número de coeficientes nulos, encontra-se o valor da incógnita e substitui o que foi encontrado na segunda equação com maior número de coeficientes nulos. Fazendo isso sucessivamente, encontra-se todos os valores das incógnitas.

Para sistemas não escalonados, no entanto, não é possível fazer a retrosubstituição. Assim, a fim de encontrar soluções de um sistema S_1 não escalonado, é preciso encontrar um sistema escalonado S_2 equivalente a S_1 . Para isso, usa-se operações elementares. As operações elementares são operações feitas sobre as equações de um sistema, de modo a encontrar sistemas equivalentes ao inicial. Abaixo, encontram-se as três operações:

- Trocar duas equações de lugares; multiplicar uma equação por um número real diferente de zero; substituir uma equação por ela somada a outra equação do sistema.

A justificativa para tais operações pode ser encontrada em [4]. Além disso, na prática pode-se substituir uma equação por uma combinação linear dela com outras, resultando em um sistema equivalente, desde que não se multiplique nenhuma das equações por zero. Isso é justificável pelas três operações.

6 Resolução do circuito com mais de uma malha

Com as operações elementares, pode-se resolver o sistema 1. Os símbolos L_1 e L_2 abaixo indicam, respectivamente, a primeira e segunda equação do quarto sistema.

$$\begin{aligned} & \begin{cases} -i_1 - i_2 + i_3 = 0 \\ -60i_2 - 30i_3 = -4 \\ -60i_1 - 30i_3 = 3 \end{cases} \iff \begin{cases} i_1 + i_2 - i_3 = 0 \\ 60i_2 + 30i_3 = 4 \\ 60i_1 + 30i_3 = -3 \end{cases} \iff \begin{cases} i_1 + i_2 - i_3 = 0 \\ 30i_2 + 15i_3 = 2 \\ 20i_1 + 10i_3 = -1 \end{cases} \\ \iff & \begin{cases} i_1 + i_2 - i_3 = 0 \\ 20i_1 + 10i_3 = -1 \\ 30i_2 + 15i_3 = 2 \end{cases} \iff \begin{cases} i_1 + i_2 - i_3 = 0 \\ -20i_2 + 30i_3 = -1 \\ 30i_2 + 15i_3 = 2 \end{cases} \xrightarrow{-20L_1+L_2} \iff \begin{cases} i_1 + i_2 - i_3 = 0 \\ -20i_2 + 30i_3 = -1 \\ i_3 = \frac{1}{120} \end{cases} \end{aligned}$$

Esse último sistema está escalonado, logo basta fazer retrossubstituição, encontrando $i_2 = \frac{1}{16}$ e $i_1 = \frac{-13}{240}$. Assim, as correntes no circuito são: $i_1 = \frac{-13}{240} A$, $i_2 = \frac{1}{16} A$ e $i_3 = \frac{1}{120} A$. Além disso, como a intensidade de i_1 é negativa, o sentido adotado para corrente na resolução do exercício é oposto ao seu sentido convencional.

7 Conclusões Finais

Com esse trabalho foi visto uma aplicação da Álgebra Linear no estudo de circuitos elétricos com mais de uma malha. Nesse sentido, a resolução de sistemas lineares são indispensáveis para encontrar as correntes nesse tipo de circuito. Além disso, foi apresentado as operações elementares utilizadas na resolução de sistemas, uma ideia na qual muitas técnicas de resoluções de sistemas lineares se baseiam.

Agradecimentos ao PICME, pela oportunidade de pesquisa e ao CNPq, pelo auxílio financeiro.

Referências

- [1] HALLIDAY, D; RESNICK, R. **Fundamentos da Física Eletromagnetismo**. 10ª Edição. Rio de Janeiro: LTC, 2016.
- [2] BISCUOLA, G.J; BÔAS, V.N; DOCA, H.R. **FÍSICA 3: Eletricidade; Física Moderna; Análise Dimensional**. 2ª Edição. São Paulo: Saraiva, 2013.
- [3] ANTON, H; RORRES, C. **Álgebra Linear com aplicações**. 10º Edição. São Paulo: Bookman, 2012.
- [4] IEZZI, G;HAZZAN, S. **Fundamentos da Matemática Elementar: sequências, matrizes, determinantes, sistemas**. 8º Edição. São Paulo: Atual Editora, 2013.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Álgebra Linear aplicada aos Jogos de estratégia Linear Algebra applied to Strategy Games

Beatriz Harumi Nishiyama

UFU, Instituto de Psicologia, Uberlândia, Minas Gerais, Brasil

beatriz.nishiyama@ufu.br

Ligia Laís Fêmina

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

ligia@ufu.br

Palavras-chave

Matrizes.
Probabilidade.
Jogos de estratégia.
Matriz de compensação.
Compensação esperada.

Resumo

A teoria de matrizes é um tópico estudado no Ensino Médio e em várias disciplinas da Graduação, como Álgebra Linear. Muitas vezes nos deparamos com a pergunta: para que estudar isso? Nesse trabalho, apresentamos uma das inúmeras aplicações de matrizes, que é o estudo dos jogos de estratégia, mostrando o quão importante e interessante é o estudo de Álgebra linear.

Keywords

Matrices.
Probability.
Strategy Games.
Compensation matrix.
Expected compensation.

Abstract

Matrix theory is a topic studied in high school and in several graduation subjects, like as Linear Algebra. We are often faced with the question: why study this? In this work, we present one of the numerous matrix applications, which is the study of strategy games, showing how important and interesting it is to study Linear Algebra.

1 Introdução

A Álgebra Matricial fornece maneiras compactas de resolver um sistema de equações lineares, a qual nos ajuda nos problemas cotidianos. Uma das aplicabilidades para teoria das matrizes é na compreensão os jogos de estratégia [1].

A Teoria dos Jogos analisa as situações em que a ação de um indivíduo ou grupo, depende das ações dos demais. Nesse sentido, para que um indivíduo consiga tomar decisões deve considerar a decisão dos outros indivíduos [2].

Considerando um jogo genérico, no qual dois oponentes escolhem estratégias distintas, a fim de alcançar objetivos opostos, o objetivo desse trabalho é introduzir alguns conceitos básicos da teoria dos jogos e analisar os casos em que a estratégia ótima de cada jogador é encontrada com o uso de técnicas matriciais.

2 O jogo

Serão participantes deste jogo, dois jogadores (*jogador L e jogador C*), os quais possuem cada qual uma roda (*roda das linhas e roda das colunas*). Essa é configuração das rodas:

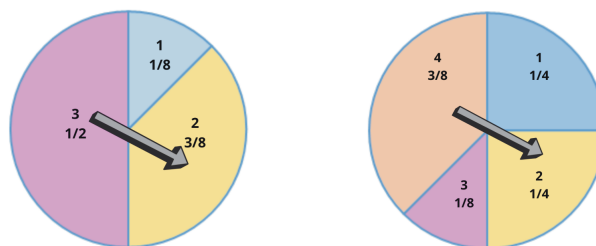


Figura 1: Roda de Linhas e Roda de Colunas

O número que o ponteiro indicar, refere-se ao movimento que cada jogador deve fazer em relação a tabela abaixo:

		Movimento do jogador C			
		1	2	3	4
movimento do jogador L	1	1	3	-3	-4
	2	-1	5	3	-2
	3	4	-4	0	6

Tabela 1: Pagamentos ao jogador L

Por exemplo, se o ponteiro da roda das linhas parar no 1, o jogador L faz o movimento 1; e o ponteiro da roda das colunas parar no 2, o jogador C faz o movimento 2. Portanto, o jogador C deve pagar 3 reais ao jogador L. Caso a combinação dos movimentos, caia em algum valor negativo, o

pagamento é inverso, o jogador L deve pagar o jogador C. Nesse sentido, à medida que os movimentos são decididos pela sorte, cada jogador precisa saber a probabilidade de ganhar ou perder, para decidir se irá jogar ou não.

3 Jogos de matriz de duas pessoas com soma zero

O termo “soma zero” significa que o mesmo valor que um jogador ganhar, o outro irá perder. Já o termo “jogo de matriz” refere-se aos jogos de duas pessoas com um número finito de movimentos, em que os resultados e ganhos dos jogadores podem ser arranjados em forma tabular ou matricial, como observado na Tabela 1 [1].

Nesse tipo de jogo, define-se m o número de movimentos do jogador L e n o número de movimentos do jogador C. Existe, portanto, uma compensação de movimentos do jogador C para o jogador L. Sendo $i = 1, 2, \dots, m$ e $j = 1, 2, \dots, n$. Portanto, a_{ij} é a compensação do jogador C para o jogador L, se o jogador L fizer o movimento i e o jogador C fizer o movimento j [1].

As mn compensações possíveis podem ser arranjadas de acordo com a **matriz de compensação** ou **matriz de pagamento**, organizada nessa forma de matriz $m \times n$, como descrito por Anton [1]:

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

O movimento de cada jogador segue uma probabilidade, a qual corresponde a área de um setor para a área da roda. Por exemplo, de acordo com a Figura 1, a probabilidade do jogador L fazer o movimento 2 é de $\frac{3}{8}$. Já a probabilidade do jogador C fazer o movimento 2, como demonstra a Figura 2, é de $\frac{1}{4}$.

Sendo assim, seguem as definições: p_i = probabilidade do jogador L fazer o movimento i ($i = 1, 2, \dots, m$) e q_j = probabilidade do jogador C fazer o movimento j ($j = 1, 2, \dots, n$) [1].

Além disso: $p_1 + p_2 + \cdots + p_m = 1$ e $q_1 + q_2 + \cdots + q_n = 1$.

Formando o vetor linha p , correspondente a **estratégia do jogador L** e o vetor coluna q , correspondente a **estratégia do jogador C**. Da seguinte forma: $p = (p_1 \ p_2 \ \cdots \ p_m)$ e $q = (q_1 \ q_2 \ \cdots \ q_n)^T$ [1].

Nesse sentido, a partir da Teoria de Probabilidades, visto que o movimento dos dois jogadores são independentes, a probabilidade do jogador L fazer o movimento i e o jogador C fazer o movimento j , em qualquer lance é $p_i q_j$ [1].

Multiplicando as compensações possíveis, pela correspondente probabilidade e somando-as, obtemos [1]:

$$a_{11}p_1q_1 + a_{12}p_1q_2 + \cdots + a_{1n}p_1q_n + a_{21}p_2q_1 + \cdots + a_{mn}p_mq_n \quad (1)$$

Essa média ponderada é denominada **compensação esperada** para o jogador L, ou seja, se o jogo for jogado muitas vezes, essa é a compensação média do jogador L. Denota-se $E(\mathbf{p}, \mathbf{q})$ a compensação esperada por L. A compensação esperada pode ser expressa em notação matricial como a seguir [1]:

$$E(\mathbf{p}, \mathbf{q}) = \begin{pmatrix} p_1 & p_2 & \cdots & p_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} = \mathbf{pAq} \quad (2)$$

Se a compensação esperada por L é $E(\mathbf{p}, \mathbf{q})$, então a compensação esperada por C é $-E(\mathbf{p}, \mathbf{q})$.

3.1 Mudança de estratégias

O jogo que vimos até agora, a estratégia estava predeterminada. A partir de agora, essa estratégia pode mudar, independentemente, ou seja, no jogo de exemplo, os jogadores podem mudar os setores das rodas. A partir disso, nenhum dos dois irá conhecer a estratégia do outro, porém eles irão fazer a melhor estratégia possível, e o outro jogador reconhece. Portanto, o jogador L, irá escolher \mathbf{p} tal que $E(\mathbf{p}, \mathbf{q})$ seja a maior possível para a melhor estratégia de \mathbf{q} , e vice-versa. Para entender as escolhas, devemos recorrer ao **teorema fundamental dos jogos de duas pessoas com soma zero** presente no livro de Anton (2012) [1].

Teorema 3.1. *Teorema fundamental dos jogos com soma zero*

Existem estratégias \mathbf{p}^ e \mathbf{q}^* tais que:*

$$E(\mathbf{p}^*, \mathbf{q}) \geq E(\mathbf{p}^*, \mathbf{q}^*) \geq E(\mathbf{p}, \mathbf{q}^*) \quad (3)$$

quaisquer que sejam as estratégias de \mathbf{p} e \mathbf{q}

As estratégias \mathbf{p}^* e \mathbf{q}^* são as estratégias ótimas de cada jogador. Nesse sentido $v = E(\mathbf{p}^*, \mathbf{q}^*)$. Pensando na desigualdade da esquerda, da inequação em (3):

$$E(\mathbf{p}^*, \mathbf{q}) \geq v, \text{ com qualquer estratégia } q.$$

Portanto, se o jogador L escolher a estratégia ótima, independente da estratégia de C, sua compensação nunca será menor que v . No entanto, o melhor que L pode fazer, é impedir que sua compensação caia abaixo de v , e o jogador C garantir que sua compensação seja pelo menor o valor de v , com a sua melhor estratégia. Anton (2012) define:

Definição 3.2. *Seguindo o teorema 3.1*

- (i) \mathbf{p}^* é a estratégia ótima para o jogador L;
- (ii) \mathbf{q}^* é a estratégia ótima para o jogador C;
- (iii) a compensação esperada $v = E(\mathbf{p}^*, \mathbf{q}^*)$ é o **valor** do jogo.

Portanto, quaisquer pares de estratégias ótimas resultam no mesmo valor v do jogo. Para encontrar as estratégias ótimas, é necessário encontrar vetores \mathbf{p}^* e \mathbf{q}^* que satisfaçam a inequação em (3). Normalmente, essas estratégias são encontradas por meio de técnicas de Programação Linear, porém em casos especiais é possível utilizar técnicas mais elementares. Segue uma segunda definição, proposta por Anton (2012):

Definição 3.3. Uma entrada a_{rs} de uma matriz de compensação A , denomina-se **ponto de sela** se

(i) a_{rs} for a menor entrada em sua linha e

(ii) a_{rs} for a maior entrada em sua coluna

Dizemos que um jogo cuja matriz de compensação possui um ponto de sela é **estritamente determinado**.

Caso a matriz tenha um ponto de sela, as estratégias ótimas para os dois jogadores são: $\mathbf{p}^* = (0 \ 0 \ \dots \ 1 \ \dots \ 0)$ e $\mathbf{q}^* = (0 \ 0 \ \dots \ 1 \ \dots \ 0)^T$. Na qual o número 1, no vetor \mathbf{p}^* , refere-se a r -ésima entrada e no vetor \mathbf{q}^* , refere-se a s -ésima entrada. Demonstrando o movimento de uma estratégia ótima. As estratégias nas quais apenas um movimento é possível, chama-se **estratégias puras**, e nas que podem ser feitas mais de um movimento, chama-se **estratégias mistas**. As estratégias acima são ótimas, pois coincidem com a inequação em (3). O valor do jogo estritamente determinado é o valor numérico do ponto de sela, uma matriz de compensação pode ter mais de um ponto de sela, mas eles devem ter o mesmo valor numérico, por conta da unicidade do valor de um jogo [1].

3.2 Jogos de matrizes 2x2

Segue o Teorema proposto por Anton (2012):

Teorema 3.4. (Estratégias ótimas para jogos com matrizes 2x2) Jogos que não sejam estritamente determinados, são estratégias ótimas para os jogadores L e C

$$\mathbf{p}^* = \left(\frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \quad \frac{a_{11} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \right) \quad \mathbf{q}^* = \begin{pmatrix} \frac{a_{22} - a_{12}}{a_{11} + a_{22} - a_{12} - a_{21}} \\ \frac{a_{11} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} \end{pmatrix}$$

O valor do jogo é:

$$v = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}}$$

Nesse sentido, as entradas dos vetores \mathbf{p}^* e \mathbf{q}^* são números entre 0 e 1. Portanto, a Equação permite que cada jogador ao escolher sua estratégia ótima, consiga alcançar a compensação esperada, independente da estratégia adversária. Vale destacar que isso só é válido em jogos com 2 movimentos.

4 Conclusão

A teoria dos jogos é importante, pois ela pode ser utilizada em diversos problemas do cotidiano, como no exemplo a seguir. Na política de vacinação do vírus da gripe, é necessário entender que existem dois sorotipos. A proporção de cada um é desconhecida. Nesse sentido, existem duas vacinas diferentes, a primeira possui 80% de eficácia contra o sorotipo 1 e 75% de eficácia contra o sorotipo 2; já a segunda possui 70% de eficácia contra o sorotipo 1 e 90% de eficácia contra o sorotipo 2. Portanto, qual deve ser a proporção de cada vacina para obter a maior eficácia?

Solução: Pode-se considerar que as duas pessoas jogando, seria o governo (vacina), para obter a maior compensação, contra o vírus, fazendo ele obter a menor compensação possível. Desse modo, a matriz de compensação é:

$$A = \begin{pmatrix} 0,80 & 0,75 \\ 0,70 & 0,90 \end{pmatrix}$$

Já que a matriz não possui ponto de sela, é possível aplicar o Teorema 3.4:

$$p_1^* = \frac{a_{22} - a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} = \frac{0,90 - 0,70}{0,80 + 0,90 - 0,75 - 0,7} = \frac{0,20}{0,25} = \frac{4}{5}$$

$$p_2^* = 1 - p_1^* = 1 - \frac{4}{5} = \frac{1}{5}$$

$$v = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11} + a_{22} - a_{12} - a_{21}} = \frac{(0,80)(0,90) - (0,75)(0,70)}{0,80 + 0,90 - 0,75 - 0,7} = \frac{0,195}{0,25} = 0,78$$

Portanto, para o governo é uma estratégia ótima, vacinar $\frac{4}{5}$ com a vacina 1 e $\frac{1}{5}$ com a vacina dois garantindo 78% de eficácia, independente da distribuição do sorotipo.

Agradecimentos

Agradeço ao PICME por me proporcionar a oportunidade de construir esse trabalho.

Referências

- [1] ANTON, Howard , **Álgebra Linear com Aplicações**, Bookman , 2012
- [2] FIGUEIREDO, R. S. **Teoria dos jogos: conceitos, formalização matemática e aplicação à distribuição de custo conjunto**. Gestão e Produção, São Carlos, v 1, n 3, 1994.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Autovalores e Autovetores de Matrizes Toeplitz Tridiagonais Simétricas Eigenvalues and Eigenvectors of Tridiagonal Simetric Toeplitz Matrices

João Vitor Caixeta Alcarria

UFU, Faculdade de Engenharia Mecânica, Uberlândia, Minas Gerais, Brasil

joao.alcarria@ufu.br

Santos Alberto Enriquez-Remigio

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

santos.er@ufu.br

Palavras-chave

Autovalores.
Autovetores.
Equações de diferenças.
Determinante.
Matriz Singular.

Resumo

Determinar autovalores e autovetores é um problema que surge nas mais diversas áreas, como na análise de vibrações, mecânica dos sólidos, estatística, entre outras. Dependendo do problema, o conhecimento dos autovalores das matrizes, que aparece na sua modelagem, pode definir a dinâmica do problema. No presente trabalho, será apresentado o processo de demonstração da fórmula dos autovalores e autovetores para matrizes Toeplitz tridiagonais simétricas. Tal processo está baseado no método de solução de equações de diferenças de segunda ordem.

Keywords

Eigenvalues.
Eigenvectors.
Differential Equations.
Determinant.
Singular Matrix.

Abstract

Determining eigenvalues and eigenvectors is a current problem that appears in various areas, for example, vibration analysis, solid mechanics, statistics and many others. Depending on the issue, the knowledge of eigenvalues of matrices, which appears in the modelling process, can define the dynamic part of the problem. In the present work, it will be shown the process of demonstration of the eigenvalues' and eigenvectors' formula for symmetric tridiagonal Toeplitz matrices. This process is based on the solution method of second order differential equations.

1 Introdução

O estudo deste artigo será direcionado à demonstração da fórmula dos autovalores e autovetores de matrizes Toeplitz tridiagonais simétricas. Tais matrizes aparecem em diferentes problemas práticos e o conhecimento de seus autovalores e autovetores permitem entender a dinâmica dos problemas em análise. Por exemplo, no caso do tópico de vibrações, eles aparecem como meio de determinar a frequência de vibração de uma estrutura e a sua variação de posição.[1]

A seguir, serão apresentados alguns conceitos necessários para o entendimento do que será mostrado.

Definição 1.1 (Autovalores e Autovetores). *Um vetor $\vec{x} \in \mathbb{R}^n$ é um vetor próprio (ou autovetor) de uma matriz A ($n \times n$) se existe um número (real ou complexo) λ*

$$A\vec{x} = \lambda\vec{x}$$

O número (real ou complexo) λ tal que $A\vec{x} = \lambda\vec{x}$ é denominado de valor próprio (ou autovalor) da matriz A associado ao vetor próprio \vec{x} .

Diante da situação e da definição apresentadas, será mostrada uma fórmula para determinar os autovalores e autovetores de matrizes do tipo Toeplitz tridiagonais simétricas de ordem N [3], as quais são expressas, de modo genérico, por:

$$A_N = \begin{bmatrix} a & b & & \cdots & 0 \\ b & a & b & & 0 \\ 0 & b & a & b & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & & b & a & b \\ 0 & \cdots & & & b & a \end{bmatrix} \quad (1)$$

onde a e b são valores determinados a partir do problema a ser solucionado, como o que foi apresentado.

2 Desenvolvimento

2.1 Cálculo dos autovalores

A seguir, apresenta-se uma propriedade que permite o cálculo dos autovalores de uma matriz, conhecendo-se os autovalores de outra matriz associada a ela.

Proposição 2.1. *Sejam A e B , duas matrizes. Se $A = P_n(B)$, onde $P_n(\lambda)$ é uma função polinomial, então as raízes do polinômio característico de A são dadas por $P_n(\lambda)$, $k = 1, 2, \dots, N$, com λ raiz do polinômio característico de B [5].*

A matriz A_N dada em (1) pode ser reescrita por $A_N = aI_N + bB$, onde as matrizes de ordem N , I_N e B , são respectivamente, a matriz identidade e a matriz dado por:

$$B = \begin{bmatrix} 0 & 1 & & \cdots & 0 \\ 1 & 0 & 1 & & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & & 1 & 0 & 1 \\ 0 & \cdots & & & 1 & 0 \end{bmatrix} \quad (2)$$

Para calcular os autovalores de B , deve-se encontrar as raízes de seu polinômio característico. Seja $\phi_N(\gamma) = \det(\gamma I - B)$ o polinômio característico de B , então, as raízes do polinômio característico serão soluções da equação $\phi_N(\gamma) = 0$. Porém, em vez de calcular a solução dessa equação diretamente, uma vez que possui um desenvolvimento muito complicado, o processo de solução passará pela determinação de uma expressão equivalente para $\phi_N(\gamma)$.

$$\phi_N(\gamma) = \det(\gamma I - B) = \begin{vmatrix} \gamma & -1 & & \cdots & 0 \\ -1 & \gamma & -1 & & 0 \\ 0 & -1 & \gamma & -1 & \cdots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \cdots & & -1 & \gamma & -1 \\ 0 & \cdots & & & -1 & \gamma \end{vmatrix}$$

A expressão para $\phi_{N+2}(\gamma) = \det(\gamma I - B)$ pode ser escrita de forma recursiva, seguindo o resultado dado em [5]:

$$\phi_{(N+2)}(\gamma) = \gamma \cdot \phi_{(N+1)}(\gamma) - \phi_N(\gamma) \quad (3)$$

Logo, a relação acima define uma equação de diferença dada por:

$$\phi_{(N+2)}(\gamma) - \gamma \cdot \phi_{(N+1)}(\gamma) + \phi_N(\gamma) = 0 \quad (4)$$

A expressão (4) é uma equação de diferença de segunda ordem e, pela teoria de métodos de solução de equações de diferenças [5], sabe-se que a expressão para $\phi_N(\gamma)$ é:

$$\phi_N(\gamma) = C_1(r_1)^N + C_2(r_2)^N \quad (5)$$

onde C_1 e C_2 são constantes a serem determinadas e, r_1 e r_2 são as raízes da equação de segundo grau:

$$r^2 - \gamma r + 1 = 0.$$

Aplicando a fórmula de Bhaskara, encontra-se a expressão para r_1 e r_2 , as quais dependem de γ .

Considerando a substituição de variável $\gamma = 2 \cos(\theta)$, temos que: $r_1 = e^{i\theta}$ e $r_2 = e^{-i\theta}$, onde i é a unidade imaginária. Para determinar as constantes C_1 e C_2 , utiliza-se o fato de se conhecer as raízes de ϕ_1 e ϕ_2 . Portanto, fazendo uso dessas informações, a expressão (5), após simplificações, pode ser reescrita como:

$$\phi_N(\gamma) = \frac{\text{sen}((N + 1)\theta)}{\text{sen}(\theta)} \tag{6}$$

A expressão (6) é uma expressão equivalente para o polinômio característico de B . Diante disso, o cálculo das raízes de $\phi_N(\gamma)$ determina de (6) que $(N + 1)\theta = k\pi$, onde $k \in \mathbb{Z}$. Logo, $\theta = \frac{k\pi}{(N+1)}$. Porém, a raiz de ϕ_N é γ , então as N raízes (autovalores da matriz B) são dadas por:

$$\gamma = 2 \cos\left(\frac{k\pi}{(N + 1)}\right), \quad k = 1, 2, \dots, N.$$

Portanto, os autovalores da matriz A são dados por:

$$\lambda_k = a + 2b \cos\left(\frac{k\pi}{(N + 1)}\right), \quad k = 1, 2, \dots, N.$$

2.2 Cálculo dos autovetores

Seja A a matriz dada em (1), λ_k seu autovalor e $\vec{x} = [x_1, x_2, \dots, x_N]^T$ seu respectivo autovetor, então $A\vec{x} = \lambda_k\vec{x}$ é equivalente a $(A - \lambda_k I)\vec{x} = 0$, onde a matriz $A - \lambda_k I$ é dada por:

$$\begin{bmatrix} -2b \cos\left(\frac{k\pi}{N+1}\right) & b & & \dots & & 0 \\ b & -2b \cos\left(\frac{k\pi}{N+1}\right) & b & & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \dots & b & -2b \cos\left(\frac{k\pi}{N+1}\right) & & b \\ 0 & \dots & & b & -2b \cos\left(\frac{k\pi}{N+1}\right) & \end{bmatrix} \tag{7}$$

Colocando em evidência a constante b de (4), a matriz $(A - \lambda_k I)$ pode ser escrita como sendo bM , em que M é representado por:

$$\begin{bmatrix} -2 \cos\left(\frac{k\pi}{N+1}\right) & 1 & & \dots & & 0 \\ 1 & -2 \cos\left(\frac{k\pi}{N+1}\right) & 1 & & \dots & 0 \\ 0 & 1 & -2 \cos\left(\frac{k\pi}{N+1}\right) & 1 & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \dots & & 1 & -2 \cos\left(\frac{k\pi}{N+1}\right) & 1 \\ 0 & \dots & & & 1 & -2 \cos\left(\frac{k\pi}{N+1}\right) \end{bmatrix} \tag{8}$$

A equação $(A - \lambda_k I)\vec{x} = 0$ é equivalente a $M\vec{x} = 0$, desde que $b \neq 0$. Logo, as equações do sistema linear $M\vec{x} = 0$ são:

$$\begin{aligned} -2 \cos\left(\frac{k\pi}{N+1}\right) x_1 + x_2 &= 0 \\ x_1 - 2 \cos\left(\frac{k\pi}{N+1}\right) x_2 + x_3 &= 0 \\ &\vdots \\ x_{N-3} - 2 \cos\left(\frac{k\pi}{N+1}\right) x_{N-2} + x_{N-1} &= 0 \\ x_{N-2} - 2 \cos\left(\frac{k\pi}{N+1}\right) x_{N-1} + x_N &= 0 \\ x_{N-1} - 2 \cos\left(\frac{k\pi}{N+1}\right) x_N &= 0 \end{aligned}$$

O sistema linear acima é um sistema linear homogêneo com matriz M singular. Logo, para encontrar a solução do sistema pode-se escolher x_1 como variável livre e encontrar a expressão das outras variáveis. Considere então $x_1 = c \operatorname{sen}\left(\frac{k\pi}{N+1}\right)$, $c \in \mathbb{R} - \{0\}$. Fazendo algumas manipulações, e aplicando algumas propriedades trigonométricas, é possível encontrar a expressão para as componentes x_2 e x_3 , dadas a seguir:

$$x_2 = c \operatorname{sen}\left(\frac{2k\pi}{N+1}\right) \quad e \quad x_3 = c \operatorname{sen}\left(\frac{3k\pi}{N+1}\right) \quad (9)$$

A expressão considerada para a componente x_1 e as encontradas para as componentes x_2 e x_3 , dadas em (9), indicam uma regra de formação, dada por:

$$x_L = c \operatorname{sen}\left(\frac{Lk\pi}{N+1}\right), \quad L = 1, 2, 3 \quad (10)$$

Deseja-se provar que é verdadeiro para todo $j \leq N$. Usando o princípio de indução matemática, supõe-se que seja verdadeira para todo $j \leq L$. Então, da equação para $L+1$, tem-se:

$$\begin{aligned} x_{L+1} &= 2 \cos\left(\frac{k\pi}{N+1}\right) x_L - x_{L-1} \\ x_{L+1} &= 2 \cos\left(\frac{k\pi}{N+1}\right) c \operatorname{sen}\left(\frac{Lk\pi}{N+1}\right) - c \operatorname{sen}\left(\frac{(L-1)k\pi}{N+1}\right) \\ x_{L+1} &= c \operatorname{sen}\left(\frac{(L+1)k\pi}{N+1}\right) \end{aligned}$$

Logo, o autovetor $\vec{x} = [x_1, x_2, \dots, x_N]^T$, associado a $\lambda_k = a + 2b \cos\left(\frac{k\pi}{N+1}\right)$ é:

$$\vec{x} = \left[c \operatorname{sen}\left(\frac{k\pi}{N+1}\right), c \operatorname{sen}\left(\frac{2k\pi}{N+1}\right), \dots, c \operatorname{sen}\left(\frac{Nk\pi}{N+1}\right) \right]^T$$

3 Conclusão

O presente trabalho apresenta a demonstração da expressão dos autovalores e autovetores de uma classe de matrizes Toeplitz tridiagonais simétricas reportada na literatura. Tais autovalores e autovetores podem servir para testar métodos numéricos que calculam autovalores e autovetores de matrizes tridiagonais simétricas.

Agradecimentos

Para o desenvolvimento deste trabalho, gostaria de agradecer ao orientador, Santos Alberto Enriquez Remigio, que confiou em mim e me proporcionou a oportunidade de começar a minha iniciação científica.

Referências

- [1] CHEEVER, E. **Using Eigenvalues and Eigenvectors to Study Vibrations**. Disponível em: <<https://lpsa.swarthmore.edu/MtrxVibe/Vibrations.html>>
- [2] CALLIOLI, C., A.; DOMINGUES, H., H.; COSTA, R., C., F. **Álgebra Linear e Aplicações**. 6ª. ed. São Paulo: Atual, 1990. Disponível em: <<https://sites.icmc.usp.br/frasson/aled/material/Algebra%20Linear%20e%20Aplicacoes%206ed%20-%20Callioli,%20Domingues,%20Costa.pdf>>
- [3] JESUS, H. N. **Método compacto de diferenças finitas para resolver equações de Schrödinger não lineares com dispersão de quarta ordem**. 2016. 194 f. Dissertação (Mestrado em Física) - Universidade Federal de Goiás, Goiânia, 2016. Disponível em: <<https://repositorio.bc.ufg.br/tede/bitstream/tede/6495/5/Disserta%3%a7%3%a3o%20-%20Hugo%20Naves%20de%20Jesus%20-%202016.pdf>>
- [4] YUEH, W. C. **Eigenvalues of Several Tridiagonal Matrices**, National Tsing Hua University, Maio 2005. Disponível em: <<http://www.kurims.kyoto-u.ac.jp/EMIS/journals/AMEN/2005/040903-7.pdf>>
- [5] ELLIOT, J., F.; KARNEY, D., L. **The Characteristic Roots of Certain Real Symmetric Matrices.**, University of Tennessee, Knoxville, August 1953. Disponível em: <https://trace.tennessee.edu/cgi/viewcontent.cgi?article=3834&context=utk_gradthes>



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Polinômios, Raios Laser e Tartarugas Polynomials, Laser Beams and Turtles

Roney Andrade

UFU, FAMAT, Uberlândia, MG, Brasil

roney.andrade@ufu.br

Germano Abud

UFU, FAMAT, Uberlândia, MG, Brasil

germano.abud@ufu.br

Resumo

Palavras-chave

Polinômio.
Raio laser.
Tartaruga.
Método de Lill.

Esta pesquisa tem como objetivo resgatar o método de Lill, famoso na década de 1930, que consiste em encontrar raízes reais de uma equação polinomial (caso existam) através de recursos geométricos. Existem diversas peculiaridades relacionadas ao método, podendo exigir as diferentes abordagens a depender das características do polinômio. Sendo assim, optou-se por uma revisão bibliográfica, partindo do artigo original de Lill e relacionando com problemas famosos (como o problema de Delos), dobraduras e outras generalizações desenvolvidas por diferentes autores.

Abstract

Keywords

Polynomial.
Laser Beam.
Turtle.
Lill's Method.

This research aims to rescue the Lill method, famous in the 1930s, which consists of finding real roots of a polynomial equation (if any) using geometric resources. There are several peculiarities related to the method, which may require different approaches depending on the characteristics of the polynomial. Therefore, a bibliographic review was chosen, starting from Lill's original article and relating to famous problems (such as Delos's problem), folding and other generalizations developed by different authors.

1 Introdução

Existem vários aspectos relacionados ao estudo de funções polinomiais, entretanto, o processo de encontrar suas raízes é um dos mais populares. Diferentes estudiosos desenvolveram abordagens diversas. A fórmula de Bháskara é possivelmente o mais difundido, entretanto, sua aplicação se resume à polinômios de grau dois. Com as fórmulas de Cardano é possível encontrar as raízes de polinômios de grau três. Ferrari oferece um método para tratar polinômios de grau quatro. Entretanto, segundo Galois, é impossível encontrar uma fórmula para calcular as raízes de um polinômio com grau cinco ou maior. Sendo assim, o "*Método das Tartarugas e Raios Laser*" se apresenta como uma poderosa ferramenta, tendo em vista que serve para qualquer grau maior ou igual a dois.

O primeiro a descrevê-lo desta maneira, em analogia com tartarugas, foi o matemático Hull [?], porém a autoria pertence ao engenheiro austríaco Eduard Lill [?] que, em 1867, descreve uma maneira geométrica (e elegante, diga-se de passagem) para se encontrar as raízes reais de um polinômio de grau 2 ou superior. Margherita Piazzola Beloch foi uma matemática italiana que, na tentativa de resolver equações cúbicas, introduziu uma operação conhecida atualmente como dobradura de Beloch. O golpe de mestre de Beloch foi perceber que o método de Lill (no caso cúbico) é apenas uma aplicação de seu quadrado de Beloch. Em [?] podemos encontrar mais detalhes e referências sobre este assunto.

Considere uma função polinomial qualquer $f(x)$, com grau n , na forma:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \text{ com coeficientes } a_i \in \mathbb{R}, i = 0, \dots, n \text{ e } a_n \neq 0.$$

O método de Lill consiste em obter geometricamente, caso exista, uma raiz real de $f(x)$ a partir do traçado de uma poligonal baseada em seus coeficientes.

2 Aplicação

Há diversas abordagens que buscam descrever o algoritmo de construção dessa poligonal. Garcia [?] analisa do ponto de vista algébrico, relacionando o sinais dos coeficientes a_i e a_{i-1} . Porém, em polinômios com coeficientes nulos consecutivos, existe margem para interpretações equivocadas, levando a poligonais que não satisfazem o teorema. O próprio Lill, assim como muitos autores, utilizam exemplos convenientes para fazer essa construção.

É notável o caráter visual que este método possui, sendo assim, a analogia com tartarugas e raios laser, utilizada por Hull [?], é a mais abrangente e de compreensão mais simples. Ela aborda a construção da poligonal do ponto de vista de uma tartaruga que, partindo de uma origem qualquer, anda a distância equivalente a a_n para frente (se $a_n > 0$) ou para trás (se $a_n < 0$). A cada novo coeficiente, a tartaruga faz um giro em torno de si mesma no valor de 90° , no sentido anti-horário, e realiza o mesmo procedimento quanto à distância percorrida e sentido.

Para exemplificar, tomemos o polinômio $f(x) = 5x^6 + 7x^5 + 7x^4 + 6x^3 + 5x^2 + 4x + 1$. Nosso objetivo é encontrar uma raiz real de $f(x)$, caso ela exista. Imagine então a tartaruga situada na origem O do plano cartesiano, olhando na direção positiva do eixo X .

1. A tartaruga caminha uma distância igual a $|a_6|$: a partir do ponto O traçamos um segmento $\overline{OA_6}$, orientado no sentido positivo do eixo X , com comprimento $\overline{OA_6} = |a_6| = 1$.
2. A partir do ponto A_6 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_5|$: traçamos um segmento vertical $\overline{A_6A_5}$, com comprimento $\overline{A_6A_5} = |a_5| = 7$.
3. A partir do ponto A_5 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_4|$: traçamos um segmento horizontal $\overline{A_5A_4}$, com comprimento $\overline{A_5A_4} = |a_4| = 7$.
4. A partir do ponto A_4 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_3|$: traçamos um segmento vertical $\overline{A_4A_3}$, com comprimento $\overline{A_4A_3} = |a_3| = 6$.
5. A partir do ponto A_3 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_2|$: traçamos um segmento horizontal $\overline{A_3A_2}$, com comprimento $\overline{A_3A_2} = |a_2| = 5$.
6. A partir do ponto A_2 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_1|$: traçamos um segmento vertical $\overline{A_2A_1}$, com comprimento $\overline{A_2A_1} = |a_1| = 4$.
7. A partir do ponto A_1 a tartaruga gira 90° no sentido anti-horário e caminha uma distância igual a $|a_0|$: traçamos um segmento horizontal $\overline{A_1A_0}$, com comprimento $\overline{A_1A_0} = |a_0| = 1$.
8. Ao final do percurso da tartaruga, nos posicionamos no ponto O e “acertamos” a tartaruga no ponto A_0 da seguinte maneira: atiramos com um raio laser que sempre é ricocheteado no próprio caminho percorrido pela tartaruga, com um ângulo de 90° . O raio deve ricochetear nas “paredes” e acertar a tartaruga no ponto A_0 .

Teorema 2.1 ($x = -tg\theta$). Se θ é o ângulo que o raio laser faz com o segmento $\overline{OA_6}$ (abertura medida no sentido anti-horário, partindo de $\overline{OA_6}$), então $x = -tg\theta$ é uma raiz de $f(x)$.

Demonstração. Na poligonal verde, o ângulo formado entre OA_6 e OB_5 é θ . Perceba que este ângulo se repete em todos os demais triângulos retângulos formados entre a poligonal vermelha e verde, ou seja, são todos triângulos semelhantes. Analisando o primeiro triângulo retângulo, temos:

$$\begin{cases} OB_5 \cdot \cos\theta = OA_6 \\ OB_5 \cdot \sen\theta = A_6B_5 \end{cases} \Rightarrow \frac{OA_6}{\cos\theta} = OB_5 = \frac{A_6B_5}{\sen\theta} \Leftrightarrow A_6B_5 = \frac{\sen\theta}{\cos\theta} \cdot OA_6 = tg\theta \cdot OA_6 \quad (1)$$

Aplicando (1) analogamente à todos os triângulos retângulos, temos:

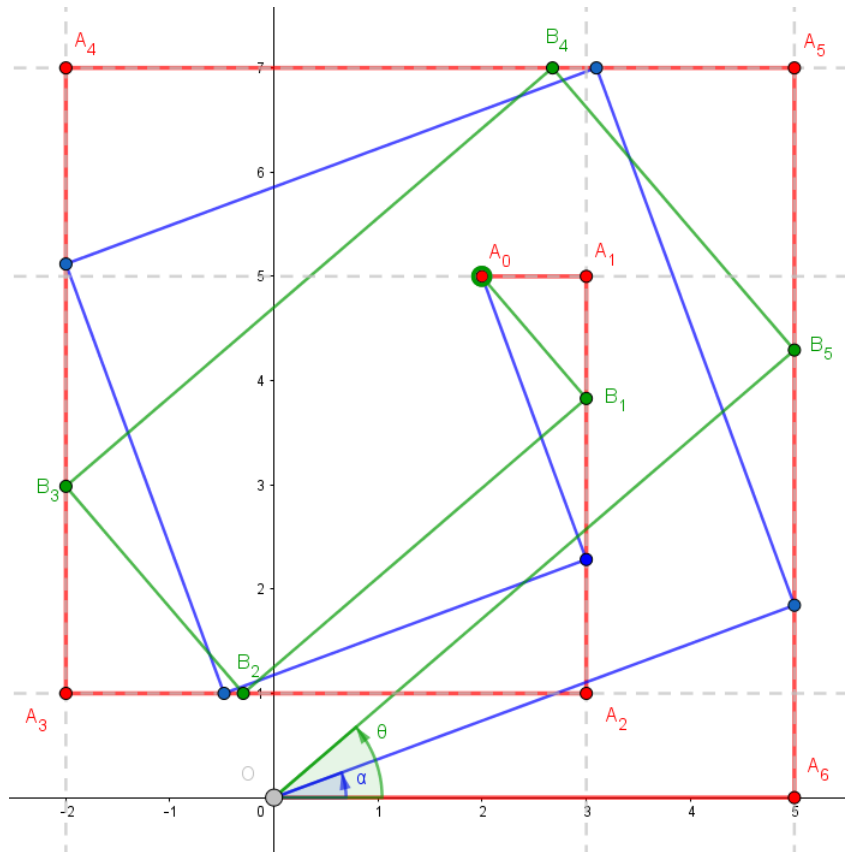


Figura 1: Caminho correspondente ao polinômio $f(x)$.

$$\left\{ \begin{array}{l}
 OB_5 = A_6B_5 \cdot \operatorname{tg}\theta \Rightarrow A_6B_5 = OA_6 \cdot (-x) \\
 A_5B_4 = A_5B_5 \cdot \operatorname{tg}\theta \Rightarrow A_5B_4 = A_5B_5 \cdot (-x) = -x(7 - A_6B_5) \\
 \qquad \qquad \qquad = -x(7 + 5x) \\
 \qquad \qquad \qquad = -7x - 5x^2 \\
 A_4B_3 = A_4B_4 \cdot \operatorname{tg}\theta \Rightarrow A_4B_3 = A_4B_4 \cdot (-x) = -x(7 - A_5B_4) \\
 \qquad \qquad \qquad = -x(7 + 7x + 5x^2) \\
 \qquad \qquad \qquad = -7x - 7x^2 - 5x^3 \\
 A_3B_2 = A_3B_3 \cdot \operatorname{tg}\theta \Rightarrow A_3B_2 = A_3B_3 \cdot (-x) = -x(6 - A_4B_3) \\
 \qquad \qquad \qquad = -x(6 + 7x + 7x^2 + 5x^3) \\
 \qquad \qquad \qquad = -6x - 7x^2 - 7x^3 - 5x^4 \\
 A_2B_1 = A_2B_2 \cdot \operatorname{tg}\theta \Rightarrow A_2B_1 = A_2B_2 \cdot (-x) = -x(5 - A_3B_2) \\
 \qquad \qquad \qquad = -x(5 + 6x + 7x^2 + 7x^3 + 5x^4) \\
 \qquad \qquad \qquad = -5x - 6x^2 - 7x^3 - 7x^4 - 5x^5 \\
 A_1A_0 = A_1A_2 \cdot \operatorname{tg}\theta \Rightarrow A_1A_0 = A_1A_2 \cdot (-x) = -x(4 - A_2B_1) \\
 \qquad \qquad \qquad = -x(4 + 5x + 6x^2 + 7x^3 + 7x^4 + 5x^5) \\
 \qquad \qquad \qquad = -4x - 5x^2 - 6x^3 - 7x^4 - 7x^5 - 5x^6
 \end{array} \right. \quad (2)$$

Sabemos ainda que $A_1A_0 = a_0 = 1$. Assim, retornamos à equação inicial:

$$A_1A_0 = -4x - 5x^2 - 6x^3 - 7x^4 - 7x^5 - 5x^6 \Rightarrow 5x^6 + 7x^5 + 7x^4 + 6x^3 + 5x^2 + 4x + 1 = 0$$

■

3 Conclusões Finais

Dada a aparente simplicidade deste método, é surpreendente o quanto o tema é pouco explorado. A pesquisa está em andamento, mas já foi possível identificar diversos desdobramentos fascinantes.

Alguns polinômios requerem abordagens diferentes, como nos casos da forma $x^n + c$, com $n \in \mathbb{N}$ e c uma constante. É notável perceber que este caso possui uma relação estreita com o famoso problema de Delos, que consiste em formular um método para encontrar um cubo cujo volume é o dobro de um cubo inicial, o qual tem aresta conhecida (utilizando apenas régua e compasso).

Em outros casos, o caminho da tartaruga pode resultar em um polígono fechado, inviabilizando (a princípio) a aplicação do método. Dan Kalman [?] demonstra a solução para esse tipo de polinômio, explicando que nessas situações o polinômio original é sempre divisível por $x^2 + 1$. De fato, ele faz uma generalização deste teorema para qualquer caminho fechado no qual o ângulo entre os segmentos (da tartaruga ou do laser) é α , em contraposição à condição inicial em que $\alpha = 90^\circ$.

Em suma, muito já se sabe sobre o método de Lill, cabendo à comunidade científica divulgá-lo, não apenas pela sua genialidade, mas para explorar as ligações entre as maneiras algébricas e geométrica de se resolver um mesmo problema.

Ainda existem questões em aberto, como apontado por Dan Kalman ou a metodologia utilizada pra encontrar raízes complexas (sim, também é possível!), que devem ser exploradas. Entretanto, este artigo está ainda em construção e tais questionamentos ainda não estão totalmente maduros para serem citados.

4 Agradecimentos

Agradeço ao PET Matemática da UFU, ao qual faço parte na condição de bolsista, diretamente vinculado ao MEC e FNDE, que atuam no pagamento de bolsas de fomento à pesquisa.

Referências Bibliográficas

- [1] Hull, T.C. Solving Cubics With Creases: The Work of Beloch and Lill. *The American Mathematical Monthly* 118.4 (2011): 307-315.
- [2] Lill, E. Résolution graphique des équations numériques de tous les degrés à une seule inconnue, et description d'un instrument inventé dans ce but. *Nouv. Annales Math. Ser. 2*, 6 (1867): 359-362.
- [3] Garcia, R. e Silva, K. Equações do terceiro grau e dobraduras de Beloch. *Revista da Olimpíada - IME - UFG*, n 14, (2019): 21-60
- [4] Kalman, D. e Verdi, M. Polynomials with Closed Lill Paths. *Mathematics Magazine*, 88:1 (2015): 3-10. Disponível em: <https://doi.org/10.4169/math.mag.88.1.3>.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Introdução às bases de Gröbner e uma aplicação Introduction to the Gröbner bases and an application

Selma Alves de Araújo

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

selma.alves@ufu.br

Cícero Carvalho

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

cicero@ufu.br

Palavras-chave

Algoritmo de Buchberger.
Bases de Gröbner.
Geometria Algébrica.
Ideais.
Sudoku.

Resumo

As bases de Gröbner são ferramentas importantes na Geometria Algébrica, na Álgebra Computacional e na Álgebra Comutativa. Faremos uma breve introdução sobre essas bases, suas propriedades e também uma aplicação no jogo Sudoku.

Keywords

Algebraic Geometry.
Buchberger algorithm.
Gröbner bases.
Ideals.
Sudoku.

Abstract

Gröbner bases are important tools on Algebraic Geometry, Computational Algebra and Commutative Algebra. We will give a brief introduction to these bases, their properties and also an application to the Sudoku game.

1 Introdução

As bases de Gröbner foram encontradas em 1965 por Bruno Buchberger (1942 –) em sua tese de doutorado e são uma ferramenta muito importante em vários campos da Álgebra. Tais bases receberam esse nome em homenagem a Wolfgang Gröbner (1899 – 1980), orientador de Buchberger. No trabalho a seguir apresentaremos uma introdução às bases de Gröbner e um método de resolução para qualquer Sudoku regular através do uso de propriedades dessas bases. A principal fonte utilizada para parte sobre bases de Gröbner deste trabalho é a referência [1].

O Sudoku é um jogo que foi criado em 1979 pelo arquiteto americano Howard Garns com o nome “Number Place”. Posteriormente, tal quebra-cabeça foi publicado no Japão em 1984, onde ganhou popularidade e passou a ser chamado de "Sudoku". Em 2004 obteve projeção mundial ao aparecer entre os passatempos do jornal britânico "The Times". A principal fonte utilizada para a aplicação das bases de Gröbner no Sudoku é a referência [2].

2 O algoritmo da divisão e variedades algébricas

Definição 2.1. Um *monômio* em X_1, \dots, X_n é um produto escrito da forma $X_1^{\alpha_1} \cdot X_2^{\alpha_2} \cdots X_n^{\alpha_n}$, onde todos os expoentes $\alpha_1, \dots, \alpha_n$ são números inteiros não negativos.

Definição 2.2. Um *polinômio* f é formado por uma combinação linear *finita* de monômios da forma $f = \sum_{\alpha} a_{\alpha} X^{\alpha}$, $a_{\alpha} \in K$. Onde a soma é sobre um número finito de n -uplas $\alpha = (\alpha_1, \dots, \alpha_n)$ e K é um corpo.

O conjunto de todos os polinômios em X_1, \dots, X_n em K é denotado por $K[X_1, \dots, X_n]$.

Definição 2.3. Uma *ordem monomial* em $K[X_1, \dots, X_n]$ é qualquer relação $>$ em $\mathbb{Z}_{\geq 0}^n$, ou equivalentemente, qualquer relação no conjunto de monômios X^{α} , $\alpha \in \mathbb{Z}_{\geq 0}^n$ satisfazendo:

- i. $>$ é de ordem total (ou linear) em $\mathbb{Z}_{\geq 0}^n$;
- ii. Se $\alpha > \beta$ e $\gamma \in \mathbb{Z}_{\geq 0}^n$, então $\alpha + \gamma > \beta + \gamma$;
- iii. A relação $>$ é bem ordenada em $\mathbb{Z}_{\geq 0}^n$. Isto é, todo subconjunto não vazio de $\mathbb{Z}_{\geq 0}^n$ tem um menor elemento sob $>$.

Definição 2.4 (Ordem Lexográfica). Sejam $\alpha = (\alpha_1, \dots, \alpha_n)$ e $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{Z}_{\geq 0}^n$. Dizemos que $\alpha >_{\text{lex}} \beta$ quando, em $\alpha - \beta \in \mathbb{Z}_{\geq 0}^n$, a entrada não nula mais à esquerda é positiva. Escreveremos $X^{\alpha} >_{\text{lex}} X^{\beta}$ se $\alpha >_{\text{lex}} \beta$.

Teorema 2.5 (Algoritmo da divisão). Fixada uma ordem monomial $>$ em $\mathbb{Z}_{\geq 0}^n$ e seja $F = (f_1, \dots, f_s)$ uma s -upla ordenada de polinômios em $K[X_1, \dots, X_n]$. Então todo $f \in K[X_1, \dots, X_n]$ pode ser escrito da forma

$$f = a_1 f_1 + \cdots + a_s f_s + r,$$

onde $a_i, r \in K[X_1, \dots, X_n]$, e $r = 0$ ou então r é uma combinação linear, com coeficientes em $K[X_1, \dots, X_n]$, de monômios tais que nenhum é múltiplo de $LM(f_i)$, para $i = 1, \dots, s$.

Demonstração. A demonstração pode ser encontrada em [1]. ■

Definição 2.6. Seja $I \subset K[X_1, \dots, X_n]$ um ideal. A **variedade algébrica** definida por I é o subconjunto de K^n definido por $V(I) := \{(a_1, \dots, a_n) \in K^n \mid f(a_1, \dots, a_n) = 0, \forall f \in I\}$.

Exemplo 2.7. Seja $I = \langle (X - Y)^2 \rangle \subset \mathbb{C}[X, Y]$ temos $V(I) = \{(a, a) \in \mathbb{C}^2 \mid a \in \mathbb{C}\}$.

Proposição 2.8. Seja $V(I)$ uma variedade algébrica. Se $\mathcal{I}(V(I))$ é o conjunto de todos os polinômios que se anulam em $V(I)$, então esse conjunto é um ideal e $I \subset \mathcal{I}(V(I))$.

Demonstração. A demonstração pode ser encontrada em [1]. ■

Exemplo 2.9. Se $I = \langle (X - Y)^2 \rangle \subset \mathbb{C}[X, Y]$ pode-se mostrar que $\mathcal{I}(V(I)) = \langle (X - Y) \rangle \supseteq I$.

Teorema 2.10 (Teorema dos Zeros de Hilbert). Se K é um corpo algebricamente fechado, então $I = \mathcal{I}(V(I))$ se, e somente se, I for radical.

Demonstração. A demonstração desse teorema, também conhecido como "Nullstellensatz" pode ser encontrada em [1]. ■

3 Bases de Gröbner e propriedades

Teorema 3.1 (Teorema da Base de Hilbert). Todo ideal $I \subset K[X_1, \dots, X_n]$ possui um conjunto gerador finito. Isto é, $I = \langle g_1, \dots, g_t \rangle$ para algum $g_1, \dots, g_t \in I$.

Demonstração. A demonstração pode ser encontrada em [1]. ■

Definição 3.2 (Bases de Gröbner). Fixada uma ordem monomial. Seja I um ideal de $K[X_1, \dots, X_n]$ tal que $I = \langle g_1, \dots, g_t \rangle$. O conjunto $G = \{g_1, \dots, g_t\}$ é uma base de Gröbner para I se todo elemento não nulo de I tem como monômio líder um múltiplo de $LM(g_i)$, para algum $i \in \{1, \dots, t\}$.

Proposição 3.3. Seja $G = \{g_1, \dots, g_t\}$ uma base de Gröbner.

- i. O resto na divisão de $a \in K[X_1, \dots, X_n]$ por $\{g_1, \dots, g_t\}$ é único;
- ii. $a \in I$ se, e somente se, o resto na divisão por $\{g_1, \dots, g_t\}$ é zero;
- iii. Um monômio M é monômio líder de algum polinômio em I se, e somente se, M é múltiplo do monômio líder de algum dos polinômios g_1, \dots, g_t .

Demonstração. A demonstração pode ser encontrada em [1]. ■

Teorema 3.4 (Algoritmo de Buchberger). *Seja um $I = \langle f_1, \dots, f_s \rangle \neq 0$ um ideal polinomial. Então uma base de Gröbner para I pode ser construída em um número finito de passos.*

Demonstração. O algoritmo e sua demonstração podem ser encontrados em [1]. ■

Definição 3.5. *Uma base de Gröbner reduzida para um ideal I é uma base de Gröbner G para I tal que*

- i. $LC(p) = 1$ para todo $p \in G$;
- ii. Para todo $p \in G$, nenhum monômio de p pertence a $\langle LT(G - \{p\}) \rangle$.

Definição 3.6 (Pegada). *Seja $I \subset K[X_1, \dots, X_n]$ um ideal. A pegada $\Delta(I)$ de I é o conjunto dos monômios M tais que $M \neq LM(f)$ para todo $f \in I \setminus \{0\}$.*

4 Resolvendo o Sudoku

Um tabuleiro de Sudoku é um quadrado 9×9 com 81 casas a serem preenchidas com números inteiros de 1 a 9, como na figura 1a. O objetivo do jogo é preencher todas as casas, a partir de uma quantidade já preenchida, de forma que cada linha, coluna e região 3×3 seja preenchida com os números de 1 a 9 sem repetição, como mostrado na figura 1b.

		7			3		5	
			9		5		8	
5	6			2	7			3
	8				2	4		9
	1			6				8
7	3	4			8			
			8	7	1			
				4		7		6
		3					9	

(a) Tabuleiro de Sudoku

1	9	7	6	8	3	2	5	4
3	4	2	9	1	5	6	8	7
5	6	8	4	2	7	9	1	3
6	8	5	1	3	2	4	7	9
2	1	9	7	6	4	5	3	8
7	3	4	5	9	8	1	6	2
9	2	6	8	7	1	3	4	5
8	5	1	3	4	9	7	2	6
4	7	3	2	5	6	8	9	1

(b) Sudoku resolvido

Figura 1: Exemplo de Sudoku

A cada uma das 81 casas do tabuleiro de Sudoku atribuiremos um inteiro do intervalo $[1, 81]$, de forma que inteiros distintos são atribuídos a casas distintas. A cada casa i , com $i = 1, \dots, 81$, associamos uma variável X_i . A ideia dessa modelagem é construir um ideal $I \in \mathbb{C}[X_1, \dots, X_{81}]$ de forma que esse ideal contenha os dados iniciais (e.g. $X_3 - 7 \in I$, $X_6 - 3 \in I$) e contenha a solução desse quebra-cabeça (e.g. $X_1 - 1 \in I$, $X_2 - 9 \in I$, etc).

Seja $F_i = \prod_{t=1}^9 (X_i - t)$, onde $i \in [1, 81] := \{1, \dots, 81\}$. Dados i e j distintos em $[1, 81]$ existe um polinômio $G_{i,j} \in \mathbb{C}[X_1, \dots, X_{81}]$ tal que $F_i - F_j = (X_i - X_j)G_{i,j}$. De fato, considerando $F_i - F_j \in (\mathbb{C}[X_j])[X_i]$, como o coeficiente de X_i em F_i é invertível, é possível dividir F_i por $X_i - X_j$ (X_j agora é considerado uma constante) obtendo $G_{i,j}, r \in (\mathbb{C}[X_j])[X_i]$ tais que $F_i - F_j = G_{i,j}(X_i - X_j) + r$ com $r = 0$ ou $\text{grau}_{X_i}(r) = 0$, mas X_j é raiz de $F_i - F_j$, então temos que $r = 0$.

Veja que X_j é uma raiz simples de $F_i - F_j$, pois não é raiz da derivada $(F_i - F_j)' = F_i'$ (com relação a X_i). Em particular, $X_i - X_j \nmid G_{i,j}$.

Seja E o conjunto dos pares ordenados de inteiros (i, j) , tais que $1 \leq i < j \leq 81$ e $(i, j) \in E$ se, e somente se, a casa i se encontra na mesma coluna, ou linha ou região 3×3 , que a casa j .

Considere agora um problema de Sudoku bem posto, ou seja, que admite uma única solução. Chamemos de $L \subset [1, 81]$ o subconjunto das casas que já aparecem preenchidas no problema apresentado, e seja $\beta_\ell \in \{1, \dots, 9\}$ o valor que aparece na casa ℓ , para todo $\ell \in L$. Denotaremos a solução única desse problema por $(\alpha_1, \dots, \alpha_{81})$. Isto é, o valor na casa i é $\alpha_i \in \{1, \dots, 9\}$, para todo $i \in [1, 81]$, de modo que $\alpha_i = \beta_i$ para todo $i \in L$.

Seja $I \subset \mathbb{C}[X_1, \dots, X_{81}]$ o ideal gerado pelo conjunto de polinômios $\{F_1, \dots, F_{81}\} \cup \{G_{i,j} \mid (i, j) \in E\} \cup \{X_\ell - \beta_\ell \mid \ell \in L\}$. Isto é,

$$I = \langle \{F_1, \dots, F_{81}\} \cup \{G_{i,j} \mid (i, j) \in E\} \cup \{X_\ell - \beta_\ell \mid \ell \in L\} \rangle$$

Para todo $i \in [1, 81]$ temos que F_i está entre os geradores, logo, as entradas de qualquer 81-upla de zeros comuns aos geradores precisam estar no conjunto $\{1, \dots, 9\}$. Como $X_\ell - \beta_\ell$ está entre os geradores, a entrada na posição ℓ será igual a β_ℓ , para todo $\ell \in L$.

Como $G_{i,j}$ está entre os geradores sempre que $(i, j) \in E$, temos de $F_i - F_j = (X_i - X_j)G_{i,j}$ que $G_{i,j}$ se anula na solução única do problema (pois $\alpha_i \neq \alpha_j$ se $(i, j) \in E$). Inversamente, suponha que $\gamma_i = \gamma_j$ para algum $(i, j) \in E$. Substituindo $X_j = \gamma_j$ em $F_i - F_j = (X_i - X_j)G_{i,j}$ temos $F_i = (X_i - \gamma_j)G_{i,j}(X_i, \gamma_j)$. Como $\gamma_i = \gamma_j$ é raiz simples de F_i , temos que $G_{i,j}(\gamma_i, \gamma_j) \neq 0$. Portanto, a única 81-upla de \mathbb{C}^{81} que é zero comum dos geradores de I é $(\alpha_1, \dots, \alpha_{81})$.

Veja que $\Delta(I)$ é finita: de fato, como F_i está entre os geradores para todo $i \in [1, 81]$, temos que os expoente de X_j num monômio de $\Delta(I)$ é menor do que 9, para todo $j \in [1, 81]$. Com isso podemos aplicar um resultado (ver o lema 8.13 em [3]) que garante que I é radical.

Como I é radical e $V(I) = (\alpha_1, \dots, \alpha_{81})$ temos que os polinômios $X_i - \alpha_i$ estão em I para todo $i \in [1, 81]$ (pois se anulam em $V(I)$). Observe que já sabíamos disso para $i \in L$.

Considere agora uma base de Gröbner reduzida \mathcal{B} para I , escolhendo ordem lexicográfica com $X_{81} > \dots > X_1$. Como $X_1 - \alpha_1 \in I$ temos que $X_1 = ML(X_1 - \alpha_1)$ é múltiplo do monômio líder de algum polinômio em \mathcal{B} que tem que ser da forma $X_1 - \alpha$ e, portanto, tem que ser $X_1 - \alpha_1$. Assim $X_1 - \alpha_1 \in \mathcal{B}$. Do mesmo modo, $X_2 = LM(X_2 - \alpha_2)$ é múltiplo do monômio líder de algum polinômio em \mathcal{B} a princípio da forma $X_2 + \alpha X_1^m + \beta$, mas como \mathcal{B} é uma base reduzida, tem que ser da forma $X_2 + \beta$, portanto, será $X_2 - \alpha_2$.

Desse modo, provamos que $X_i - \alpha_i \in \mathcal{B}$ para todo $i \in [1, 81]$. Isso mostra que se pode calcular a solução de um Sudoku bem posto utilizando um software que calcule a base de Gröbner reduzida de I segundo a ordem acima. De fato, é possível provar que a base de Gröbner reduzida de I em qualquer ordem irá conter os polinômios $X_i - \alpha_i$, com $i \in [1, 81]$, que dão a solução do Sudoku.

5 Considerações finais

As bases de Gröbner são ferramentas relativamente recentes, e que podem auxiliar em diversos problemas da Geometria Algébrica, Álgebra Computacional e Álgebra Comutativa. Além de suas aplicações na matemática pura, é possível encontrar aplicações na matemática aplicada, como mostrado nesse trabalho através de um método para a resolução do Sudoku.

Agradecimentos

Agradeço ao auxílio financeiro oferecido pela bolsa do Instituto Tim em parceria com o Instituto de Matemática Pura e Aplicada (IMPA) e a Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP). Também agradeço ao Programa de Iniciação Científica e Mestrado (PICME) oferecido pelo IMPA em parceria com o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela oportunidade de participar de uma iniciação científica com remuneração.

Referências

- [1] COX, D.; LITTLE, J.; O'SHEA, D. **Ideals, Varieties, and Algorithms: an introduction to computational algebraic geometry and commutative algebra**. 2ª edição: Springer, 1997.
- [2] STEENPASS, A. **Singular tutorial**. Disponível em: <<http://math.ipm.ac.ir/conferences/2011/cca2011/SlidesShow-Steenpass.pdf>>. Acesso em: 17 de abril de 2021.
- [3] BECKER, T.; WEISPFENNING, V. **Gröbner Bases: a computational approach to commutative algebra**. Berlim, Alemanha. 2ª edição: Springer, 1998.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Identificação de Parâmetros Envolvendo a Equação da Difusão Bidimensional com Termo Fonte via Métodos Adjuntos

Parameter Identification Involving the Bidimensional Diffusion Equation with Source Term via Adjoint Methods

Thaís Barbosa Caetano Souza

UFU, Faculdade de Engenharia Elétrica, Uberlândia, Minas Gerais, Brasil
thais.barbosa.caetano.souza@ufu.br

Alessandro Alves Santana

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
alessandro.santana@ufu.br

Resumo

Palavras-chave

Identificação de parâmetros.
Equação do Calor.
Gradientes Conjugados.
Métodos Adjuntos.
Diferenças Finitas.

Esse trabalho tem como finalidade a estimação dos parâmetros relativos às funções de difusividade térmica e de distribuição do termo fonte da Equação do Calor. Esse tipo de problema faz parte dos chamados problemas inversos, os quais podem ser resolvidos por meio de métodos baseados em gradientes. Assim, a fim de determinar estes parâmetros, foi empregado o Método dos Gradientes Conjugados, utilizando para o cálculo do vetor gradiente o Método Adjunto. Fez-se, ainda, uma comparação deste com o Método de Diferenças Finitas. Observou-se, nas simulações, que as diferenças entre as estimativas com a aplicação desses dois métodos ocorreram, especialmente, nas proximidades dos valores exatos dos coeficientes. Além disso, a utilização do Método Adjunto gerou aproximações mais precisas dos parâmetros com uma menor quantidade de iterações.

Abstract

Keywords

Parameter identification.
Heat Equation.
Conjugate Gradients.
Adjoint methods.
Finite Differences.

This work aims to estimate the parameters related to thermal diffusivity and source term distribution functions of the Heat Equation. This type of problem is part of the so-called inverse problems, which can be solved using methods based on gradients. So, in order to determine these parameters, it was used the Conjugate Gradient Method, applying the Adjoint Method for calculation of the gradient vector. A comparison was also made between this and the Finite Difference Method. It was observed, in the simulations, that the differences between the estimates with the application of these two methods occurred, especially, in the vicinity of the exact values of the coefficients. Besides, the use of the Adjoint Method has generated more precise approximations of the parameters with lesser iterations.

1 Introdução

A identificação de parâmetros, no mundo real, está presente em inúmeras situações: no cálculo da influência destrutiva dos íons H+ no tecido saudável a partir da densidade do tecido normal e do neoplásico [1]; na determinação do hipocentro de um terremoto com base nos tempos de chegada das ondas sísmicas identificados pelos sismógrafos [2]. Além dos exemplos citados, os problemas de identificação envolvendo transferências de calor têm sido historicamente estudados, considerando suas aplicações, por exemplo, nas áreas mecânica, aeroespacial e química [3].

A estimação de parâmetros faz parte dos chamados problemas inversos. Estes são relacionados à determinação de variáveis de interesse a partir de uma análise indireta [2], considerando como ponto de partida os efeitos que tais variáveis geram em um determinado modelo [4]. Esse tipo de problema é considerado quando os parâmetros de interesse são mais difíceis de quantificar experimentalmente do que os resultados deles sobre um sistema.

Nesse contexto, o projeto em questão tem como foco a identificação de parâmetros relativos à Equação da Difusão Bidimensional com Termo Fonte, comparando-se dois métodos para a obtenção dos gradientes necessários para a busca dos parâmetros ótimos: Método de Diferenças Finitas (MDF) e Método Adjunto (MA). A equação supracitada é modelada matematicamente por [3]:

$$\frac{\partial \phi}{\partial t} = \frac{\partial}{\partial x} \left(\alpha \frac{\partial \phi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\alpha \frac{\partial \phi}{\partial y} \right) + F, \quad (1)$$

sendo

$$\alpha(x, y) = \sum_{k=1}^n \beta_k C_k(x, y) \quad \text{e} \quad F(x, y, t) = \sum_{k=1}^{n_F} \beta_{F_k} C_{F_k}(x, y, t), \quad (2)$$

com domínio espacial $[0, L_x] \times [0, L_y]$ e domínio temporal $[0, T]$. Nessa formulação, ϕ indica a temperatura do objeto no ponto analisado; x e y as dimensões espaciais; t a dimensão temporal; α a difusividade térmica; F a distribuição do termo fonte, representativa da geração de calor; β_k e β_{F_k} os coeficientes relativos à α e F , respectivamente; C_k e C_{F_k} , em ordem, as funções utilizadas para a determinação de α e F ; n e n_F o número de coeficientes de α e F , respectivamente. A condição inicial foi definida como $\phi(x, y, 0) = 100 \cdot \text{sen}(\pi x) \cdot \text{sen}(\pi y)$ e considerou-se, arbitrariamente, a condição de fronteira de Neumann, com $\phi_x(0, y, t) = 0$, $\phi_x(L_x, y, t) = 0$, $\phi_y(x, 0, t) = 0$ e $\phi_y(x, L_y, t) = 0$.

O problema inverso, em que o objetivo primordial é a identificação dos parâmetros β_k e β_{F_k} , sendo C_k e C_{F_k} funções conhecidas, depende de uma solução de referência, ϕ_R , obtida, em situações reais, por meio de experimentações, a partir das quais a estimação de parâmetros é orientada. Destaca-se que, em estudos de metodologias, é comum a utilização de simulações numéricas para a geração da referência ϕ_R , prática adotada nesse trabalho [3, 4]. Com isso, os valores a serem encontrados para os parâmetros devem produzir a solução de referência quando aplicados ao problema direto, descrito pela Eq. (1) e resolvido pelo Método de Runge-Kutta (MRK) de 4ª ordem (após uma semi-discretização).

Assim, essa questão pode ser tratada como um problema de otimização, na medida em que busca-se minimizar a diferença entre a solução de referência e o resultado obtido por meio da aplicação dos

parâmetros estimados ao problema direto, considerando-se o domínio analisado. Essa diferença, é dada

pelo chamado funcional, apresentado da seguinte forma [4]:
$$\mathcal{F}(\phi) = \frac{1}{2} \int_0^T \int_0^{L_y} \int_0^{L_x} (\phi - \phi_R)^2 dx dy dt.$$

Pode-se observar que o funcional é sensível às variações nos parâmetros β_k e β_{F_k} , uma vez que as alterações deles sobre α e F resultam em mudanças em ϕ . Da mesma maneira que em um problema de otimização com restrição, o funcional, calculado a partir da Regra 1/3 de Simpson Repetida (RSR), deve ser minimizado sujeito ao chamado problema primal, apresentado na Eq. (1). Nesse sentido, o problema inverso é definido por:
$$\min_{\substack{\beta_k, k=1, \dots, n \\ \beta_{F_k}, k=1, \dots, n_F}} \mathcal{F}(\phi) \quad \text{sujeito à} \quad \frac{\partial \phi}{\partial t} = \frac{\partial}{\partial x} \left(\alpha \frac{\partial \phi}{\partial x} \right) + \frac{\partial}{\partial y} \left(\alpha \frac{\partial \phi}{\partial y} \right) + F.$$

Como trata-se de um problema de otimização multidimensional, uma vez que ϕ é dependente de α e F (e, portanto, de β_k e de β_{F_k}), o Método dos Gradientes Conjugados (MGC) foi utilizado para a resolução. Este é um método iterativo, em que a direção de busca dos parâmetros ideais, a cada iteração, é baseada no sentido oposto do vetor gradiente do funcional (sentido de minimização) [5], sendo que esse vetor pode ser determinado utilizando MDF ou MA. Assim, procura-se minimizar $\mathcal{F}(\phi)$, por meio do Método da Seção Áurea (MSA), a cada direção determinada.

2 Metodologia

2.1 Método dos Gradientes Conjugados

Conforme expresso na definição do problema de otimização, o número parâmetros a serem estimados depende do valor definido para n e para n_F . No MGC aplicado nesse projeto, baseado em [5], os subíndices α e F representam os termos relativos, respectivamente, à $\alpha(x, y)$ e $F(x, y, t)$, enquanto o superíndice w representa a iteração em questão. Com isso, g_α e g_F são os vetores gradientes do funcional (cujas componentes são as derivadas parciais de \mathcal{F} em relação a cada parâmetro β_k e β_{F_k} , nessa ordem); s_α e s_F são as direções de busca determinadas a cada iteração, as quais são, inicialmente, iguais ao oposto de g_α e g_F ; b_α e b_F são escalares cujas variações são utilizadas para a determinação do ponto de mínimo encontrado por meio do MSA; p_α e p_F são os vetores relacionados aos parâmetros encontrados, ou seja, $p_\alpha = (\beta_1, \beta_2, \dots, \beta_n)$ e $p_F = (\beta_{F_1}, \beta_{F_2}, \dots, \beta_{n_F})$. Além disso, considera-se ν_α e ν_F variáveis auxiliares. O MGC consiste na repetição dos seguintes passos:

1. Procura do menor valor do funcional na direção de busca definida por s_α e s_F , com a aplicação do MSA para $\mathcal{F}(\phi^{w+1})$, dado que $p_\alpha^{w+1} = p_\alpha^w + b_\alpha^w s_\alpha^w$ e $p_F^{w+1} = p_F^w + b_F^w s_F^w$, sendo $\phi^{w+1} = \phi(\alpha^{w+1}, F^{w+1})$, $\alpha^{w+1} = \alpha(p_\alpha^{w+1})$ e $F^{w+1} = F(p_F^{w+1})$;
2. Cálculo do gradiente do funcional por MDF ou MA, sendo que, para $\mathcal{F} = \mathcal{F}(\phi^{w+1})$,

$$g_\alpha^{w+1} = \left(\frac{\partial \mathcal{F}}{\partial \beta_1}, \frac{\partial \mathcal{F}}{\partial \beta_2}, \dots, \frac{\partial \mathcal{F}}{\partial \beta_n} \right) \quad \text{e} \quad g_F^{w+1} = \left(\frac{\partial \mathcal{F}}{\partial \beta_{F_1}}, \frac{\partial \mathcal{F}}{\partial \beta_{F_2}}, \dots, \frac{\partial \mathcal{F}}{\partial \beta_{F_{n_F}}} \right); \quad (3)$$

3. Determinação das variáveis auxiliares ν_α e ν_F pela Fórmula de Polak e Ribiere (FPR), resultando em $\nu_\alpha^{w+1} = \frac{(g_\alpha^{w+1} - g_\alpha^w)^T g_\alpha^{w+1}}{(g_\alpha^w)^T g_\alpha^w}$ e $\nu_F^{w+1} = \frac{(g_F^{w+1} - g_F^w)^T g_F^{w+1}}{(g_F^w)^T g_F^w}$;
4. Cálculo das direções de busca s_α e s_F , com a substituição dos valores encontrados nos passos anteriores em $s_\alpha^{w+1} = -g_\alpha^{w+1} + \nu_\alpha^{w+1} s_\alpha^w$ e em $s_F^{w+1} = -g_F^{w+1} + \nu_F^{w+1} s_F^w$.

Os passos descritos são repetidos a cada iteração do método até a tolerância desejada.

2.2 Método de Diferenças Finitas

Para o cálculo de g_α e g_F a partir do MDF, foi aplicada a Fórmula de Diferenças Finitas Progressiva (FDF) de Ordem 2. Considerando notação $x_i = a + i\varepsilon$ (dado um intervalo $[a, b]$), essa FDF é dada por [6]: $f'_i = \frac{-f_{i+2} + 4f_{i+1} - 3f_i}{2\varepsilon}$. A aplicação de tal relação na Eq. (3), além da adoção de $n = 2$ e $n_F = 2$, resulta em: $g_\alpha^{w+1} = \left(\frac{-\mathcal{F}(\alpha(2\varepsilon + \beta_1^{w+1}, \beta_2^{w+1}), F^{w+1}) + 4\mathcal{F}(\alpha(\varepsilon + \beta_1^{w+1}, \beta_2^{w+1}), F^{w+1}) - 3\mathcal{F}(\alpha^{w+1}, F^{w+1})}{2\varepsilon}, \frac{-\mathcal{F}(\alpha(\beta_1^{w+1}, 2\varepsilon + \beta_2^{w+1}), F^{w+1}) + 4\mathcal{F}(\alpha(\beta_1^{w+1}, \varepsilon + \beta_2^{w+1}), F^{w+1}) - 3\mathcal{F}(\alpha^{w+1}, F^{w+1})}{2\varepsilon} \right)$. O processo de obtenção de g_F^{w+1} ocorre de modo análogo ao apresentado.

Conforme pode ser percebido, com essa FDF, são necessárias 2 resoluções do problema primal a cada derivada presente no cálculo dos vetores gradientes em relação à $\alpha(x, y)$ e $F(x, y, t)$ (considerando que $\mathcal{F}(\alpha^{w+1}, F^{w+1})$ será determinado e armazenado). Isso significa que faz-se $2(n + n_F) + 1$ MRKs para o segundo passo apresentado do MGC, questão inviável especialmente quando a quantidade de parâmetros a serem estimados é alta [4]. Para contornar esse problema, a outra forma de cálculo do g_α e do g_F envolve o MA, com a introdução do chamado problema dual, o qual relaciona-se à equação adjunta, tema da próxima seção.

2.3 Método Adjunto

No caso do MA, o cálculo do gradiente provém do resultado dos problemas primal e dual, relacionado à determinação do multiplicador de Lagrange ψ . Para se chegar nas equações relativas, faz-se: 1) Lagrangiano aumentado; 2) Equações de sensibilidade; 3) Equação adjunta e condições de fronteira; 4) Equação do gradiente do funcional [3, 4]. Tais cálculos produzem a equação adjunta:

$$-\frac{\partial \psi}{\partial t} - \frac{\partial}{\partial x} \left(\alpha \frac{\partial \psi}{\partial x} \right) - \frac{\partial}{\partial y} \left(\alpha \frac{\partial \psi}{\partial y} \right) = \phi_R - \phi, \quad (4)$$

sendo a condição final $\psi(x, y, T) = 0$ ($x \in [0, L_x], y \in [0, L_y]$) e a condição de contorno dada por $\psi_x(0, y, t) = 0, \psi_x(L_x, y, t) = 0, \psi_y(x, 0, t) = 0$ e $\psi_y(x, L_y, t) = 0$. Além disso, a partir do desenvolvimento do problema dual, chega-se que cada componente dos vetores g_α e g_F é dado, respectivamente,

$$\text{por: } \delta_{\beta_k} \mathcal{F} = - \int_0^T \int_0^{L_y} \int_0^{L_x} \phi \left[\frac{\partial}{\partial x} \left(C_k \frac{\partial \psi}{\partial x} \right) + \frac{\partial}{\partial y} \left(C_k \frac{\partial \psi}{\partial y} \right) \right] dx dy dt \text{ e } \delta_{\beta_{F_k}} \mathcal{F} = - \int_0^T \int_0^{L_y} \int_0^{L_x} \psi C_{F_k} dx dy dt.$$

As equações expressas referentes a $\delta_{\beta_k} \mathcal{F}$ e $\delta_{\beta_{F_k}} \mathcal{F}$ podem ser resolvidas, assim como o próprio funcional $\mathcal{F}(\phi)$, a partir da RSR. Para isso, determina-se ψ em todo o domínio por meio da resolução da Eq. (4) utilizando métodos, como o MRK de 4º Ordem, adaptados para problemas de valor final. Destaca-se que, com o MA, independente da quantidade de parâmetros a serem estimados, é necessária apenas duas aplicações do MRK para a estimação do gradiente, diferentemente do que se observa com o cálculo de g_α e g_F a partir do MDF.

3 Resultados

Considerando a FPR; $\alpha(x, y) = \beta_1 x + \beta_2 y$; $F(x, y, t) = e^{-t} [\beta_{F_1} \text{sen}(\pi x) + \beta_{F_2} \text{sen}(\pi y)]$; $\varepsilon = 0.001$; ϕ_R gerado por $\beta_1 = 0.04$, $\beta_2 = 0.03$, $\beta_{F_1} = 0.05$ e $\beta_{F_2} = 0.07$ (soluções exatas); e a condição de parada do MGC dada por $\mathcal{F}(\phi^{w+1}) < 10^{-9}$, os resultados encontrados para β_1 , β_2 , β_{F_1} e β_{F_2} com MDF e MA estão apresentados nas Figuras 1 e 2.

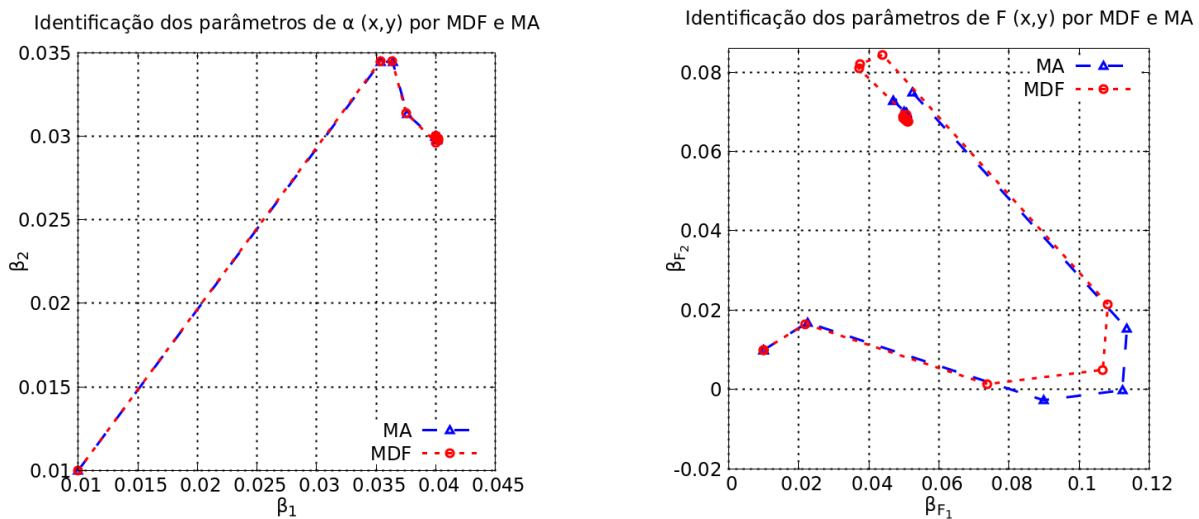


Figura 1: Resultado da estimação dos parâmetros β_1 , β_2 , β_{F_1} , β_{F_2} por MDF (com o final $\beta_1 = 0.039996$, $\beta_2 = 0.029986$, $\beta_{F_1} = 0.049809$ e $\beta_{F_2} = 0.068759$), em vermelho, e por MA (com o final $\beta_1 = 0.040000$, $\beta_2 = 0.030000$, $\beta_{F_1} = 0.049996$ e $\beta_{F_2} = 0.070053$), em azul, ao longo das iterações do MGC.

Por meio desses resultados, destaca-se a convergência da solução com a utilização do MA após 16 iterações (com um Tempo de Execução Médio (TEM) a cada iteração de (394 ± 14) s), enquanto, para o caso do MDF, parou-se manualmente os cálculos após 30 iterações (sendo o TEM de (435 ± 21) s), sem que tenha se chegado na condição de parada. Além disso, apesar das semelhanças em relação à estimação dos parâmetros nas iterações iniciais, conforme Fig. 1, a aproximação em direção à solução real torna a estimação com MDF menos precisa, devido à maior influência da difusão numérica dos métodos envolvidos, como pode ser visto na Fig. 2. Os resultados vão ao encontro ao que era esperado devido à maior quantidade de cálculos necessários ao MDF. No entanto, deve-se ponderar sobre essa conclusão a influência tanto da condição de contorno adotada, quanto de cada um dos métodos utilizados na aplicação: FPR, MRK, RSR e MSA.

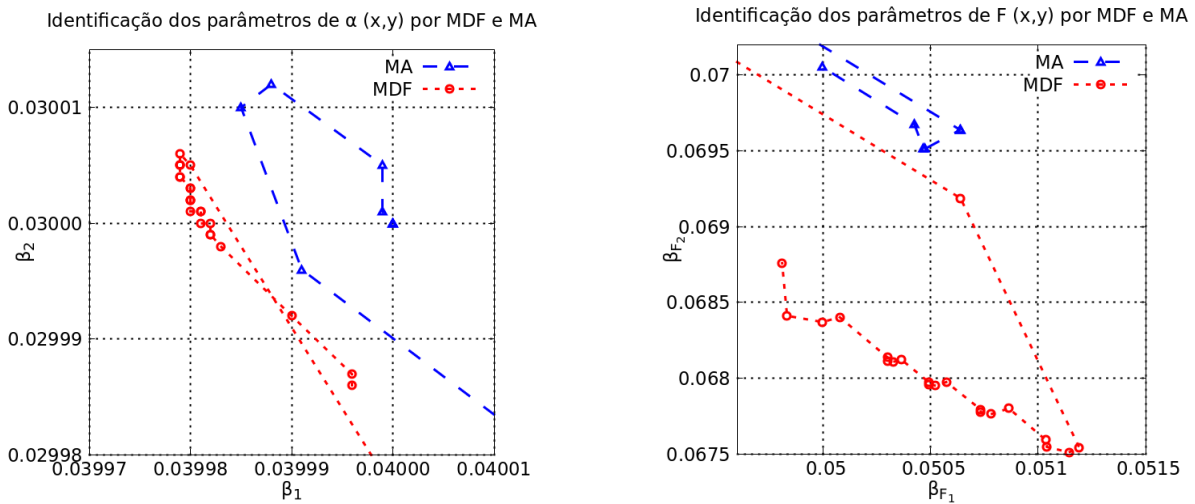


Figura 2: Ampliação do resultado da estimação dos parâmetros para as últimas iterações do MGC.

4 Conclusão

A identificação de parâmetros de equações diferenciais é encontrada em diversas aplicações, justificando a necessidade da análise dos métodos existentes para a resolução de problemas com essa tipificação. No caso estudado, destaca-se, para esse fim, que a implementação do MGC com o cálculo dos vetores gradientes por meio do MA apresentou resultados melhores do que esse cálculo realizado a partir do MDF em relação ao TEM, à quantidade de iterações e à influência da difusão numérica.

Referências

- [1] QUIROGA, A. A. I; FERNÁNDEZ, D.; TORRES, G. A.; TURNER, C.V. Adjoint method for a tumor invasion PDE-constrained optimization problem in 2D using adaptive finite element method. **Applied Mathematics and Computation**, v. 270, p. 358–368, novembro 2015.
- [2] ASTER, R. C.; BORCHERS, B.; THURBER, C. H. **Parameter Estimation and Inverse Problems**. 3ª Edição. Elsevier, 2019. ISBN 9780128046517.
- [3] OSIZIK, M. N; ORLANDE, H. R. B. **Inverse Heat Transfer: Fundamentals and Applications**. Nova York: Taylor & Francis, 2000.
- [4] SANTANA, A. A. **Identificação de parâmetros em problemas de advecção-difusão combinando a técnica do operador adjunto e métodos de volumes finitos de alta ordem**. 2007, 125 f. Tese (Doutorado em Matemática Aplicada) - Instituto de Matemática e Estatística, Universidade de São Paulo, 2007.
- [5] HEATH, M. **Scientific Computing: An Introductory Survey**. 2ª Edição Revisada. Filadélfia: SIAM, 2018.
- [6] HOFFMAN, K. A.; CHIANG, S. T. **Computational Fluid Dynamics**. Wichita: Engineering Education System, 2000. ISBN 0962373109.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O ponto e seus grupos de homologia The point and their homology groups

Alef Alves Fidelis

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

alef.fidelis222530@gmail.com

Francielle Rodrigues de Castro Coelho

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

francielle@ufu.br

Resumo

Palavras-chave

Espaço topológico.
Homologia singular.
Operador bordo.
Complexo de cadeias.
Homologia do ponto.

Este trabalho traz a construção dos grupos de homologia de um espaço topológico e o cálculo dos grupos de homologia de um espaço constituído de um único ponto. Esta construção, por sua vez, será feita de forma natural, a fim de tornar familiar ao leitor conceitos por vezes abordados somente em nível de pós-graduação. Porém, mesmo adotando esta abordagem, não obtemos uma forma efetiva, em um primeiro momento, de como efetuar os cálculos dos grupos de homologia de um espaço topológico. Com isso, necessitamos transportar um problema topológico para o universo algébrico, enxergando esses grupos de homologia como os grupos de homologia de um complexo de cadeias.

Abstract

Keywords

Topological space.
Singular Homology.
Boundary operator.
Chain complex.
Homology of a point.

This work deals with the construction of the homology groups of a topological space and the calculation of the homology groups of a space made up of a single point. This construction, in turn, will be done in a natural way, in order to make the reader familiar with concepts that are sometimes exclusive at the graduate level. However, even adopting this approach, we did not obtain an effective way, at first, of how to carry out the calculations of the homology groups of a topological space. With that, we need to transport a topological problem to the algebraic universe, seeing these homology groups as the homology groups of a chain complex.

1 Introdução

Na Topologia Algébrica, transformam-se problemas topológicos em algébricos e, ao resolver estes últimos, obtêm-se informações sobre o problema original. Na teoria de Homologia Singular de Espaços Topológicos pode-se caracterizar os espaços topológicos em termos de seus “buracos” n -dimensionais. Para um espaço topológico X , o seu grupo de homologia está associado a sua estrutura geométrica, mais especificamente está associado a maneira como os “buracos” de X estão arrançados.

Nosso objetivo neste trabalho é fazer a construção dos grupos de homologia singular de um espaço topológico qualquer X . Compreender também que podemos enxergar a sequência de grupos abelianos livres cujas bases são simplexes singulares de X junto com o operador bordo nos dá um complexo de cadeias. E reduzir o cálculo da homologia de X ao cálculo da homologia desse complexo de cadeias. Além disso, para elucidar o que apresentaremos, será feito o cálculo da homologia do ponto.

Este trabalho se baseia no estudo da referência [1] no que tange módulos, sequências exatas, complexo de cadeias e homologia de complexos de cadeias; e da referência [2] no que tange homologia de um espaço topológico e seus cálculos.

2 Conceitos preliminares

Nesta seção apresentaremos alguns conceitos de topologia geral e de álgebra de grupos, essenciais para o pleno entendimento deste trabalho.

Definição 2.1. *Uma topologia em um conjunto X não vazio é uma coleção \mathcal{T} de subconjuntos de X tais que:*

- (i) *O conjunto vazio \emptyset e o conjunto todo X pertencem à \mathcal{T} ;*
- (ii) *A união dos elementos de qualquer subcoleção de \mathcal{T} pertence à \mathcal{T} ;*
- (iii) *A interseção dos elementos de qualquer subcoleção finita de \mathcal{T} pertence à \mathcal{T} .*

O par ordenado (X, \mathcal{T}) , onde X é um conjunto não vazio e \mathcal{T} é uma topologia em X , é chamado de espaço topológico.

Definição 2.2. *Um grupo abeliano G é livre se existe um subconjunto $A \subset G$ tal que todo elemento $g \in G$ possui uma representação única*

$$g = \sum_{x \in A} n_x x,$$

onde n_x é um inteiro igual a zero exceto para uma quantidade finita de x 's.

Definição 2.3. *Um complexo de cadeias é uma sequência de grupos abelianos e homomorfismos*

$$\cdots \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots$$

na qual a composição $\partial_{n-1} \circ \partial_n = 0$, para cada $n \in \mathbb{Z}$, ou seja, $\text{Im}(\partial_n) \subset \text{Ker}(\partial_{n-1})$, para cada $n \in \mathbb{Z}$.

3 Grupos de homologia

Nesta seção faremos a apresentação de conceitos específicos de topologia algébrica, os quais utilizamos na construção dos grupos de homologia de um espaço topológico.

Definição 3.1. *Seja $A \subset \mathbb{R}^n$. A envoltória convexa de A é a interseção de todos os conjuntos convexos de \mathbb{R}^n que contém A .*

Definição 3.2. *Um p -simplexo $s \subset \mathbb{R}^n$ é a envoltória convexa de uma coleção de $p + 1$ pontos $\{x_0, \dots, x_p\} \subset \mathbb{R}^n$, no qual os pontos $x_i - x_0$, com $i \in \{1, \dots, p\}$, formam um conjunto linearmente independentes.*

Note que a definição de um p -simplexo é independente da designação de qual ponto é x_0 .

Proposição 3.3. *Seja $\{x_0, \dots, x_p\} \subset \mathbb{R}^n$. Então, são equivalentes:*

(a) $x_1 - x_0, \dots, x_p - x_0$ são linearmente independentes.

(b) Se $\sum_{i=0}^p s_i x_i = \sum_{i=0}^p t_i x_i$ e $\sum_{i=0}^p s_i = \sum_{i=0}^p t_i$, com $s_i, t_i \in \mathbb{R}$, então $s_i = t_i$, para todo $i \in \{0, \dots, p\}$.

Demonstração. Veja [2, p. 1]. ■

Agora, seja s um p -simplexo em \mathbb{R}^n e considere o conjunto de todos os pontos da forma $t_0 x_0 + \dots + t_p x_p$, onde $\sum_{i=0}^p t_i = 1$ e $t_i \geq 0$, para cada $i \in \{0, \dots, p\}$. Note que esse conjunto de pontos é a envoltória convexa do conjunto $\{x_0, \dots, x_p\}$ e portanto, pela Proposição 3.3, temos:

Proposição 3.4. *Se o p -simplexo s é a envoltória convexa de $\{x_0, \dots, x_p\}$, então cada ponto de s possui uma única representação, distintas entre si, na forma $\sum_{i=0}^p t_i x_i$, com $t_i \in \mathbb{R}$, onde $t_i \geq 0$, para*

cada $i \in \{0, \dots, p\}$, e $\sum_{i=0}^p t_i = 1$.

Demonstração. Veja [2, p. 2]. ■

Os pontos x_i são os vértices de s . A Proposição 3.4 nos permite associar os pontos de s com $(p + 1)$ -uplas (t_0, \dots, t_p) , com uma escolha adequada das coordenadas t_i . Por sua vez, se os vértices receberem uma ordem específica, então s é um simplexo ordenado.

Definição 3.5. *Seja s um simplexo ordenado com vértices x_0, \dots, x_p . Defina σ_p como sendo o conjunto de todos os pontos $(t_0, \dots, t_p) \in \mathbb{R}^{p+1}$ com $\sum_{i=0}^p t_i = 1$ e $t_i \geq 0$, para cada $i \in \{0, \dots, p\}$.*

Se a função

$$f: \sigma_p \longrightarrow s$$

é dada por $f(t_0, \dots, t_p) = \sum_{i=0}^p t_i x_i$, então f é contínua. Além disso, da unicidade dos representantes e do fato de σ_p e s serem espaços compactos e de Hausdorff, segue que f é um homeomorfismo. Logo, cada p -simplexo ordenado é uma imagem homeomórfica natural de σ_p . Note também que σ_p é um p -simplexo com vértices $x'_0 = (1, 0, \dots, 0)$, $x'_1 = (0, 1, \dots, 0)$, \dots , $x'_p = (0, 0, \dots, 1)$. Nos referenciaremos à σ_p como um p -simplexo usual com a ordem natural.

Definição 3.6. *Seja X um espaço topológico. Um p -simplexo singular em X é uma função contínua*

$$\varphi: \sigma_p \longrightarrow X.$$

Note que os 0-simplexos singulares podem ser identificados com os pontos de X , enquanto que os 1-simplexos singulares com os caminhos em X .

Definição 3.7. *Se φ é um p -simplexo singular e i um inteiro com $0 \leq i \leq p$, defina a i -ésima face de φ , $\partial_i(\varphi)$, como um $(p-1)$ -simplexo singular em X dado por*

$$[\partial_i(\varphi)](t_0, \dots, t_{p-1}) = \varphi(t_0, \dots, t_{i-1}, 0, t_i, \dots, t_{p-1}).$$

Definição 3.8. *Se X é um espaço topológico, defina $S_n(X)$ como sendo o grupo abeliano livre cuja base é o conjunto de todos os n -simplexos singulares de X . Um elemento de $S_n(X)$ é uma n -cadeia singular de X e possui a forma*

$$\sum_{\varphi} n_{\varphi} \varphi,$$

onde n_{φ} é um inteiro igual a zero exceto para uma quantidade finita de φ 's.

Como a i -ésima face ∂_i é uma função do conjunto dos n -simplexos singulares no conjunto dos $(n-1)$ -simplexos singulares, existe uma única extensão para o homomorfismo

$$\partial_i: S_n(X) \longrightarrow S_{n-1}(X),$$

dada por $\partial_i \left(\sum_{\varphi} n_{\varphi} \varphi \right) = \sum_{\varphi} n_{\varphi} \partial_i(\varphi)$.

Definição 3.9. *O operador bordo é o homomorfismo*

$$\partial: S_n(X) \longrightarrow S_{n-1}(X),$$

dado por

$$\partial = \partial_0 - \partial_1 + \partial_2 + \dots + (-1)^n \partial_n = \sum_{i=0}^n (-1)^i \partial_i.$$

Proposição 3.10. A composição $\partial \circ \partial$ em

$$S_n(X) \xrightarrow{\partial} S_{n-1}(X) \xrightarrow{\partial} S_{n-2}(X)$$

é o homomorfismo nulo.

Demonstração. Veja [2, p. 4]. ■

Definição 3.11. Um elemento $c \in S_n(X)$ é um n -ciclo se $\partial(c) = 0$. Um elemento $d \in S_n(X)$ é um n -bordo se $d = \partial(e)$, para algum $e \in S_{n+1}(X)$. Como ∂ é um homomorfismo, seu núcleo (o conjunto de todos os n -ciclos) é um subgrupo de $S_n(X)$, denotado por $Z_n(X)$. De forma similar, a imagem de ∂ em $S_n(X)$ é o subgrupo $B_n(X)$ de todos os n -bordos.

Note que, pela Proposição 3.10, $B_n(X) \subset Z_n(X)$.

Definição 3.12. Seja X um espaço topológico. Consideremos $\{S_n(X)\}$, $\partial_n: S_n(X) \rightarrow S_{n-1}(X)$, $Z_n(X)$ e $B_n(X)$. O n -ésimo grupo de homologia singular de X é

$$H_n(X) = Z_n(X)/B_n(X).$$

Temos que $\{S_n(X), \partial_n\}$ é um complexo de cadeias e o grupo de homologia de X é a homologia desse complexo de cadeias.

4 Os grupos de homologia do ponto

Nesta seção apresentaremos o cálculo dos grupos de homologia do ponto.

Seja X um espaço topológico formado por um único ponto. Então, para cada $p \in \mathbb{Z}$, com $p \geq 0$, existe um único p -simplexo singular $\varphi_p: \sigma_p \rightarrow X$. Note também que para $p > 0$, temos $\partial_i(\varphi_p) = \varphi_{p-1}$. Considere o seguinte complexo de cadeias:

$$\cdots \rightarrow S_2(X) \rightarrow S_1(X) \rightarrow S_0(X) \rightarrow 0.$$

Cada $S_n(X)$ é um grupo cíclico infinito gerado por φ_n . O operador bordo é dado por

$$\partial(\varphi_n) = \sum_{i=0}^n (-1)^i \partial_i(\varphi_n) = \sum_{i=0}^n (-1)^i \varphi_{n-1}.$$

Logo, $\partial(\varphi_{2n-1}) = 0$ e $\partial(\varphi_{2n}) = \varphi_{2n-1}$, para $n \in \mathbb{Z}$, com $n > 0$. Perceba que a parcela i está associada com a parcela $2n - 1 - i$, ou seja, elas se anulam. Assim, para $0 \leq i \leq 2n$, a parcela $2n$ não se anulará, logo $\partial(\varphi_{2n}) = \varphi_{2n-1}$, e além disso, para $0 \leq i \leq 2n - 1$, todas as parcelas se anulam, logo $\partial(\varphi_{2n-1}) = 0$.

Aplicando isso no complexo de cadeias, é evidente que

$$Z_n(X) = B_n(x),$$

para $n > 0$. Contudo, $Z_0(X) = S_0(X)$ é cíclico infinito, onde $B_0(X) = 0$. Além disso, concluímos que os grupos de homologia do ponto são dados por

$$H_n(X) = \begin{cases} \mathbb{Z}, & \text{se } n = 0, \\ 0, & \text{se } n > 0. \end{cases}$$

5 Conclusão

Pode-se pensar, em um primeiro momento, que o cálculo dos grupos de homologia do ponto é algo muito básico. E de fato, este é o segundo espaço topológico mais simples que existe (o primeiro é o espaço vazio), porém, com mais alguns resultados, ficaremos hábeis à calcular os grupos de homologia de vários espaços reduzindo-os ao do ponto. Um resultado que nos dá isso de imediato diz que se X e Y são espaços topológicos tal que existe uma equivalência homotópica entre eles, então seus grupos de homologia são isomorfos em todos os níveis.

Nota-se que com poucos resultados de topologia e álgebra, somos capazes de construir o início do que é chamado de Teoria de Homologia Singular de Espaços Topológicos. Conceitos esses que são abordados somente em cursos de pós-graduação no Brasil. Além disso, podemos ver que somente com esses resultados iniciais, foi possível calcular os grupos de homologia do ponto e, como comentado no parágrafo anterior, podemos calcular os grupos de homologia de qualquer espaço topológico que possuir uma equivalência homotópica com o ponto, a saber, por exemplo, qualquer subespaço convexo de \mathbb{R}^n .

Agradecimentos

Na condição de bolsista do PET Matemática da Universidade Federal de Uberlândia, agradeço o PET-SESu-MEC.

Referências

- [1] HU, Sze-tsen. **Introduction to Homological Algebra**. San Francisco: Holden-Day, 1968.
- [2] VICK, J.W. **Homology Theory**. 2ª Edição. New York: Springer Verlag, 1994.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Uma representação do número π em Frações Contínuas A representation of the number π in Continuous Fractions

Aloisio da Silva Teixeira

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
aloisio.teixeira@ufu.br

Lígia Laís Fêmina

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
ligia@ufu.br

Palavras-chave

Frações contínuas.
Algoritmo de Euclides.
Números racionais.
Número irracionais.
Número Pi.

Resumo

A teoria de frações contínuas é um dos temas mais interessantes da matemática. Seu estudo data de pelo menos três séculos antes de Cristo, iniciando com o Algoritmo de Euclides para o máximo divisor comum. O objetivo desse trabalho é entender como representar números racionais e irracionais através de frações contínuas, em particular apresentar uma representação do número π em frações contínuas.

Keywords

Continuous fractions.
Euclid's Algorithm.
Rational numbers.
Irrational numbers .
Pi number.

Abstract

The theory of continuous fractions is one of the most interesting topics in mathematics. His study dates back at least three centuries before Christ, starting with Euclid's Algorithm for the greatest common divisor. The objective of this work is to understand how to represent rational and irrational numbers through continuous fractions, in particular to present a representation of the number π in continuous fractions.

1 Introdução

O número π , com duas casas decimais vale 3,14. Se queremos achar uma fração equivalente e irredutível a este valor, podemos fazer $3,14 = \frac{314}{100} = \frac{157}{50}$. Mas veja que os antigos utilizavam a fração irredutível $\frac{22}{7}$, com o mesmo grau de precisão. Esta última fração é resultado de uma fração contínua que, de modo geral, fornece uma aproximação de um número irracional com uma precisão maior ou igual do que aquelas aproximações racionais com denominadores de base 10. Outra particularidade das frações contínuas é que as mesmas nos garante sempre representações finitas de números racionais em contrapartida às dízimas periódicas. Neste trabalho, veremos o que é uma fração contínua, o modo como a mesma se relaciona com o algoritmo de Euclides e como podemos usá-las para representar números reais racionais e irracionais.

2 Conceitos Fundamentais

Nessa seção veremos as principais definições para o entendimento do trabalho, para maiores detalhes consultar [1].

Definição 2.1. Seja x um número real positivo e m um inteiro positivo tais que $m < x < m + 1$. Sejam agora $a_0, a_1, a_2, \dots, a_n$ inteiros positivos. Fração contínua é a expressão

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_n}}}$$

também denotada por $[a_0; a_1, a_2, \dots, a_n]$ que representa o número real x de forma exata ou aproximada, caso x seja racional ou irracional, respectivamente.

Observação 2.2 . Pela definição, o número real x pode ser representado por uma fração imprópria, ou seja, $x = \frac{a}{b}$, com $a, b \in \{2, 3, \dots\}$, $a > b$ e $b \nmid a$. Portanto, $x > 1$ e não é inteiro.

Denotaremos por $[x]$ a parte inteira de x e por $\{x\}$ a parte fracionária de x .

Definição 2.3. (Parte inteira e parte fracionária do número real x) Dados k inteiro e $d_1, d_2, d_3, \dots \in \{0, 1, 2, \dots, 9\}$, todo número real x pode ser representado na forma decimal, como $x = k, d_1d_2d_3\dots$. O número inteiro k é único de forma que $k < x < k + 1$. Como $x = k + 0, d_1d_2d_3\dots$, definimos $[x] = k$ e $\{x\} = 0, d_1d_2d_3\dots$

Caso x seja racional, temos a seguinte particularidade.

Definição 2.4. Parte inteira e parte fracionária do número racional $x = \frac{a}{b}$, conforme observação 2.2. Pelo Algoritmo de Euclides, existem inteiros positivos q e r tais que $a = bq + r$, com $0 \leq r < b$.

Definimos $[x] = q$ e $\{x\} = \frac{r}{b}$ que é equivalente com a definição 2.3.

Observação 2.5 . Note que temos as relações $[x] = x - \{x\}$ e $\{x\} = x - [x]$.

3 Montagem de frações contínuas

3.1 Fração contínua de número racional

Para eliminar as trivialidades e sem prejuízo de generalidade, interessam os números racionais positivos que representam frações impróprias de números que não são inteiros.

Seja $x = \frac{a}{b}$, com $a, b \in \{2, 3, \dots\}$, $a > b$ e $b \nmid a$. Pelo Algoritmo de Euclides, existem inteiros positivos q_1 e r_1 tais que $a = bq_1 + r_1$, com $0 < r_1 < b$. Logo,

$$x = \frac{a}{b} = q_1 + \frac{r_1}{b}$$

No segundo membro desta igualdade, temos $0 < \frac{r_1}{b} < 1$ (fração própria). Mas, $\frac{b}{r_1} > 1$ (fração imprópria). Então, podemos escrever

$$x = \frac{a}{b} = q_1 + \frac{1}{\frac{b}{r_1}}$$

Novamente, pelo algoritmo de Euclides e em relação a fração imprópria $\frac{b}{r_1} > 1$, existem inteiros positivos q_2 e r_2 tais que $b = r_1q_2 + r_2$, com $0 \leq r_2 < r_1$. Então, podemos reescrever

$$x = \frac{a}{b} = q_1 + \frac{1}{q_2 + \frac{r_2}{r_1}}$$

Procedendo de forma análoga, ou seja, aplicando o algoritmo de Euclides na fração imprópria $\frac{1}{\frac{r_2}{r_1}} = \frac{r_1}{r_2} > 1$ e assim sucessivamente, obteremos restos $r_1 > r_2 > r_3 \dots$ cada vez menores de tal forma que existirá um índice n , onde $r_{n+1} = 0$ com o qual se finda a construção da fração contínua, com a mesma assumindo o aspecto

$$x = \frac{a}{b} = q + \frac{1}{r_1 + \frac{1}{r_2 + \dots + \frac{1}{r_n}}}$$

Observação 3.1.1. O número de coeficientes $a_0 = q, a_1 = r_1, a_2 = r_2, \dots, a_n = r_n$ de uma fração contínua que representa um número racional é finito. Portanto, é interessante que um número racional pode ter representação decimal infinita, como é o caso das dízimas periódicas, no entanto, sua representação em forma de fração contínua é sempre finita.

Exemplo 3.1.2.

$$\frac{35}{8} = 4 + \frac{3}{8} = 4 + \frac{1}{8/3} = 4 + \frac{1}{2 + \frac{2}{3}} = 4 + \frac{1}{2 + \frac{1}{3/2}} = 4 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2}}}$$

ou

$$\frac{35}{8} = [4; 2, 1, 2]$$

Como $\frac{35}{8} = 4,375$, veja que o ponto e vírgula, assim como a vírgula, tem o papel de separar a parte inteira da fracionária.

3.2 Fração contínua de número irracional

Seja $x = \alpha_0 > 1$ um número irracional. A representação decimal deste número é uma composição de sua parte inteira com sua parte fracionária, ou seja, $\alpha_0 = [\alpha_0], \{\alpha_0\}$ o que indica

$$\alpha_0 = [\alpha_0] + \{\alpha_0\},$$

com $[\alpha_0] > 1$ inteiro e $0 < \{\alpha_0\} < 1$ irracional.

Mas veja que $\alpha_1 = \frac{1}{\{\alpha_0\}} > 1$ é também irracional. Segue que $\alpha_1 = [\alpha_1] + \{\alpha_1\}$. Logo,

$$\alpha_0 = [\alpha_0] + \{\alpha_0\} = [\alpha_0] + \frac{1}{1/\{\alpha_0\}} = [\alpha_0] + \frac{1}{\alpha_1} = [\alpha_0] + \frac{1}{[\alpha_1] + \{\alpha_1\}}$$

Por sua vez, $\alpha_2 = \frac{1}{\{\alpha_1\}} > 1$ é irracional. Prossegue que $\alpha_2 = [\alpha_2] + \{\alpha_2\}$. Logo,

$$\alpha_0 = [\alpha_0] + \frac{1}{[\alpha_1] + \{\alpha_1\}} = [\alpha_0] + \frac{1}{[\alpha_1] + \frac{1}{1/\{\alpha_1\}}} = [\alpha_0] + \frac{1}{[\alpha_1] + \frac{1}{\alpha_2}} = [\alpha_0] + \frac{1}{[\alpha_1] + \frac{1}{[\alpha_2] + \{\alpha_2\}}}$$

Repetindo este processo com $\alpha_i = \frac{1}{\alpha_{i-1}} > 1$, $i \in \{0, 1, 2, \dots, n\}$, um número finito de vezes, digamos, n vezes, temos a fração contínua

$$\alpha_0 = [\alpha_0] + \frac{1}{[\alpha_1] + \frac{1}{[\alpha_2] + \dots + \frac{1}{[\alpha_n]}}}$$

Esta expressão é uma boa aproximação para o número irracional $x = \alpha_0 > 1$. Dizemos, então que

$$x : \lim_{n \rightarrow \infty} [[\alpha_0]; [\alpha_1], [\alpha_2], \dots, [\alpha_n]]$$

representa, de fato, o número real $x = a_0$.

Exemplo 3.2.1. Vamos aproximar o número π com $n = 3$.

$$\begin{aligned}\pi &= 3,141\dots = 3 + 0,141\dots = 3 + \frac{1}{7,062\dots} = 3 + \frac{1}{7 + 0,062\dots} = 3 + \frac{1}{7 + \frac{1}{15,996\dots}} = \\ &= 3 + \frac{1}{7 + \frac{1}{15 + 0,996\dots}} = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1,003\dots}}} = 3 + \frac{1}{7 + \frac{1}{15 + \frac{1}{1 + 0,003\dots}}} \simeq \frac{355}{113} = 3,1415929204\end{aligned}$$

Ótima precisão até a sexta casa decimal! Com $[\alpha_0 = 3], [\alpha_1 = 7], [\alpha_2 = 15]$ e $[\alpha_3 = 1]$. Assim, $\pi \simeq [3; 7, 15, 1]$.

4 Considerações finais

O estudo das frações contínuas é importante em vários aspectos, não só em questão de representação e computação de números como também serve de base para estudos avançados em Teoria dos Números e Análise.

Referências

- [1] Moreira, C. **Representações de Números e Aproximações Diofantinas**. I Colóquio da Região Sudeste. Primeira edição. São Paulo: IMPA, 2011.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O Teorema da Aplicação Aberta e a essencialidade de suas hipóteses The Open Application Theorem and the essentiality of its hypotheses

Ana Laura Mendonça Marangoni

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
almmarangoni@gmail.com

Fábio José Bertoloto

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
bertoloto@ufu.br

Palavras-chave

Transformações lineares contínuas.
Aplicação aberta.
Espaços de Banach.
Interior de um conjunto.
Fecho de um conjunto.

Resumo

Neste trabalho é apresentada uma demonstração do Teorema da Aplicação Aberta utilizando o Teorema de Baire e, portanto, os conceitos de espaços topológicos de primeira categoria e segunda categoria. Também é demonstrado que a hipótese de sobrejetividade de uma aplicação, bem como o fato desta aplicação se dar entre espaços de Banach, são essenciais na obtenção do resultado do teorema principal.

Keywords

Continuous linear transformations.
Open application.
Banach spaces.
Interior of a set.
Closure of a set.

Abstract

This paper presents a demonstration of the Open Application Theorem using Baire's Theorem and, therefore, the concepts of topological spaces of first and second category. It is also proved that the hypothesis of surjectivity of the application, as well as the fact that this application occurs between Banach spaces, are both essential in obtaining the main theorem's result.

1 Introdução

Estudos sobre Análise Funcional começaram no início do século XX e têm suas raízes históricas marcadas por trabalhos sobre transformações e equações diferenciais. Por volta de 1920, quando foram publicados o Teorema de Hanh-Banach e sua demonstração, a teoria ganhou força, devido ao ilustre austríaco Hans Hanh e ao celebrado polonês Stefan Banach. Além destes, outros grandes matemáticos marcaram a trajetória da Análise Funcional, como Hugo Steinhaus, com o Princípio da Limitação Uniforme (ou Teorema de Banach-Steinhaus), e Juliusz Pawel Schauder, com o Teorema da Aplicação Aberta (ou Teorema de Banach-Schauder) - o qual será o foco deste trabalho.

O Teorema da Aplicação Aberta, demonstrado por Banach e Schauder em 1929, classifica funções abertas entre espaços de Banach e garante que todo operador linear, contínuo e sobrejetor $T: E \rightarrow F$ é uma aplicação aberta, sempre que E e F forem espaços de Banach. Ou seja, $T(A)$ é aberto em F sempre que A for aberto em E . Uma importante consequência deste resultado é o também bastante conhecido Teorema do Gráfico Fechado (ver [1, p. 34]).

2 Uma demonstração do Teorema da Aplicação Aberta

Para a compreensão do Teorema da Aplicação Aberta e de sua demonstração, apresentaremos, a priori, o seguinte resultado.

Lema 2.1. *Sejam E e F espaços de Banach e seja $T \in L(E; F)$. Então, as seguintes afirmações são equivalentes:*

- (a) T é sobrejetora.
- (b) $\overline{T(B_E(0; 1))} \supset B_F(0; \delta)$, para algum $\delta > 0$.
- (c) $T(B_E(0; 1)) \supset B_F(0; \delta)$, para algum $\delta > 0$.

Demonstração. Devemos mostrar a equivalência das afirmações acima.

Suponha T sobrejetora. De $E = \bigcup_{n=1}^{\infty} B_E(0; n)$, vale que:

$$F = T(E) = T\left(\bigcup_{n=1}^{\infty} B_E(0; n)\right) = \bigcup_{n=1}^{\infty} T(B_E(0; n)) = \bigcup_{n=1}^{\infty} \overline{T(B_E(0; n))}.$$

Pelo Teorema de Baire (ver [2], p. 54), F é de segunda categoria em si mesmo, isto é, existe $n \in \mathbb{N}$, tal que o conjunto $\overline{T(B_E(0; n))}$ tem interior não vazio. Então, existem $b \in F$ e $r > 0$, tais que $B_F(b; r) \subset \overline{T(B_E(0; n))}$. Pela simetria de $\overline{T(B_E(0; n))}$, segue que $B_F(-b; r) \subset \overline{T(B_E(0; n))}$. Pela convexidade de $\overline{T(B_E(0; n))}$, vale que

$$\overline{T(B_E(0; n))} \supset \text{co}(B_F(b; r) \cup B_F(-b; r)) \supset B_F(0; r).$$

Logo, $\overline{T(B_E(0; 1))} \supset B_F(0; \frac{r}{n})$. De fato, seja $y \in B_F(0; \frac{r}{n})$. Então, $ny \in B_F(0; r) \subset \overline{T(B_E(0; n))}$. Logo, $ny = \lim y_j, y_j \in T(B_E(0; n))$, isto é, $ny = \lim T(x_j), x_j \in B(0; n)$. Daí,

$$y = \lim T\left(\frac{x_j}{n}\right), x_j \in B(0; 1) \Rightarrow y \in \overline{T(B_E(0; 1))},$$

o que prova **(b)**.

Suponha, agora, que $\overline{T(B_E(0; 1))} \supset B_F(0; \delta)$, para algum $\delta > 0$. Então, podemos afirmar que

$$\overline{T\left(B_E\left(0; \frac{1}{2^n}\right)\right)} \supset B_F\left(0; \frac{\delta}{2^n}\right), \text{ para cada } n.$$

A partir disso, provaremos que

$$T(B_E(0; 1)) \supset B_F\left(0; \frac{\delta}{2}\right).$$

Seja $y \in B_F(0; \frac{\delta}{2}) \subset \overline{T(B_E(0; \frac{1}{2}))}$. Logo, existe $x_1 \in B_E(0; \frac{1}{2})$ tal que

$$y - T(x_1) \in B_F\left(0; \frac{\delta}{2^2}\right) \subset \overline{T\left(B_E\left(0; \frac{1}{2^2}\right)\right)}.$$

Logo, existe $x_2 \in B_E(0; \frac{1}{2^2})$ tal que

$$y - T(x_1) - T(x_2) \in B_F\left(0; \frac{\delta}{2^3}\right) \subset \overline{T\left(B_E\left(0; \frac{1}{2^3}\right)\right)}.$$

Logo, existe $x_3 \in B_E(0; \frac{1}{2^3})$ tal que

$$y - T(x_1) - T(x_2) - T(x_3) \in B_F\left(0; \frac{\delta}{2^4}\right) \subset \overline{T\left(B_E\left(0; \frac{1}{2^4}\right)\right)}.$$

Procedendo indutivamente, obtemos uma sequência (x_n) em E tal que

$$x_n \in B_E\left(0; \frac{1}{2^n}\right) \text{ e } y - \sum_{j=1}^n T(x_j) \in B_F\left(0; \frac{\delta}{2^{n+1}}\right), \text{ para cada } n.$$

Como $\sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=1}^{\infty} 2^{-n} = 1$ e E é de Banach, segue que

$$\sum_{n=1}^{\infty} x_n \in B_E(0; 1) \text{ e } T\left(\sum_{n=1}^{\infty} x_n\right) = y,$$

o que prova **(c)**.

Suponha, finalmente, que $T(B_E(0; 1)) \supset B_F(0; \delta)$, para algum $\delta > 0$. Observe que $nT(B_E(0; 1)) \supset$

$B_F(0; n\delta)$ e

$$\begin{aligned} nT(B_E(0; 1)) &= \{nT(x); x \in B(0; 1)\} \\ &= \{T(nx); x \in B(0; 1)\} \\ &= \{T(z); z \in B(0; n)\} \subset T(E). \end{aligned}$$

Portanto,

$$F = \bigcup B_F(0; n\delta) \subset T(E) \subset F,$$

isto é, T é sobrejetora e acabamos. ■

Teorema 2.2. (Teorema da Aplicação Aberta) *Sejam E e F espaços de Banach e seja $T: E \rightarrow F$ linear, contínuo e sobrejetor. Então, T é uma aplicação aberta. Em particular, todo operador linear contínuo e bijetor entre espaços de Banach é um isomorfismo topológico.*

Demonstração. Tomemos A um aberto em E . Queremos mostrar que $T(A)$ é aberto em F .

Seja $y \in T(A)$, então $y = T(x)$, $x \in A$. Portanto, existe $\epsilon > 0$ tal que $B_E(x; \epsilon) \subset A$, isto é, $y \in T(B_E(x; \epsilon)) \subset T(A)$.

Vamos mostrar, agora, que $\frac{T(B_E(x; \epsilon)) - \{T(x)\}}{\epsilon} = T(B_E(0; 1))$. Dividimos a demonstração:

1. Se $z \in B_E(x; \epsilon)$, então $T(z) - T(x) = T(z - x) \in T(B_E(0; \epsilon))$, pois $z - x \in B_E(0; \epsilon)$.
Provamos, então, que $\frac{T(B_E(x; \epsilon)) - \{T(x)\}}{\epsilon} \subset T(B_E(0; 1))$.
2. Observe, agora, que $T(B_E(0; \epsilon)) \subset T(B_E(x; \epsilon)) - \{T(x)\}$. De fato, se $y \in T(B_E(0; \epsilon))$, então $y = T(s)$, onde $s \in B_E(0; \epsilon)$, ou seja, $\|s\| = \|(s + x) - x\| < \epsilon$. Logo, $s + x \in B_E(x; \epsilon)$ e podemos concluir que $s \in B_E(x; \epsilon) - \{x\}$. Como T é linear, segue que $y = T(s) \in T(B_E(x; \epsilon)) - \{T(x)\}$.

Por fim, provamos que $T(B_E(0; 1)) = \frac{T(B_E(x; \epsilon)) - \{T(x)\}}{\epsilon}$ ou, equivalentemente, que $T(B_E(0; \epsilon)) = T(B_E(x; \epsilon)) - \{T(x)\}$.

Pelo Lema 2.1, como T é sobrejetora, resulta que $B_F(0; \delta) \subset T(B_E(0; 1))$, para algum $\delta > 0$. Segue que $B_F(0; \delta) \subset T(B_E(0; 1)) = \frac{T(B_E(x; \epsilon)) - \{T(x)\}}{\epsilon}$. Daí,

$$\begin{aligned} B_F(0; \epsilon\delta) \subset T(B_E(x; \epsilon)) - \{T(x)\} &\Rightarrow B_F(T(x); \epsilon\delta) \subset T(B_E(x; \epsilon)) \subset T(A) \\ &\Rightarrow B_F(y; \epsilon\delta) \subset T(B_E(x; \epsilon)) \subset T(A). \end{aligned}$$

Portanto, $y \in \text{int}(T(A))$ e podemos concluir que $T(A)$ é aberto em F . ■

3 Essencialidade das Hipóteses

O Teorema da Aplicação Aberta exige que o operador $T: E \rightarrow F$, definido entre espaços de Banach, seja linear, contínuo e sobrejetor. O resultado e o exemplo a seguir ressaltam a essencialidade de duas das hipóteses desse teorema: a sobrejetividade de T e a completude dos espaços E e F . No que segue, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Proposição 3.1. *Toda aplicação linear aberta $T: E \rightarrow F$ é sobrejetora sempre que E e F são espaços normados.*

Demonstração. Seja $r > 0$ e considere $B_E(0; r)$. Como T é aberta, segue que, para algum $\delta > 0$, $T(B_E(0; r)) \supset B_F(0; \delta)$. Isto é, $nT(B_E(0; r)) \supset B_F(0; n\delta)$. Como T é linear, vale que:

$$\begin{aligned} nT(B_E(0; r)) &= \{nT(x); x \in B_E(0; r)\} \\ &= \{T(nx); nx \in B_E(0; nr)\} \\ &= \{T(z); z \in B_E(0; nr)\} \subset T(E). \end{aligned}$$

Portanto, $F = \bigcup B_F(0; n\delta) \subset T(E) \subset F$, isto é, T é sobrejetora. ■

Definição 3.2. Definimos o espaço vetorial c_{00} formado por seqüências eventualmente nulas da seguinte maneira: $c_{00} = \{(a_k)_{k=1}^\infty \in \mathbb{K}; \text{ existe } k_0 \in \mathbb{N} \text{ tal que } a_k = 0 \text{ para todo } k \geq k_0\}$.

A função $\|x\|_\infty = \sup_{n \in \mathbb{N}} \{|a_n|\}$ define uma norma em c_{00} , para todo $x = (a_k)_{k=1}^\infty \in c_{00}$.

Exemplo 3.3. *Seja $T: c_{00} \rightarrow c_{00}$ o operador linear dado por $T((a_n)_{n=1}^\infty) = \left(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \dots\right)$. Veremos a seguir que T é linear, bijetor e contínuo, mas que o operador inverso T^{-1} não é contínuo.*

Sejam $x = (a_n)_{n=1}^\infty, y = (b_n)_{n=1}^\infty \in c_{00}$ e $\alpha \in \mathbb{K}$. Observe que

$$\begin{aligned} T((a_n)_{n=1}^\infty + (b_n)_{n=1}^\infty) &= T((a_n + b_n)_{n=1}^\infty) \\ &= \left(a_1 + b_1, \frac{a_2 + b_2}{2}, \frac{a_3 + b_3}{3}, \dots\right) \\ &= \left(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \dots\right) + \left(b_1, \frac{b_2}{2}, \frac{b_3}{3}, \dots\right) \\ &= T((a_n)_{n=1}^\infty) + T((b_n)_{n=1}^\infty) \end{aligned}$$

e

$$\begin{aligned} T(\alpha(a_n)_{n=1}^\infty) &= T((\alpha a_n)_{n=1}^\infty) \\ &= \left(\alpha a_1, \alpha \frac{a_2}{2}, \alpha \frac{a_3}{3}, \dots\right) \\ &= \alpha \left(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \dots\right) \\ &= \alpha T((a_n)_{n=1}^\infty) \end{aligned}$$

mostram que T é linear.

Ademais,

$$\|T(x)\| = \left\| \left(a_1, \frac{a_2}{2}, \frac{a_3}{3}, \dots \right) \right\| = \sup_{n \in \mathbb{N}} \left\{ |a_1|, \left| \frac{a_2}{2} \right|, \left| \frac{a_3}{3} \right|, \dots \right\} \leq \sup_{n \in \mathbb{N}} \{ |a_1|, |a_2|, |a_3|, \dots \} = \|x\|,$$

isto é, T é contínua. Mais ainda, T é bijetora e sua inversa é dada por

$$T^{-1}((a_n)_{n=1}^{\infty}) = (a_1, 2a_2, 3a_3, \dots)$$

. Considere os vetores $e_n = (0, \dots, 0, \underbrace{1}_{n\text{-ésima}}, 0, \dots)$, $n \in \mathbb{N}$. Note que, para cada $n \in \mathbb{N}$, $\|e_n\|_{\infty} = 1$, mas $\|T^{-1}(e_n)\|_{\infty} = n$, ou seja, T^{-1} não é contínua.

Como consequência do Teorema da Aplicação Aberta, é possível garantir que sempre que os espaços E e F forem completos, a inversa de uma bijeção linear $T: E \rightarrow F$ sempre é contínua. É importante observar que c_{00} não é completo, (ver [1], Exemplo 1.1.7) e, por isso, não vale a continuidade de T^{-1} , algo equivalente a dizer que T é aberta.

4 Conclusões Finais

Por meio dos fatos mostrados, percebemos que nenhuma das hipóteses do Teorema da Aplicação Aberta pode ser desconsiderada. De fato, elas são fundamentais para a demonstração do resultado como obtido.

5 Agradecimentos

Na condição de bolsista do PET Matemática da Universidade Federal de Uberlândia, agradeço ao PET-SESu-MEC pelo apoio financeiro.

Referências

- [1] BOTELHO, G.; PELLEGRINO, D.; TEIXEIRA, E. **Fundamentos de Análise Funcional**. Rio de Janeiro: SBM, 2012.
- [2] MUJICA, J. **Notas de Análise Funcional**. Unicamp, 2003.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Completamento de Espaços Métricos Completion of Metric Spaces

Ana Lúdia da Silva Lício

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
analidiasl@hotmail.com

Elisa Regina dos Santos

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
elisars@ufu.com

Palavras-chave

Métrica.
Espaços Métricos.
Sequências convergentes.
Sequências de Cauchy.
Completamento.

Resumo

Este trabalho é baseado no estudo de conjuntos que possuem uma métrica associada, denominados espaços métricos, e tem por objetivo principal apresentar e demonstrar o Teorema do Completamento de Espaços Métricos. Veremos que todo espaço métrico pode ser ampliado a um espaço métrico completo, que será denominado um completamento do espaço original, e também veremos que tal completamento é único a menos de isometria.

Keywords

Metric.
Metric Spaces.
Convergent sequences.
Cauchy sequences.
Completion.

Abstract

This work is based on the study of sets with an associated metric, called metric spaces, and its main purpose is to present and demonstrate the Metric Spaces Completion Theorem. We will see that every metric space can be expanded to a complete metric space, which will be called a completion of the original space, and we will also see that such completion is unique up to isometry.

1 Introdução

Um espaço métrico é um conjunto não vazio munido de uma noção de distância, denominada métrica. Segundo [1], o estudo dos espaços métricos foi iniciado pelo matemático francês Maurice Fréchet em 1905. Um espaço métrico é dito completo se toda sequência de pontos que cujos termos se tornam arbitrariamente próximos (conhecidas como sequências de Cauchy) for convergente para um ponto do espaço métrico. Por exemplo, o conjunto dos números racionais com a noção de distância usual é não um espaço métrico completo, enquanto o conjunto dos números reais com a noção de distância usual é completo. Mais do que isso, é possível provar que toda sequência de Cauchy de números racionais converge para um número real. Assim, o conjunto dos números reais formam um completamento do conjunto dos números racionais. Tal fato foi provado pela matemático alemão Felix Hausdorff em 1914.

Este trabalho objetiva explicar sobre os conceitos de espaços métricos completos e não completos. Inicialmente, introduziremos alguns importantes conceitos explorados no curso de espaços métricos, tais como métrica, sequências de Cauchy e espaços métricos completos. Tendo conhecimento de tais definições, investigaremos o completamento de um espaço métrico não completo e a unicidade deste completamento.

2 Preliminares

Apresentaremos a seguir algumas definições preliminares.

Definição 2.1. *Seja M um conjunto qualquer não vazio. Uma **métrica** em M é uma função $d : M \times M \rightarrow \mathbb{R}$ que satisfaz as seguintes condições:*

D1) $d(x, y) \geq 0$, para todos $x, y \in M$, e $d(x, y) = 0$ se, e somente se, $x = y$;

D2) $d(x, y) = d(y, x)$, para todos $x, y \in M$;

D3) $d(x, z) \leq d(x, y) + d(y, z)$, para todos $x, y, z \in M$.

O par (M, d) é dito **espaço métrico**.

Iremos nos referir a um espaço métrico (M, d) simplesmente por M quando não houver risco de confusão.

Definição 2.2. *Seja M um espaço métrico. Uma **sequência** em M é uma função $f : \mathbb{N} \rightarrow M$. Denotaremos $f(n)$ por $x_n \in M$. Usaremos as notações $(x_1, x_2, \dots, x_n, \dots)$, $(x_n)_{n=1}^{\infty}$ ou (x_n) para representar uma sequência. Usaremos também a notação $(x_n) \subset M$ para indicar que os termos da sequência (x_n) são pontos de M .*

Definição 2.3. Seja M um espaço métrico. Uma sequência (x_n) em M é dita **convergente** para $x \in M$ se, para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$d(x_n, x) < \varepsilon, \forall n \geq n_0.$$

Neste caso, x é dito **limite** de (x_n) e escrevemos $x_n \rightarrow x$.

Definição 2.4. Seja M um espaço métrico. Uma sequência (x_n) em M é dita uma **sequência de Cauchy** se, para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$d(x_n, x_m) < \varepsilon, \forall n, m \geq n_0.$$

Definição 2.5. Um espaço métrico M é dito **espaço métrico completo** ou apenas **completo** se toda sequência de Cauchy em M é convergente para um limite que também está em M .

Exemplo 2.6. O espaço métrico \mathbb{Q} não é completo.

Exemplo 2.7. Vemos no curso de Análise que toda sequência de Cauchy em \mathbb{R} é convergente. Logo, \mathbb{R} é completo.

Para mais informações sobre espaços métricos, sugerimos a referência [2].

3 Teorema do Completamento

Mostraremos a seguir que todo espaço métrico pode ser estendido a um espaço métrico completo, mais precisamente, podemos adicionar pontos a M de maneira que ele se torne completo.

Definição 3.1. Um **completamento** do espaço métrico (M, d) é um par $((\widehat{M}, \widehat{d}), \varphi)$ tal que:

- i) $(\widehat{M}, \widehat{d})$ é um espaço métrico completo;
- ii) $\varphi : M \rightarrow \widehat{M}$ é uma imersão isométrica;
- iii) $\varphi(M)$ é denso em \widehat{M} .

Exemplo 3.2. O completamento de \mathbb{Q} é \mathbb{R} . De fato, temos que \mathbb{R} com a métrica usual d é um espaço métrico completo, a função inclusão $J : \mathbb{Q} \hookrightarrow \mathbb{R}$ é uma imersão isométrica, uma vez que $d(J(x), J(y)) = d(x, y)$, e $J(\mathbb{Q}) = \mathbb{Q}$ é denso em \mathbb{R} .

Para demonstrarmos o Teorema do Completamento precisaremos do seguinte resultado.

Lema 3.3. Seja M um espaço métrico. Suponha que M contenha um subconjunto A tal que:

- (a) $A \neq \emptyset$;

(b) A é denso em M ;

(c) toda sequência de Cauchy em A , converge em M .

Então M é completo.

Demonstração. Veja [3, Lema 8.21]. ■

Teorema 3.4 (Teorema do Completamento). *Todo espaço métrico possui um completamento.*

Demonstração. Considere um espaço métrico (M, d) . Vamos demonstrar o teorema por partes.

a) Construção do espaço métrico $(\widehat{M}, \widehat{d})$:

Defina

$$S = \{(x_n) \subseteq M : (x_n) \text{ é de Cauchy}\}.$$

Considere a seguinte relação em S :

$$(x_n) \sim (y_n) \iff \lim_{n \rightarrow \infty} d(x_n, y_n) = 0.$$

É fácil notar que \sim é uma relação de equivalência. Defina então $\widehat{M} = S/\sim = \{\overline{(x_n)} : (x_n) \in S\}$, onde $\overline{(x_n)} = \{(w_n) \in S : (w_n) \sim (x_n)\}$, e $\widehat{d} : \widehat{M} \times \widehat{M} \rightarrow \mathbb{R}$ por

$$\widehat{d}(\overline{(x_n)}, \overline{(y_n)}) = \lim_{n \rightarrow \infty} d(x_n, y_n).$$

É possível mostrar que:

(1) a sequência $(d(x_n, y_n))$ é convergente;

(2) a definição de \widehat{d} independe dos representantes, isto é, se $(w_n) \sim (x_n)$ e $(z_n) \sim (y_n)$, então $\lim_{n \rightarrow +\infty} d(x_n, y_n) = \lim_{n \rightarrow \infty} d(w_n, z_n)$;

(3) \widehat{d} é uma métrica.

Assim, $(\widehat{M}, \widehat{d})$ é um espaço métrico.

b) Construção da imersão isométrica φ :

Defina $\varphi : M \rightarrow \widehat{M}$ por $\varphi(a) = \overline{(a, a, \dots)}$. Observe que φ está bem definida, pois $(a, a, \dots) \in S$ para todo $a \in M$. Além disso, para $a, b \in M$ temos

$$\begin{aligned} \widehat{d}(\varphi(a), \varphi(b)) &= \widehat{d}(\overline{(a, a, \dots)}, \overline{(b, b, \dots)}) \\ &= \lim_{n \rightarrow \infty} d(a, b) = d(a, b). \end{aligned}$$

Logo, φ é uma imersão isométrica.

c) $\varphi(M)$ é denso em \widehat{M} :

Sejam $\overline{(x_n)} \in \widehat{M}$ e $\varepsilon > 0$. Como (x_n) é de Cauchy, existe $n_1 \in \mathbb{N}$ tal que

$$m, n \geq n_1 \Rightarrow d(x_m, x_n) < \frac{\varepsilon}{2}.$$

Assim,

$$\widehat{d}(\overline{(x_n)}, \overline{(x_{n_1}, x_{n_1}, \dots)}) = \lim_{x \rightarrow \infty} d(x_n, x_{n_1}) \leq \frac{\varepsilon}{2} < \varepsilon.$$

Daí,

$$\widehat{d}(\overline{(x_n)}, \varphi(x_{n_1})) = \widehat{d}(\overline{(x_n)}, \overline{(x_{n_1}, x_{n_1}, \dots)}) < \varepsilon,$$

ou seja, $B(\overline{(x_n)}; \varepsilon) \cap \varphi(M) \neq \emptyset$. Como $\varepsilon > 0$ é arbitrário, segue que $\overline{(x_n)} \in \overline{\varphi(M)}$. Portanto, $\overline{\varphi(M)} = \widehat{M}$.

d) \widehat{M} é completo:

Mostremos que toda sequência de Cauchy em $\varphi(M)$ converge em \widehat{M} . Seja (β_n) uma sequência de Cauchy em $\varphi(M)$. Então, para cada $n \in \mathbb{N}$, existe $a_n \in M$ tal que $\beta_n = \varphi(a_n) = \overline{(a_n, a_n, \dots)}$. Como

$$\widehat{d}(\beta_m, \beta_n) = \widehat{d}(\overline{(a_m, a_m, \dots)}, \overline{(a_n, a_n, \dots)}) = d(a_m, a_n)$$

e (β_n) é de Cauchy, temos (a_n) de Cauchy em M . Logo, $(a_n) \in S$ e $\alpha = \overline{(a_n)} \in \widehat{M}$.

Mostremos que $\beta_n \rightarrow \alpha$ em \widehat{M} . Dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$m, n \geq n_0 \implies d(a_m, a_n) < \frac{\varepsilon}{2}.$$

Assim, para $n \geq n_0$, temos

$$\widehat{d}(\beta_n, \alpha) = \widehat{d}(\overline{(a_n, a_n, \dots)}, \overline{(a_1, a_2, \dots)}) = \lim_{k \rightarrow \infty} d(a_n, a_k) \leq \frac{\varepsilon}{2} < \varepsilon.$$

Logo, $\beta_n \rightarrow \alpha \in \widehat{M}$. Portanto, toda sequência de Cauchy em $\varphi(M)$ converge em \widehat{M} .

Como $\varphi(M) \neq \emptyset$ e $\varphi(M)$ é denso em \widehat{M} , segue do Lema 3.3 que \widehat{M} é completo, o que conclui a demonstração do teorema. ■

Teorema 3.5 (Unicidade do Completamento). *Sejam $((\widehat{M}, \widehat{d}), \varphi)$ e $(\widetilde{M}, \widetilde{d}), \psi)$ dois completamentos do espaço métrico (M, d) . Então, existe uma única isometria $f : \widehat{M} \rightarrow \widetilde{M}$ tal que $f \circ \varphi = \psi$.*

Demonstração. Veja [2, Capítulo 7, Proposição 14]. ■

Exemplo 3.6. *Sejam $P([0, 1]; \mathbb{R}) = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ é um polinômio}\}$ e $C([0, 1]; \mathbb{R}) = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ é contínua}\}$. É possível mostrar que $P([0, 1]; \mathbb{R})$ não é completo, mas é denso em $C([0, 1]; \mathbb{R})$ que é completo. Logo, $C([0, 1]; \mathbb{R})$ é o completamento de $P([0, 1]; \mathbb{R})$.*

4 Considerações finais

O foco principal deste trabalho foi demonstrar o Teorema do Completamento de Espaços Métricos e garantir que sempre é possível ampliar qualquer espaço métrico não completo para que se torne completo. Além de obter a existência de tal completamento, também vimos que ele é único a menos de uma isometria.

Agradecimentos

Enquanto bolsista do PICME, agradeço ao CNPq pelo apoio financeiro.

Referências

- [1] CARLSON, Stephan C.. **Metric space**. Encyclopedia Britannica. Disponível em: <<https://www.britannica.com/science/metric-space>>. Acesso em: 17 abr. 2021.
- [2] LIMA, Elon L.. **Espaços Métricos**. Coleção Projeto Euclides. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 2013.
- [3] FÁVARO, Vinícius V.. **Notas de Aula de Topologia dos Espaços Métricos**. Notas de aula, UFU-FAMAT, 2019.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O toro e a esfera são homeomorfos? Are the torus and the sphere homeomorphic?

Anderson Martins Rocha do Prado

UFU, FAMAT, Uberlândia, MG, Brasil

andersonrochadoprado@gmail.com

Ligia Lais Fêmina

UFU, FAMAT, Uberlândia, MG, Brasil

ligia@ufu.br

Palavras-chave

Característica de Euler.
Invariante topológico.
Espaços não homeomorfos.
Esfera.
Toro.

Resumo

A característica de Euler é um importante invariante topológico, que associa aos espaços um número inteiro, através dela é possível obter uma classificação das superfícies compactas, conexas e sem bordo. O principal objetivo desse trabalho é apresentar a característica de Euler e como aplicação desse invariante, descobrir se existe um homeomorfismo entre o toro e a esfera.

Keywords

Euler's characteristic.
Topological Invariant.
Non-homeomorphic spaces.
Torus .
Sphere .

Abstract

Euler's characteristic is an important topological invariant, which associates spaces with an integer, through which it is possible to obtain a classification of compact, connected and borderless surfaces. The main objective of the work is to present an Euler characteristic and as an application of this invariant, find out if there is a homeomorphism between the torus and a sphere.

1 Introdução

Há vários invariantes topológicos, como grupo fundamental, grupos de homologia e a característica de Euler. O propósito é o mesmo: permitir identificar espaços que não são homeomorfos. O interessante da característica de Euler é que classifica importantes espaços topológicos através somente de um número inteiro.

O objetivo desse trabalho é responder a seguinte pergunta: existe um homeomorfismo entre o toro e a esfera?

2 Preliminares

Nessa seção, serão abordados alguns conceitos fundamentais para o entendimento do trabalho. Para maiores detalhes consultar [1].

Definição 2.1. *Uma função $f : X \rightarrow Y$ é um homeomorfismo se ela é injetora, sobrejetora, contínua e possui inversa contínua. Nesse caso, dizemos que X e Y são homeomorfos.*

Definição 2.2. *Um espaço topológico X é de Hausdorff se, dados dois pontos distintos a e b de X , existem abertos disjuntos A contendo a e B contendo b .*

É relevante termos em mente que todo espaço métrico é de Hausdorff.

Definição 2.3. *Dizemos que um espaço topológico X é conexo se os únicos subconjuntos simultaneamente abertos e fechados de X são \emptyset e o próprio X . Equivalentemente, podemos dizer que se X é conexo, A e B são abertos disjuntos tais que $X = A \cup B$, então $A = \emptyset$ ou $B = \emptyset$.*

Definição 2.4. *Um espaço topológico X é compacto se toda família de abertos, cuja reunião é X , admite um número finito de abertos dessa família, cuja reunião ainda é X . Em outras palavras, X é compacto se toda cobertura aberta de X admite uma subcobertura finita.*

3 Característica de Euler

O número inteiro $\chi(P)$ associado a um poliedro recebe o nome de característica de Euler, em homenagem ao matemático Leonhard Euler (1707 - 1783). Euler escreveu uma carta para o matemático Christian Goldbach em 1750, onde afirmava que $\chi(P) = 2$ para qualquer poliedro. Alguns historiadores asseguram que um manuscrito de Descartes de 1639 contém resultados a partir dos quais poderia se chegar na relação de Euler. A relação de Euler é válida para poliedros convexos, para poliedros não convexos pode se ter $\chi(P) \neq 2$.

Definição 3.1. *Seja P um poliedro e denotemos por v , o número de vértice, f o número de faces e a o número de arestas de P . O número $\chi(P) = v - a + f$ é chamado **característica de Euler de P** .*

Iremos calcular a característica de Euler das superfícies: toro e esfera. Como sabemos que a fórmula depende do número de vértices, arestas e faces, a maneira de calcularmos irá depender de uma triangulação da superfície e aplicarmos a fórmula. Com esse intuito, definimos a seguir superfícies e triangularização.

Definição 3.2. Uma superfície é um espaço de Hausdorff, compacto e conexo, tal que para cada ponto existe um aberto que o contém e que é homeomorfo a uma bola aberta do \mathbb{R}^2 .

Definição 3.3. Uma superfície é triangulável quando podemos dividi-la em um número finito de vértices, arestas e faces, com as seguintes condições:

- 1) Qualquer aresta é aresta de dois triângulos;
- 2) Dados dois triângulos existe uma sequência de triângulos começando em um e terminando no outro, de modo que dois triângulos consecutivos possuem uma aresta em comum.

Teorema 3.4. Se S_1 e S_2 são duas superfícies fechadas, então S_1 é homeomorfa a S_2 se, e somente se $\chi(S_1) = \chi(S_2)$ e ambas são orientáveis ou ambas são não orientáveis.

Demonstração: [2], teorema 5.16.

Exemplo 3.5. Considere a seguinte triangularização do toro T^2 :

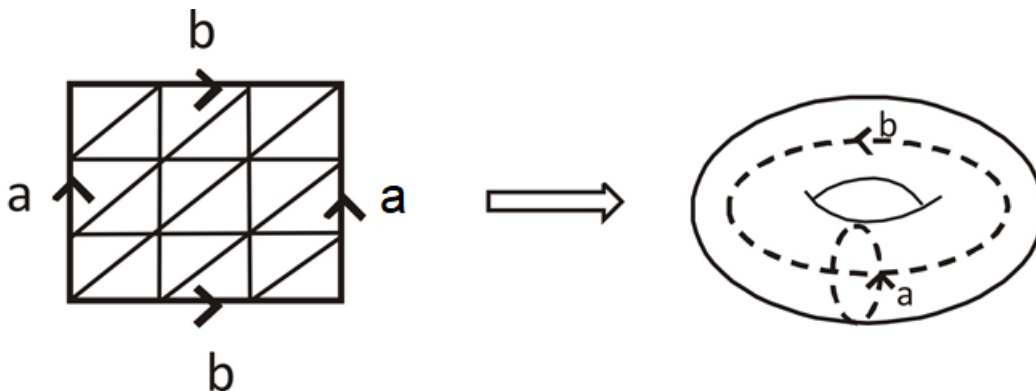


Figura 1: Triangularização do toro

Nota-se que temos $v = 9$, $a = 27$ e $f = 18$. Logo,

$$\chi(T^2) = v - a + f = 9 - 27 + 18 = 0$$

Exemplo 3.6. Considere a seguinte triangularização da esfera S^2 :

Nota-se que temos $v = 4$, $a = 6$ e $f = 4$. Logo,

$$\chi(S^2) = v - a + f = 4 - 6 + 4 = 2$$

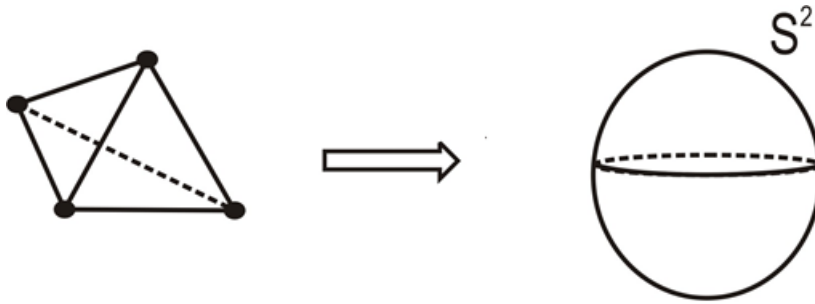


Figura 2: Triangularização da esfera

4 Conclusão

O toro e a esfera são homeomorfos?

A resposta é não!

De fato, pelo exemplo 3.5 obtemos $\chi(T^2) = 0$ e pelo exemplo 3.6, temos $\chi(S^2) = 2$. Como $\chi(T^2) \neq \chi(S^2)$, segue pelo teorema 3.4 que a esfera e o toro não são homeomorfos.

5 Considerações finais

Existe uma generalização para o cálculo da característica de Euler:

Seja S um complexo simplicial, a característica de Euler-Poincaré de S é dada por

$$\chi(S) = \alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 + \dots + (-1)^n \alpha_n,$$

onde α_i é o número de simplexes de dimensão i .

Entender esse resultado é o próximo objetivo dessa pesquisa.

Agradecimentos

À PIBIC/CNPq por ter financiado o presente estudo, tornando possível sua realização, além de exercer um trabalho fundamental na vida de diversos pesquisadores e por confiar na ciência brasileira, que tanto carece de instituições de apoio e fomento.

Referências

- [1] MUNKRES, J.R. **Topology: A First Course**. Prentice Hall, Inc., 1975.
- [2] LIBARDI, A. K. M.; VIEIRA, J. P.; MELO, T. **Invariantes Topológicos**. São Paulo: Cultura Acadêmica, 2012.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Diagonalização de Matrizes Simétricas na Identificação de Cônicas Diagonalization of Symmetric Matrices on Identification of Conics

Bruno Henrique Viana de Moraes

UFU, Instituto de Ciências Agrárias, Uberlândia, Minas Gerais, Brasil

brunohvmorais@ufu.br

Catiana Casonatto

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

ccasonatto@ufu.br

Palavras-chave

Cônica
Matriz
Diagonalização

Resumo

Descrevemos um método para identificar uma cônica cuja equação não está na sua forma padrão. Para isso, usamos o fato de que toda matriz simétrica A é diagonalizável por uma matriz ortogonal, isto é, existe uma matriz diagonal D e uma matriz ortogonal P (cujas colunas são autovetores ortonormais de A), tais que $A = PDP^t$. A partir dessas matrizes, é possível determinar uma mudança de coordenadas para reduzir a equação à forma padrão.

Keywords

Conic
Matrix
Diagonalization

Abstract

We describe a method to identify a conic section by its equation when its not in standard form. For this, we use the fact that every symmetric matrix A is diagonalizable with an orthogonal matrix, that is, there is a diagonal matrix D and an orthogonal matrix P (whose columns are orthonormal eigenvectors of A), such that $A = PDP^t$. From these matrices, it is possible to determine a change in coordinates in order to reduce the equation to the standard form.

1 Introdução

Reconhecer uma cônica em \mathbb{R}^2 a partir de sua equação é bastante fácil, desde que esteja em sua forma padrão. Caso contrário, tal tarefa se torna mais complexa, sendo necessário realizar mudanças de coordenadas com o intuito de reduzi-la à sua forma padrão. Neste trabalho, aplicamos conceitos e resultados sobre diagonalização de matrizes simétricas para descrevermos um método eficiente de se efetuar tais simplificações. Assumiremos conceitos básicos da teoria de Geometria Analítica e Álgebra Linear que podem ser estudados com detalhes em [F. Coelho, R. Santos].

2 Diagonalização de Matrizes

Uma matriz $n \times n$ A é dita **diagonalizável** se existirem matrizes P e D tais que $A = PDP^{-1}$ ou, equivalentemente, $D = P^{-1}AP$, onde D é diagonal.

Se A é diagonalizável e P, D são as matrizes

$$P = \begin{bmatrix} V_1 & V_2 & \cdots & V_n \end{bmatrix} \text{ e } D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

onde V_j é a j -ésima coluna de P , tais que $A = PDP^{-1}$, então multiplicando à direita por P em ambos os membros da equação anterior, obtemos $AP = PD$. Assim,

$$AP = A \begin{bmatrix} V_1 & V_2 & \cdots & V_n \end{bmatrix} = \begin{bmatrix} AV_1 & AV_2 & \cdots & AV_n \end{bmatrix}$$

e

$$PD = \begin{bmatrix} V_1 & V_2 & \cdots & V_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 V_1 & \lambda_2 V_2 & \cdots & \lambda_n V_n \end{bmatrix}.$$

Logo,

$$AV_j = \lambda_j V_j,$$

para $j = 1, 2, \dots, n$, isto é, V_j são **autovetores** de A associados aos **autovalores** λ_j , para $j = 1, 2, \dots, n$. E mais, como a matriz P é invertível, os autovetores V_1, \dots, V_n são linearmente independentes. Portanto, se A é diagonalizável e $A = PDP^{-1}$, então os autovalores de A formam a diagonal de D e n autovetores linearmente independentes associados aos autovalores formam as colunas de P .

O próximo resultado nos dá informações sobre os autovetores de uma matriz A **simétrica** (aquela

que satisfaz a igualdade $A^t = A$).

Proposição 2.1. *Se uma matriz A é simétrica, então os autovetores associados a autovalores diferentes são ortogonais.*

Demonstração. Sejam V_1 e V_2 autovetores de A associados aos autovalores λ_1 e λ_2 , respectivamente, com $\lambda_1 \neq \lambda_2$. Então $AV_1 = \lambda_1 V_1$ e $AV_2 = \lambda_2 V_2$. Se escrevermos os autovetores como matrizes colunas, então o produto escalar corresponde ao produto matricial da transposta da primeira pela segunda e obtemos

$$AV_1 \cdot V_2 = (AV_1)^t \cdot V_2 = V_1^t A^t V_2 = V_1 \cdot A^t V_2.$$

Como A é simétrica e V_1 e V_2 são autovetores de A , temos que

$$\lambda_1 V_1 \cdot V_2 = AV_1 \cdot V_2 = V_1 \cdot A^t V_2 = \lambda_2 V_1 \cdot V_2 \Leftrightarrow (\lambda_1 - \lambda_2) V_1 \cdot V_2 = 0, \text{ com } \lambda_1 \neq \lambda_2.$$

Logo, $V_1 \cdot V_2 = 0$ e, portanto, V_1 e V_2 são ortogonais. ■

Veremos a seguir que uma matriz A simétrica é sempre diagonalizável com uma matriz **ortogonal**, P (aquela em que $P^t = P^{-1}$; equivalentemente, as colunas de P são ortonormais).

Teorema 2.2. *Se A é uma matriz simétrica, então ela é diagonalizável com uma matriz ortogonal, ou seja, existe uma matriz ortogonal P e uma matriz diagonal D tal que*

$$A = PDP^t.$$

Demonstração. Veja [R. Santos], Apêndice V. ■

Uma vez que, pela Proposição 2.1, os autovetores de uma matriz simétrica A associados a autovalores distintos já são ortogonais, para encontrarmos a matriz P ortonormal que realiza a diagonalização da matriz A , cuja existência é garantida pelo Teorema 2.2, basta tomarmos, para cada autovalor, autovetores ortonormais associados a eles. Para isso, podemos aplicarmos o processo de ortogonalização de Gram-Schmidt (veja [F. Coelho]).

3 Identificação das cônicas

Chamamos de **cônica** o lugar geométrico dos pontos $P = (x, y)$ no plano que satisfazem a equação

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

onde $a, b, c, d, e, f \in \mathbb{R}$ e a, b e c não simultaneamente nulos. Quando b é zero, é simples reconhecermos qual é a cônica, pois basta realizarmos uma translação de eixos e chegaremos na sua forma padrão.

Agora, quando b for diferente de zero, devemos realizar uma rotação de eixos para obtermos sua forma padrão. O próximo resultado nos fornece a mudança de coordenadas apropriada que realiza tal rotação.

Teorema 3.1. *Considere a equação*

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (1)$$

com $a, b, c, d, e, f \in \mathbb{R}$, sendo $b \neq 0$ e a, c não simultaneamente nulos. Então existe um sistema de coordenadas (x', y') , onde a equação (1) tem a forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0,$$

sendo λ_1, λ_2 autovalores de

$$A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}.$$

Mais ainda,

$$X' = QX,$$

onde $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \end{bmatrix}$ e Q é uma matriz ortogonal.

Demonstração. Podemos expressar a equação (1) da forma

$$X^tAX + KX + f = 0,$$

com $A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$, $K = \begin{bmatrix} d & e \end{bmatrix}$ e $X = \begin{bmatrix} x \\ y \end{bmatrix}$.

Como A é simétrica, pelo Teorema 2.2, A é diagonalizável com uma matriz ortogonal, isto é, $A = PDP^t$, equivalentemente $D = P^tAP$, onde P é ortogonal e D é diagonal. Mas ainda, podemos considerar

$$P = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \text{ e } D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

onde U_1 e U_2 são autovetores ortonormais de A associados aos autovalores λ_1 e λ_2 , respectivamente.

Considerando $X' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ e realizando a mudança de variáveis $X = PX'$ em $X^tAX + KX + f = 0$ obtemos

$$(PX')^tA(PX') + KPX' + f = 0 \Leftrightarrow X'^t(P^tAP)X' + KPX' + f = 0 \Leftrightarrow X'^tDX' + KPX' + f = 0,$$

a qual pode ser escrita da forma

$$\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' + f = 0, \text{ onde } \begin{bmatrix} d' & e' \end{bmatrix} = KP.$$

Por fim, basta tomarmos $Q = P^t$ e teremos $X' = QX$, onde Q é uma matriz ortogonal. ■

Exemplo 3.2. Identificar a cônica definida por

$$4x^2 - 20xy + 25y^2 - 15x - 6y = 0. \quad (2)$$

Podemos reescrever a equação acima da forma

$$X^t AX + KX = 0,$$

onde $A = \begin{bmatrix} 4 & -10 \\ -10 & 25 \end{bmatrix}$, $K = \begin{bmatrix} -15 & -6 \end{bmatrix}$ e $X = \begin{bmatrix} x & y \end{bmatrix}$. O polinômio característico de A é

$$p(\lambda) = \det(A - \lambda I_2) = \det \begin{bmatrix} 4 - \lambda & -10 \\ -10 & 25 - \lambda \end{bmatrix} = \lambda^2 - 29\lambda.$$

Portanto, os autovalores de A são $\lambda_1 = 29$ e $\lambda_2 = 0$. Para $\lambda_1 = 29$, temos o sistema $(A - 29I_2)X = 0$:

$$\begin{bmatrix} -25 & -10 \\ -10 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

que possui solução $\mathbb{W}_1 = \left\{ \left(\alpha, -\frac{5}{2}\alpha \right), \alpha \in \mathbb{R} \right\}$. Assim, um autovetor unitário associado ao autovalor $\lambda_1 = 29$ é $U_1 = \left(\frac{2}{\sqrt{29}}, -\frac{5}{\sqrt{29}} \right)$. Analogamente, para $\lambda_2 = 0$, temos o sistema $(A - 0I_2)X = 0$, que possui solução $\mathbb{W}_2 = \left\{ \left(\alpha, \frac{2}{5}\alpha \right), \alpha \in \mathbb{R} \right\}$. Desta maneira, $U_2 = \left(\frac{5}{\sqrt{29}}, \frac{2}{\sqrt{29}} \right)$ é um autovetor unitário associado ao autovalor $\lambda_2 = 0$. Como U_1 e U_2 são autovetores unitários associados a autovalores diferentes, pela Proposição 2.2, segue que são ortonormais. Assim, a matriz A é dada por $A = PDP^t$, onde

$$D = \begin{bmatrix} 29 & 0 \\ 0 & 0 \end{bmatrix} \text{ e } P = \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{29}} & \frac{5}{\sqrt{29}} \\ \frac{-5}{\sqrt{29}} & \frac{2}{\sqrt{29}} \end{bmatrix}.$$

Pelo Teorema 2.2, podemos escrever a equação (2) da forma $\lambda_1 x'^2 + \lambda_2 y'^2 + d'x' + e'y' = 0$, onde d', e' são tais que $\begin{bmatrix} d' & e' \end{bmatrix} = KP$, isto é,

$$29x'^2 - 3\sqrt{29}y' = 0,$$

ou ainda

$$x'^2 = \frac{3\sqrt{29}}{29}y',$$

a qual define uma **parábola**.

Por fim, para esboçar o gráfico desta parábola, devemos primeiramente traçar os eixos x' e y' . O eixo x' é obtido fazendo $y' = 0$ em $X = PX'$, ou seja, o eixo x' passa pela origem e tem o mesmo

sentido e direção do vetor U_1 . Já o eixo y' é obtido fazendo $x' = 0$ em $X = PX'$, o que nos dá que o eixo y' passa pela origem e tem o mesmo sentido e direção do vetor U_2 . A parábola está representada na Figura 1.

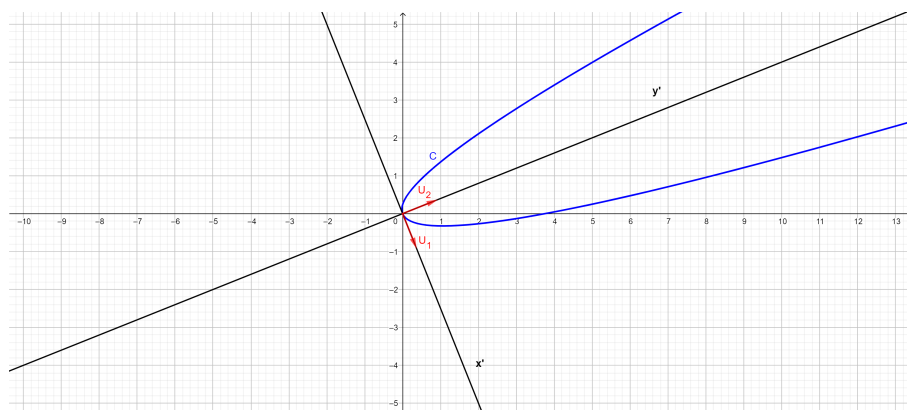


Figura 1: Parábola do Exemplo 3.2

4 Considerações finais

O método apresentado neste trabalho se mostrou bastante prático na identificação de uma cônica a partir de sua equação, bem como na obtenção de um sistema de coordenadas que facilita o esboço da mesma. Vale a pena ressaltar, que tal aplicação só foi possível, pois a matriz P , que realiza a diagonalização da matriz A definida no Teorema 2.2, é ortogonal.

Agradecimentos

Agradecimentos ao CNPq pelo apoio financeiro via Programa de Iniciação Científica PICME-UFU.

Referências

- [R. Santos] Santos, Reginaldo J., *Um Curso de Geometria Analítica e Álgebra Linear*. Belo Horizonte: Imprensa Universitária da UFMG, 2020. Disponível em: <<https://regijs.github.io/livros.html>>. Acesso em: out. 2020.
- [F. Coelho] Coelho, F. U e Lourenço, M. L., *Um Curso de Álgebra Linear*. São Paulo: Editora da Universidade de São Paulo, 2005.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O problema da Trisseção do Ângulo The problem of Angle Trisection

Cinira Aparecida de Oliveira

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil

ciniraapoliveira@ufu.br

Adriana Rodrigues da Silva

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil

adrianafamat@ufu.br

Palavras-chave

Trisseção do ângulo.
Problema clássico.
Geometria Plana.
Extensões algébricas.
Origami.

Resumo

Os problemas matemáticos clássicos da Grécia Antiga formavam um conjunto de construções geométricas em que se procuravam as soluções usando apenas régua e compasso. Um desses problemas é a Trisseção do Ângulo, que consiste em dividir um ângulo qualquer em três ângulos congruentes. Neste trabalho, vamos mostrar por meio de uma solução algébrica, a impossibilidade desta construção e que utilizando a Geometria com Origami, trissectar um ângulo qualquer se torna possível.

Keywords

Trisection of the angle.
Classic problem.
Plan Geometry.
Algebraic extensions.
Origami.

Abstract

The classical mathematical problems of Ancient Greece formed a set of geometric constructions in which solutions were sought using only a ruler and compass. One of these problems is the Angle Trisection, which consists of dividing any angle into three congruent angles. In this work, we will show by means of an algebraic solution, the impossibility of this construction and that using Geometry with Origami, trisecting any angle becomes possible.

1 Introdução

Na Geometria Plana, alguns ângulos particulares, como por exemplo, os de 135° e 90° , podem ser trissectados recorrendo a uma régua não graduada e a um compasso, mas, no caso geral, é impossível dividir um ângulo em três partes iguais dessa forma, pois demonstra-se que a equação usada para resolver o problema é cúbica e tem a forma:

$$a^3 - 3a + 3b = 0, [1].$$

Logo, iremos recorrer a métodos para solucionar esse problema utilizando Extensões Algébricas e a Geometria com Origami.

2 Desenvolvimento

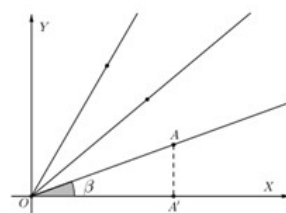
2.1 Extensões Algébricas:

Teorema 2.1. *O ângulo $\frac{\pi}{3}$ não pode ser trissectado utilizando régua e compasso.*

Demonstração:

Suponha, por absurdo, que seja possível trissectar um ângulo qualquer utilizando régua e compasso. Logo, podemos fazer a trissecção de um ângulo cuja medida é $\theta = \frac{\pi}{3}$, e trissectar tal ângulo é equivalente a obtermos $\beta = \frac{\pi}{9}$.

A figura nos mostra um plano coordenado XY , e uma semirreta OA a partir da origem, que forma um ângulo de medida β no primeiro quadrante sobre o eixo positivo OX . É possível realizarmos uma projeção no eixo X pelo ponto A , obtendo o ponto A' , dessa forma: $OA' = x = \cos(\frac{\pi}{9}) = \cos(\beta)$.



Agora, usando a identidade trigonométrica: $\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta)$ ao substituirmos θ por $\frac{\pi}{9}$ encontramos

$$\cos(\frac{\pi}{3}) = 4\cos^3(\frac{\pi}{9}) - 3\cos(\frac{\pi}{9}) = \frac{1}{2}, \text{ e daí segue, } 8\cos^3(\frac{\pi}{9}) - 6\cos(\frac{\pi}{9}) - 1 = 0, [1].$$

Logo, $\cos(\frac{\pi}{9})$ é raiz do polinômio $p(x) = 8x^3 - 6x - 1$ onde $p(x) \in \mathbb{Q}[x]$.

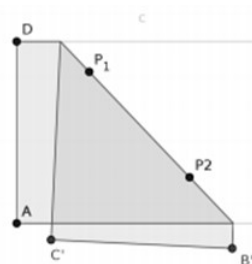
Tal polinômio é irredutível e caracteriza $\cos(\frac{\pi}{9}) = \cos(\beta)$ como um número algébrico sobre os racionais \mathbb{Q} . Portanto, podemos construir uma extensão algébrica do tipo $\mathbb{Q}(\cos(\beta)) : \mathbb{Q}$ e o grau dessa extensão será 3, isto é, $[\mathbb{Q}(\cos(\beta)) : \mathbb{Q}] = 3$.

Contradição, pois se $\frac{\pi}{9}$ é construtivo então $\cos(\frac{\pi}{9})$ também será construtivo, entretanto, o grau de extensão para números construtivos será sempre 1 ou 2, [2].

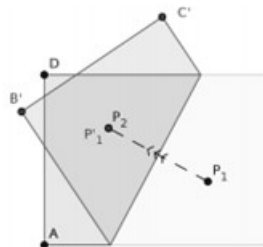
2.2 Origami:

Primeiramente, apresentaremos os Axiomas de Huzita, que são de suma importância para as construções com Origamis.

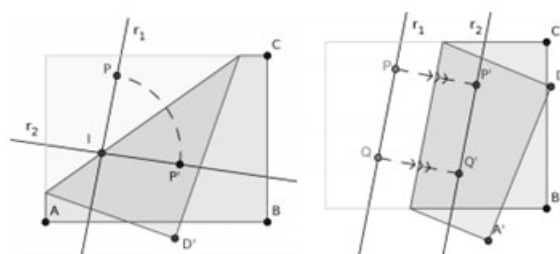
Axioma 2.1. *Dados dois pontos distintos P_1 e P_2 , existe apenas uma dobra que passa por eles.*



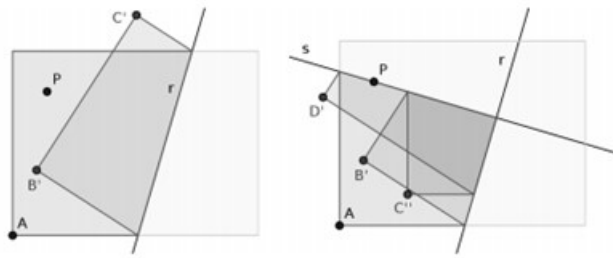
Axioma 2.2. *Dados dois pontos distintos P_1 e P_2 , existe apenas uma dobra que coloca P_1 sobre P_2 .*



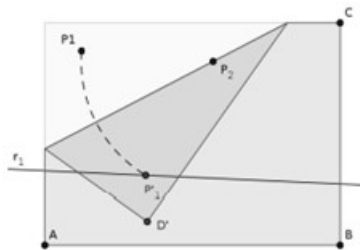
Axioma 2.3. *Dadas as retas r_1 e r_2 , existe uma única dobra que coloca r_1 sobre r_2 .*



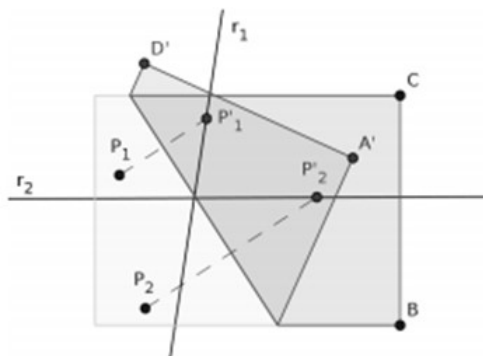
Axioma 2.4. *Dados um ponto P e uma reta r , existe uma única dobra que é perpendicular a r e que passa por P .*



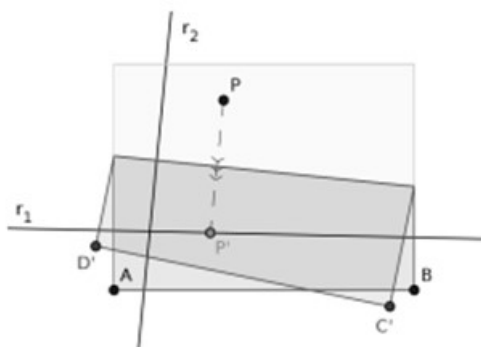
Axioma 2.5. Dados dois pontos P_1 e P_2 e uma reta r_1 , existe uma dobra que coloca P_1 sobre r_1 e que passa por P_2 .



Axioma 2.6. Dados dois pontos P_1 e P_2 e duas retas r_1 e r_2 , existe uma dobra que leva simultaneamente P_1 sobre r_1 e P_2 sobre r_2 .

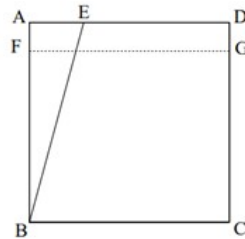


Axioma 2.7. Dados um ponto P e duas retas r_1 e r_2 , existe uma dobra que coloca P sobre r_1 e que é perpendicular a r_2 .

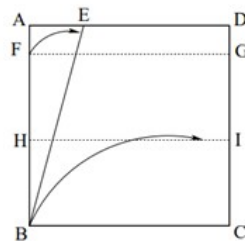


Trissecção de um ângulo qualquer com Origami:

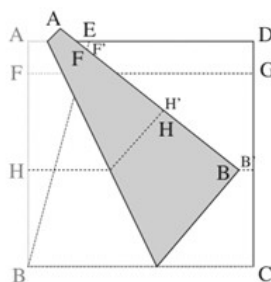
Seja um ângulo $\angle EBC$ menor do que 90° , vamos determinar uma paralela FG a AD .



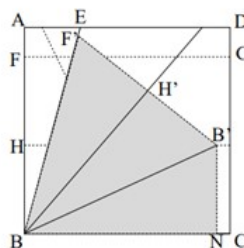
Vamos construir uma paralela HI , onde H e I são os respectivos pontos médios de FB e GC . dobre de modo a levar o ponto F ao segmento EB e o ponto B ao segmento HI . Esta última dobra é dada pelo Axioma 2.6, o qual não pode ser realizado com auxílio de régua e compasso.



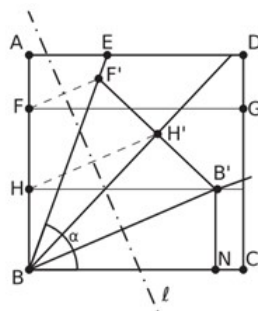
Para uma melhor visualização, vamos marcar os pontos H', F' e B' sobre o papel, e vamos traçar o segmento $F'B'$.



Abrindo novamente o papel, vamos traçar os segmentos $B'B$ e $H'B$. Após isso, traçamos pelo ponto B' uma paralela a HB , com extremidade N . Temos que os triângulos $\triangle BB'N$, $\triangle BB'H'$ e $\triangle BF'H'$ são congruentes, com os ângulos em B congruentes.



De fato, os triângulos $\triangle BB'N$ e $\triangle BB'H'$ são congruentes, pois possuem a hipotenusa BB' em comum e os catetos opostos aos ângulos no vértice B são congruentes, já que $\overline{NB'} = \overline{BH} = \overline{B'H'}$. Já os triângulos $\triangle BB'H'$ e $\triangle BF'H'$ são congruentes, pois possuem um cateto BH' em comum e os catetos opostos aos ângulos no vértice B são congruentes, pois $\overline{H'B'} = \overline{HB} = \overline{HF} = \overline{H'F'}$.



Portanto, o vértice dos triângulos que estão em B têm os mesmos ângulos, assim, $\angle EBC$ está dividido em três partes congruentes, [3].

3 Considerações finais

A demonstração da impossibilidade da trissecção de um ângulo qualquer nos mostra as limitações existentes nas ferramentas Euclidianas, [4].

Dessa forma, a utilização de soluções algébricas e a Geometria com Origami, se tornam alternativas para esse tipo de problema.

Referências

- [1] GUSMAI, R. **Um estudo sobre três problemas clássicos da geometria euclidiana.** Disponível em: <https://teses.usp.br/teses/disponiveis/55/55136/tde-29112016-141932/publico/RafaelMartinsGusmai_revisada.pdf>. Acesso em: 10 abr. 2021.
- [2] GONÇALVES, A. **Introdução à álgebra.** Projeto Euclides. 5ª Edição. Rio de Janeiro: IMPA, 2011.
- [3] CAVACAMI, E.; FURUYA, Y. **Explorando Geometria com Origami.** Apostila OBMEP, 2010.
- [4] REZENDE, E.; QUEIROZ, M. **Geometria Euclidiana Plana e construções geométricas.** 2ª Edição. Campinas: UNICAMP, 2008.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

24 A 26 DE MAIO DE 2021

Controlabilidade em Sistemas Dinâmicos Lineares

Controllability of Linear Dynamical Systems

Dhara Cristina de Freitas Lago Grande

UFU, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil
dharaacristina1999@gmail.com

Luciana Aparecida Alves

UFU, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil
luciana.alves@ufu.br

Resumo

Palavras-chave

Sistemas Dinâmicos.
Sistemas de Controle.
Sistema Autônomo.
Controlabilidade.

Desde a antiguidade já existia a necessidade do controle de sistemas físicos a fim de obter um resultado desejado. Os conceitos de controlabilidade e as noções de estabilidade desempenham um papel essencial nas soluções de muitos problemas importantes de controle. Neste trabalho, portanto, apresentaremos alguns dos conceitos presentes na teoria de sistemas de controle linear e que serão essenciais para o entendimento de uma de suas aplicações, a qual será desenvolvida por meio da solução de um exemplo ao final do texto aqui percorrido. Assim sendo, iniciaremos definindo o que é a controlabilidade de um sistema para, posteriormente, introduzirmos algumas propriedades sobre os conjuntos controláveis e outras definições necessárias para o estudo do tema, tais como as definições de conjunto aberto e conexidade por caminhos.

Abstract

Keywords

Dynamical Systems.
Control Systems.
Autonomous Systems.
Controllability.

Since ancient times, there has been a need to control physical systems in order to obtain a desired result. The concepts of controllability and the notions of stability play an essential role in the solutions of many important control problems. In this work we will present some of these concepts in the theory of linear control systems that will be essential for the understanding of one of its applications, which will be developed through the solution of an example at the end of the text discussed here. Therefore, we will start by defining what is the controllability of a system and, subsequently, we will introduce some properties on the controllable sets and other necessary definitions for the study of the theme, such as the definitions of open set and path connectedness.

1 Introdução

Esta seção será utilizada para introduzir algumas definições e proposições que serão de suma importância para o desenvolvimento do exemplo a ser apresentado. Vale observar que os conceitos a seguir e as devidas demonstrações podem ser encontradas nas referências [1], [2] e [3]. Isto posto, a primeira informação que precisamos saber é que trabalharemos com sistemas lineares autônomos da forma

$$x' = Ax + Bu, \quad (1)$$

em que $A \in \mathbb{R}^{n \times n}$ e $B \in \mathbb{R}^{n \times m}$ são matrizes constantes; x é um vetor de dimensão n (chamado de vetor estado); e u é uma variável de controle m -dimensional, com n e m arbitrários.

1.1 Controlabilidade

Para definirmos controlabilidade de um sistema, suponhamos $x' = f(x, u)$ equação de estado análoga a (1), em que f é continuamente diferenciável em x e em u .

Consideraremos, inicialmente, a controlabilidade à origem, ou seja, o nosso *estado final (ou estado alvo)* será $x = 0$, dessa forma, temos $f(0, 0) = 0$ e podemos garantir que o sistema atingirá o estado final desejado e permanecerá naquele ponto mesmo com as mudanças de controle. Além disso, como estamos trabalhando com sistemas autônomos, podemos definir um tempo inicial $t = 0$, de modo que o *estado inicial* seja da forma: $x(0) = x_0$.

Assim como mencionado anteriormente, as variáveis de controle são funções de t , integráveis, e as classificaremos de acordo com seus diferentes tipos. Dados \mathcal{U}_{bb} , \mathcal{U}_b e \mathcal{U}_u **conjuntos de controle** e u variável de controle, temos que:

- se estas funções são *ilimitadas*, então $u \in \mathcal{U}_u$;
- se *limitadas*, ou seja, se $|u_i(t)| \leq 1$, para $i = 1, 2, \dots, m$, então $u \in \mathcal{U}_b$;
- se $|u_i(t)| = 1$, para $i = 1, 2, \dots, m$, então $u \in \mathcal{U}_{bb}$.

Logo, $\mathcal{U}_{bb} \subset \mathcal{U}_b \subset \mathcal{U}_u$.

Definição 1.1. *Definimos o conjunto controlável no tempo t_1 como sendo o conjunto dos estados iniciais x_0 que podem ser levados até a origem em um tempo t_1 , usando um controle que seja admissível, ou seja, que pertença ao conjunto de controle \mathcal{U} escolhido.*

A notação a ser utilizada para nos referirmos ao conjunto definido acima será $\mathcal{C}(t_1)$ ou, caso seja necessário um detalhamento maior, $\mathcal{C}(t_1, u, 0)$, com t_1 , u e 0 , o tempo, o controle e o estado alvo, respectivamente.

Definição 1.2. O conjunto controlável \mathcal{C} é o conjunto de pontos que podem ser levados à origem em tempo finito qualquer. Escrevemos:

$$\mathcal{C} = \bigcup_{t_1 \geq 0} \mathcal{C}(t_1).$$

Se todos os estados iniciais forem controláveis à origem, ou seja, se $\mathcal{C} = \mathbb{R}^n$, dizemos que o sistema é **completamente controlável**.

1.2 Conjunto Controlável

Definida a controlabilidade de um sistema, trabalharemos, agora, com algumas propriedades dos conjuntos controláveis que serão utilizadas na resolução do exemplo a ser apresentado.

Proposição 1.3. Se x_0 pertence ao conjunto controlável \mathcal{C} e temos y um ponto na trajetória, ou seja, no caminho percorrido, de x_0 até 0, então $y \in \mathcal{C}$.

Proposição 1.4. \mathcal{C} é aberto, se e somente se, $0 \in \text{int } \mathcal{C}$ (em que $\text{int } \mathcal{C}$ denota o interior do conjunto \mathcal{C}).

Demonstração. Inicialmente, provemos que se \mathcal{C} é aberto, então $0 \in \text{int } \mathcal{C}$. Sabemos que $0 \in \mathcal{C}$ e, além disso, pela hipótese, temos que $\mathcal{C} = \text{int } \mathcal{C}$, logo, $0 \in \text{int } \mathcal{C}$.

Basta, agora, provarmos que se $0 \in \text{int } \mathcal{C}$, então \mathcal{C} é aberto. Por definição, se $0 \in \text{int } \mathcal{C}$, então existe $r > 0$, tal que $\mathcal{B}(0, r) \subset \mathcal{C}$. Considere $x_0 \in \mathcal{C}$ e seja $u(t)$, um controle que leva x_0 até 0 em um determinado tempo t_1 e seja $y(0) = y_0$, uma solução da equação (1) tal que $y_0 \in \mathcal{B}(x_0, r_0)$, como já vimos que f é contínua, temos que $y(t_1) = y_1$, com $y_1 \in \mathcal{B}(0, r)$ para um r_0 suficientemente pequeno.

Como sabemos que $\mathcal{B}(0, r) \subset \mathcal{C}$, conseguimos encontrar um outro controle $\hat{u}(t)$ que leve y_1 até 0 em um tempo t_2 desejado. Tomemos, assim, o controle $v(t)$ definido da seguinte forma:

$$v(t) = \begin{cases} u(t), & \text{para } 0 \leq t \leq t_1 \\ \hat{u}(t - t_1), & \text{para } t_1 \leq t \leq t_2. \end{cases}$$

Logo, se usarmos o controle $v(t)$, conseguimos levar y_0 até 0 no tempo $t_1 + t_2$, de forma que $y_0 \in \mathcal{C}(t_1 + t_2)$ e $\mathcal{B}(x_0, r_0) \subset \mathcal{C}$, para todo $x_0 \in \mathcal{C}$. Concluimos, assim, que \mathcal{C} é aberto, se $0 \in \mathcal{C}$, como queríamos demonstrar. ■

Definição 1.5. Seja $X \subset \mathbb{R}^n$ um espaço métrico e $I \subset \mathbb{R}$ um intervalo. Um **caminho** em X é uma função contínua $f : I \rightarrow X$, definida no intervalo I . Em outras palavras, f é o caminho que liga os pontos iniciais e finais de I . E dizemos que X é **conexo por caminhos** quando, para quaisquer dois pontos $x_1, x_2 \in X$, existe um caminho f (definido anteriormente) ligando x_1 e x_2 .

O conjunto controlável \mathcal{C} também possui as propriedades dadas nas proposições a seguir. As provas desses resultados podem ser encontradas em [2].

Proposição 1.6. \mathcal{C} é conexo por caminhos, ou seja, para quaisquer dois pontos $x_0, x_1 \in \mathcal{C}$, existe uma função contínua $f : [0, 1] \rightarrow \mathcal{C}$ tal que $f(0) = x_0$ e $f(1) = x_1$.

Proposição 1.7. Se $t_1 \leq t_2$, então $\mathcal{C}(t_1) \subset \mathcal{C}(t_2)$.

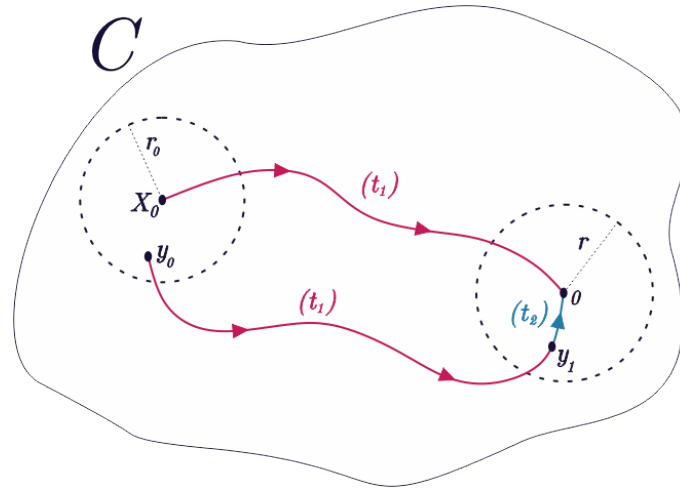


Figura 1: Ilustração da Proposição 1.4

2 Aplicação

Exemplo 2.1. Consideraremos um sistema instável com dois componentes, ou seja, existe um desvio da origem. Se o sistema for **incontrolável**, ele será levado a um estado de desequilíbrio. Se o sistema for **controlável**, teremos as seguintes equações de estado:

$$\begin{cases} x'_1 = x_1 + u_1 \\ x'_2 = x_2 + u_1 \end{cases},$$

com $u_1 \in \mathcal{U}_b$, ou seja, $|u_1| \leq 1$.

O sistema acima é equivalente a (1), com $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. De [2], sabemos que a solução de um sistema da forma (1) é dada por:

$$x_0 = - \int_0^{t_1} e^{-sA} B u(s) ds,$$

com $x_0 \in \mathcal{C}(t_1)$. Do exemplo, como $A = I_d$, matriz identidade de ordem 2, temos:

$$x_1 = x_2 = - \int_0^{t_1} e^{-I_d s} B u_1(s) ds.$$

Mas sabemos que $e^{-I_d s} = \sum_{k=0}^{\infty} \frac{(-s)^k I_d^k}{k!} = I_d \sum_{k=0}^{\infty} \frac{(-s)^k}{k!} = I_d e^{-s}$. Logo, $x_1 = x_2 = - \int_0^{t_1} e^{-s} u_1(s) ds$ e, como $|u_1| \leq 1$:

$$|x_1| = \left| - \int_0^{t_1} e^{-s} u_1(s) ds \right| \leq \int_0^{t_1} |e^{-s}| |u_1(s)| ds = \int_0^{t_1} e^{-s} ds = -e^{-t_1} + 1.$$

Dessa forma, podemos concluir que $\mathcal{C}(t_1) = \{(x_1, x_2); |x_1| \leq -e^{-t_1} + 1\}$ e $\mathcal{C} = \{(x_1, x_2); |x_1| < 1\}$.

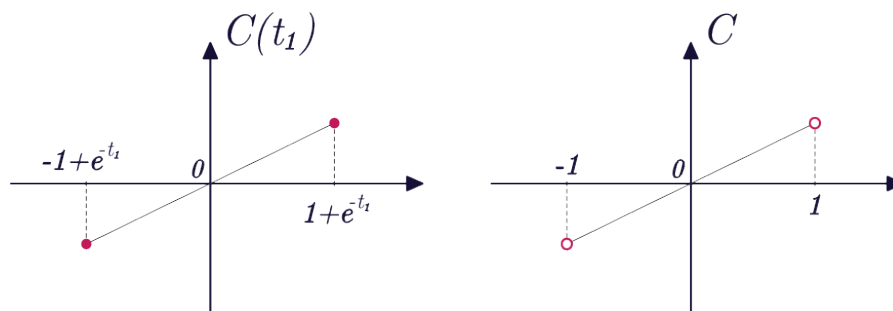


Figura 2: Ilustração do Exemplo 2.1

3 Considerações finais

Esse trabalho teve como objetivo apresentar uma teoria inicial sobre o estudo da controlabilidade em sistemas dinâmicos lineares e familiarizar o leitor com algumas de suas aplicações. A teoria mais aprofundada, juntamente com outros exemplos, podem ser facilmente encontrados nas referências aqui citadas.

Agradecimentos

Agradeço a Secretaria de Educação Superior (Sesu/MEC) pelo apoio financeiro em relação ao benefício que recebo enquanto bolsista do grupo PET Matemática, ao próprio grupo PETMAT pela oportunidade e a Universidade Federal de Uberlândia.

Referências

- [1] BOYCE, W. E. e DIPRIMA, R. C. **Equações Diferenciais Elementares: e Problemas de Valores de Contorno**. Livros Técnicos e Científicos. Rio de Janeiro: Editora Ltda, 1999.
- [2] LEITÃO, A e SILVA, G. N. **Tópicos em Teoria de Controle** Disponível em: <<http://www.bienasbm.ufba.br/M44.pdf>>. Acesso em: 20 abr. 2021.
- [3] MENDES, A. M. **Uma breve introdução ao estudo da análise do \mathbb{R}^n** . 2017. 67 f. Dissertação (Mestrado Profissional em Matemática) - Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, João Pessoa - PB, 2017. Disponível em: <<https://repositorio.ufpb.br/jspui/bitstream/tede/9834/2/Arquivototal.pdf>>. Acesso em: 20 abr. 2021.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

24 A 26 DE MAIO DE 2021

O Inverso de Dirichlet e a Fórmula de Inversão de Möbius

The Dirichlet inverse and the Möbius inversion formula

Fernanda de Andrade Flor

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

fernandaflor@ufu.br

Victor Gonzalo Lopez Neumann

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

victor.neumann@ufu.br

Resumo

Palavras-chave

Inverso de Dirichlet.
Fórmula de Inversão de Möbius.
Produto de Dirichlet.
Funções Aritméticas Multiplicativas.
Função Totiente de Euler.

Este trabalho apresenta um aprofundamento na área da Teoria dos Números, sendo o foco o inverso de Dirichlet e a demonstração da fórmula de inversão de Möbius. A apresentação se inicia definindo funções aritméticas multiplicativas, funções nas quais $f(mn) = f(m)(n)$ com m e n primos entre si, e mostrando alguns exemplos e resultados sobre tais funções, além de relacionar a função totiente de Euler com a função de Möbius, exemplificando a fórmula de inversão de Möbius. Com isso conseguimos definir uma nova operação chamada “produto de Dirichlet”, a qual relaciona duas funções multiplicativas em um somatório. Definimos suas propriedades e a operação inversa deste produto, apresentando também a inversa de algumas funções inicialmente definidas. Com todos esses resultados podemos demonstrar a fórmula de inversão de Möbius, já enunciada, de maneira mais intuitiva, concluindo o objetivo do trabalho.

Abstract

Keywords

Dirichlet Inverse.
Möbius Inversion Formula.
Dirichlet Product.
Multiplicative Arithmetic Functions.
Euler's Totient Function.

This academic work presents a deepening in the area of Number Theory, focusing on the Dirichlet inverse and the Möbius inversion formula demonstration. The presentation begins by defining multiplicative arithmetic functions, functions in which $f(mn) = f(m)(n)$ with m and n relative primes, and showing some examples and results about such functions, in addition to relating Euler's totient function to the Möbius function, exemplifying the Möbius inversion formula. As a result, we're able to define a new operation called “Dirichlet product”, which relates two multiplicative functions in a summation. We define its properties and the inverse operation of this product, also proving the inverse of some initially defined functions. With all these results, we can demonstrate the Möbius inversion formula, already enounced, in a more intuitive way, concluding the purpose of the work.

1 Introdução

Para uma melhor compreensão deste estudo é necessário apresentarmos algumas definições e resultados a respeito das funções aritméticas multiplicativas.

Definição 1.1 (Funções aritméticas multiplicativas). *Uma função multiplicativa é uma função aritmética (não nula) tal que $f(mn) = f(m)f(n)$ para todo par m, n de inteiros positivos primos entre si. Já uma função aritmética na qual vale $f(mn) = f(m)f(n)$ para quaisquer m e n é chamada de completamente multiplicativa.*

Além de ser um exemplo de função aritmética, a função ϕ de Euler também é multiplicativa e isso nos permite provar diversas propriedades dessa função.

Definição 1.2 (Função ϕ de Euler). *Seja n um inteiro positivo. Definimos a função $\phi(n)$ como sendo o número de inteiros j tal que $\text{mdc}(j, n) = 1$ com $1 \leq j \leq n$. Essa função é conhecida como função totiente de Euler.*

Com essa definição podemos dizer que a quantidade de números j menores que 3 tais que $\text{mdc}(j, 3) = 1$ é igual a 2, pois teremos $j = 1$ e $j = 2$ que cumprem os requisitos. Em outras palavras $\phi(3) = 2$.

Teorema 1.3. *Para qualquer inteiro positivo n temos que $\sum_{d|n} \phi(d) = n$.*

A demonstração do teorema acima pode ser encontrada em [1], Teorema 2.2.

Definição 1.4 (Função μ de Möbius). *Seja $n = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ a fatoração em primos distintos p_i com $i \in \{1, \dots, r\}$. A função μ de Möbius é definida por*

$$\mu(n) = \begin{cases} 1, & \text{se } n = 1; \\ (-1)^r, & \text{se } a_j = 1 \text{ para todo } j; \\ 0, & \text{se } a_j > 1 \text{ para algum } j. \end{cases} \quad (1)$$

Portanto, essa definição nos diz que $\mu(n) = 0$ se n for divisível pelo quadrado de algum primo. Por exemplo, temos que $\mu(2) = -1$, pois $2 = 2^1$. Também é possível, a partir dessa definição, desenvolver um teorema sobre o somatório da função de Möbius dos divisores de um n inteiro positivo.

Teorema 1.5. *Se n um inteiro positivo, temos*

$$F(n) = \sum_{d|n} \mu(d) = \begin{cases} 1, & \text{se } n = 1, \\ 0, & \text{se } n > 1. \end{cases} \quad (2)$$

A demonstração do teorema acima pode ser encontrada em [2], Teorema 4.7.

Definição 1.6 (Função λ de Liouville). *Definimos $\lambda(1) = 1$ e, se $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ é a fatoração de n em primos distintos, definimos*

$$\lambda(n) = (-1)^{a_1+a_2+\dots+a_k}.$$

A definição nos mostra que λ é completamente multiplicativa.

1.1 Uma relação entre as funções ϕ e μ

É possível desenvolver relações entre as funções aritméticas, nesta seção estudaremos as relações encontradas entre as funções ϕ de Euler e μ de Möbius.

Teorema 1.7. *Para todo inteiro $n \geq 1$, temos*

$$\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}. \quad (3)$$

A demonstração do teorema acima pode ser encontrada em [2], Teorema 4.11.

Teorema 1.8 (Fórmula de inversão de Möbius). *Se duas funções aritméticas $f(n)$ e $g(n)$ satisfazem uma das duas condições abaixo*

$$f(n) = \sum_{d|n} g(d), \quad g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) \quad (4)$$

para todo n , então elas satisfazem as duas condições.

A demonstração do teorema acima pode ser encontrada em [2], Teorema 4.12.

Esse tipo de soma do Teorema 1.8 será mostrada na próxima seção.

2 O Produto de Dirichlet e as funções aritméticas

A soma mostrada anteriormente $\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}$ é uma soma que aparece frequentemente na Teoria dos Números. Essas somas possuem a forma

$$\sum_{d|n} f(d) g\left(\frac{n}{d}\right), \quad (5)$$

onde f e g são funções aritméticas. É importante tratarmos essas somas como um novo tipo de multiplicação de funções aritméticas.

Definição 2.1 (Produto de Dirichlet). *Se f e g são duas funções aritméticas, definimos seu produto de Dirichlet como sendo a função h definida pela equação:*

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right). \quad (6)$$

Escrevemos $f * g$ para h e $(f * g)(n)$ para $h(n)$. O símbolo N será usado para a função aritmética na qual $N(n) = n$ para todo n .

Nessa notação, o Teorema 1.7 da relação entre as funções ϕ e μ pode ser denotado como:

$$\phi = \mu * N. \quad (7)$$

Teorema 2.2. *O produto de Dirichlet é comutativo e associativo. Ou seja, para quaisquer funções aritméticas f, g e k temos:*

$$\begin{aligned} f * g &= g * f \quad (\text{comutativa}); \\ (f * g) * k &= f * (g * k) \quad (\text{associativa}). \end{aligned}$$

A demonstração do teorema acima pode ser encontrada em [1], Teorema 2.6.

Agora iremos introduzir o elemento identidade para essa multiplicação:

Definição 2.3 (Identidade do Produto de Dirichlet). *A função aritmética I é dada por*

$$I(n) = \left[\frac{1}{n} \right] = \begin{cases} 1, & \text{se } n = 1, \\ 0, & \text{se } n > 1 \end{cases} \quad (8)$$

e é chamada de função identidade.

Teorema 2.4. *Para todo f temos que $I * f = f * I = f$.*

A demonstração do teorema acima pode ser encontrada em [1], Teorema 2.7.

3 O inverso de Dirichlet e fórmula de inversão de Möbius

3.1 O inverso de Dirichlet

A partir da definição da função identidade na última seção apresentada, podemos desenvolver uma nova propriedade do produto de Dirichlet, o inverso de Dirichlet.

Teorema 3.1. *Se f é uma função aritmética com $f(1) \neq 0$, há apenas uma função aritmética f^{-1} , chamada inverso de Dirichlet de f , tal que*

$$f * f^{-1} = I. \quad (9)$$

Mais ainda, f^{-1} é dada pelas fórmulas:

$$f^{-1}(1) = \frac{1}{f(1)}, \quad f^{-1}(n) = \frac{-1}{f(1)} \sum_{\substack{d|n \\ d < n}} f\left(\frac{n}{d}\right) f^{-1}(d), \quad \text{para } n > 1. \quad (10)$$

A demonstração do teorema acima pode ser encontrada em [1], Teorema 2.8.

Abaixo temos uma tabela na qual relacionamos algumas funções aritméticas multiplicativas com suas inversas em relação ao produto de Dirichlet.

Função aritmética	Inverso de Dirichlet
Função constante de valor 1	Função μ de Möbius
Função λ de Liouville	Valor absoluto da função de Möbius $ \mu $
Função ϕ totiente de Euler	$\sum_{d n} (d \cdot \mu(d))$

Tabela 1: Algumas funções aritméticas multiplicativas e suas inversas

Podemos agora exemplificar a primeira linha da tabela. Seja $h(n)$ o produto de Dirichlet das funções, logo temos que seu valor é a função identidade, de fato:

i) se $n = 1$,

$$h(1) = \sum_{d|1} \mu(d) \cdot 1 = 1 \cdot 1 = 1 = I(1); \quad (11)$$

ii) se $n > 1$,

$$h(n) = \sum_{d|n} \mu(d) \cdot 1 = 0 \cdot 1 = 0 = I(n); \quad (12)$$

no qual a segunda igualdade de cada situação segue do Teorema 1.5.

3.2 Demonstração da fórmula de inversão de Möbius

Definição 3.2. Definimos a função u como sendo a função aritmética tal que $u(n) = 1$ para todo n .

O Teorema 1.5 nos mostra que $\sum_{d|n} \mu(d) = I(n)$. Utilizando o produto de Dirichlet podemos enunciá-lo da seguinte maneira:

$$\mu * u = I.$$

Logo u e μ são os inversos de Dirichlet um do outro:

$$u = \mu^{-1} \quad e \quad \mu = u^{-1}.$$

De acordo com essa propriedade da função de Möbius e a propriedade associativa do produto de Dirichlet podemos demonstrar a fórmula de inversão de Möbius (Teorema 1.8) de uma maneira direta e intuitiva:

Demonstração. Podemos escrever a primeira equação do teorema como sendo $f = g * u$. Multiplicando ambos os lados por μ temos $f * \mu = (g * u) * \mu$. Agora utilizando a associativa do produto de Dirichlet, como u e μ são inversos um do outro, concluímos que:

$$f * \mu = g * (u * \mu) = g * I = g.$$

Portanto $g = f * \mu$, ou seja, a segunda equação.

Para provar o que falta, basta multiplicarmos a segunda equação em produto de Dirichlet por u , logo teremos:

$$\begin{aligned} g &= f * \mu \\ \Rightarrow g * u &= f * (\mu * u) \\ \Rightarrow g * u &= f * I = f, \end{aligned}$$

ou seja, a primeira equação, comprovando assim o teorema. ■

A fórmula de inversão de Möbius já tinha sido utilizada em um par de fórmulas nos Teoremas 1.5 e 1.7:

$$n = \sum_{d|n} \phi(d) \quad \text{e} \quad \phi(n) = \sum_{d|n} \mu(d) \left(\frac{n}{d}\right).$$

4 Conclusão

Ao final deste trabalho concluímos que o objetivo inicial (demonstração intuitiva da fórmula de inversão de Möbius) foi cumprido, além de possibilitar um grande e importante estudo sobre a Teoria dos Números. Para resultados e demonstrações mais detalhadas consultar as referências ao final do trabalho.

Agradecimentos

Agradeço à SESu-MEC pelo auxílio financeiro, estando na condição de bolsista no grupo PET Matemática da Universidade Federal de Uberlândia.

Referências

- [1] APOSTOL, T. M.. **Introduction to Analytic Number Theory**. Berlin: Springer-Verlag, 2 ed., 1976.
- [2] SANTOS, J. P. O.. **Introdução à Teoria dos Números**. Rio de Janeiro: SBM, 1 ed., 1998.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Núcleos e centro de um quasegrupo Nuclei and center of a quasigroup

Gabriel Simão Mucci

UFU, Faculdade de Matemática/UFU, Uberlândia, Minas Gerais, Brasil

gabriel.mucci@ufu.br

Dylene Agda Souza de Barros

UFU, Faculdade de Matemática/UFU, Uberlândia, Minas Gerais, Brasil

dylene@ufu.br

Palavras-chave

Grupóides.
Quasegrupos.
Loops.
Núcleos.
Centro.

Resumo

Neste trabalho apresentaremos os conceitos de núcleo e centro de um quasegrupo e o que a existência destes implica na estrutura do quasegrupo.

Keywords

Grupoids.
Quasigroups.
Loops.
Nuclei.
Center.

Abstract

In this work we shall present the concepts of nucleous and center for a quasigroup and what their existence implies into the structure of the quasigroup.

1 Introdução

Para duas das operações básicas não vale a propriedade associativa e esse fato, por si só, já justificaria o estudo das estruturas não associativas. Entretanto, a teoria de quasegrupos e loops é uma área da Matemática razoavelmente recente cujas raízes vêm da geometria, álgebra e combinatória.

Apesar do estudo de estruturas não associativas ser uma área da álgebra rica, independente e com vida própria, nesse trabalho queremos mostrar que a existência de elementos que associam com todos os elementos da estrutura (que mais adiante os chamaremos de *elementos nucleares*) implicam em relevantes consequências para a estrutura do quasegrupo como um todo.

O assunto tratado neste trabalho é baseado no texto [1]. Para mais detalhes, sugerimos a consulta do primeiro capítulo da referência citada.

2 Grupóides, Quasegrupos e Loops

Nesta seção apresentaremos os primeiros objetos do nosso trabalho, que é base para o seu desenvolvimento.

Um **grupóide** é um conjunto não vazio, munido de uma operação binária. Sempre que falamos de um **grupóide** (G, \cdot) , queremos dizer que G denota um conjunto não vazio e que " \cdot " é o símbolo usado para denotar a operação binária em G .

Seja (G, \cdot) um grupóide e seja a um elemento fixado em G . Definimos as **funções de translação à esquerda** e **à direita**, $L(a), R(a) : G \rightarrow G$, respectivamente, por

$$xL(a) = a \cdot x \text{ e } xR(a) = x \cdot a, \text{ para todo } x \in G.$$

Definição 2.1. O grupóide (G, \cdot) é comutativo quando $L(a) = R(a)$ para todo $a \in G$.

Definição 2.2. O grupóide (G, \cdot) é associativo é quando $R(a \cdot b) = R(a)R(b)$ para todo $a, b \in G$.

Definição 2.3. Seja (G, \cdot) um grupóide e seja $\varrho \in G$. Então ϱ ser um elemento de **identidade à esquerda** (**à direita**) para (G, \cdot) significa que $L(\varrho) : G \rightarrow G$ ($R(\varrho) : G \rightarrow G$) é a função identidade de G . Dizemos que, ϱ é um **elemento identidade** para (G, \cdot) se ϱ é um elemento identidade à esquerda e à direita para (G, \cdot) .

Assim,

$$\begin{aligned} L(\varrho) = id_{\varrho} &\implies xL(\varrho) = x, \forall x \in G \implies \varrho x = x, \forall x \in G \\ R(\varrho) = id_{\varrho} &\implies xR(\varrho) = x, \forall x \in G \implies x\varrho = x, \forall x \in G \end{aligned}$$

Um grupóide (G, \cdot) é chamado **quasegrupo** se as funções $L(a) : G \rightarrow G$ e $R(a) : G \rightarrow G$ são bijetoras, para todo $a \in G$. Equivalentemente, dada a equação $x \cdot y = z$ e sabendo quaisquer dois elementos de $\{x, y, z\} \subset G$, pode-se determinar o terceiro elemento.

Em um quasegrupo satisfaz são válidas as leis de cancelamento à direita e à esquerda, para $a, x, y \in G$

$$\begin{aligned}x \cdot a = y \cdot a &\implies x = y \quad \text{e} \\ a \cdot x = a \cdot y &\implies x = y\end{aligned}$$

De fato, se $a \cdot x = a \cdot y = b$, então x e y são soluções de $a \cdot u = b$, donde concluímos que $x = y$. De forma análoga, se $x \cdot a = y \cdot a = b$, então x e y são soluções de $v \cdot a = b$ e então $x = y$.

O próximo resultado nos fala das consequências da associatividade em um quasegrupo.

Teorema 2.4. *Se (G, \cdot) é um quasegrupo que é associativo, então (G, \cdot) tem um (único) elemento identidade.*

Observe que em um quasegrupo associativo, se ϱ é seu elemento identidade, as equações $ax = \varrho$ e $ya = \varrho$ possuem respectivas soluções a^ρ e a^λ . Daí, pela associatividade, temos

$$a^\rho = \varrho a^\rho = (a^\lambda a) a^\rho = a^\lambda (a a^\rho) = a^\lambda \varrho = a^\lambda$$

Portanto, **grupos** são, precisamente, aqueles quasegrupos que são associativos. Entretanto, existem quasegrupos, com elemento identidade, que não são associativos.

Definição 2.5. *Um loop é um quasegrupo (G, \cdot) que possui elemento identidade.*

3 Núcleos e centro de um quasegrupo

Seja (G, \cdot) um grupóide e seja $a \in G$. Às vezes, o papel que a desempenha como um membro de (G, \cdot) pode ser descrito ou caracterizado pelo comportamento das funções de translação $L(a) : G \rightarrow G$ e $R(a) : G \rightarrow G$, quais foram introduzidos na Seção 2.

Definição 3.1. *Seja (G, \cdot) um grupóide e seja $a \in G$.*

a) Dizemos que, a é nuclear à esquerda (central ou à direita) se $L(a \cdot x) = L(x) \cdot L(a)$ ($L(x \cdot a) = L(a)L(x)$), $R(x \cdot a) = R(x)R(a)$ para todo $x \in G$. O elemento a é dito nuclear em (G, \cdot) se a é nuclear à esquerda, central e à direita em (G, \cdot) ;

b) O núcleo à esquerda, N_λ , (o núcleo central N_μ , o núcleo à direita N_ρ) de (G, \cdot) é o conjunto de todos os elementos nucleares à esquerda (centrais, à direita) em (G, \cdot) . O núcleo N de (G, \cdot) é definido por $N = N_\lambda \cap N_\mu \cap N_\rho$.

Pela definição acima, temos que

$$\begin{aligned}N_\lambda &= \{a \in G \mid a \cdot (xy) = (ax) \cdot y, \forall x, y \in G\}. \\ N_\mu &= \{a \in G \mid (xa) \cdot y = x \cdot (ay), \forall x, y \in G\}. \\ N_\rho &= \{a \in G \mid (xy) \cdot a = x \cdot (ya), \forall x, y \in G\}.\end{aligned}$$

Em um grupóide (G, \cdot) não existe garantia que os elementos nucleares existam e, portanto, qualquer núcleo N_λ, N_μ, N_ρ ou N pode ser o conjunto vazio. Por exemplo, para o grupóide $A = \{1, 2, 3\}$ com a operação (\cdot) definida pela tabela de Cayley

\cdot	1	2	3
1	1	2	2
2	1	1	2
3	2	2	2

todos os três núcleos N_λ, N_μ e N_ρ são vazios.

Sempre que $N_\lambda (N_\mu, N_\rho)$ é não vazio, então $N_\lambda (N_\mu, N_\rho)$ é um subgrupóide (associativo) de (G, \cdot) . O próximo resultado nos mostra que a existência de elementos nucleares influencia a estrutura de um quasegrupo. A demonstração deste resultado pode ser encontrada no Teorema I.3.4 de [1].

Teorema 3.2. *Seja (G, \cdot) um quasegrupo.*

1. *Se $N_\mu \neq \emptyset$, então N_μ é um subgrupo de (G, \cdot) e o elemento identidade ϱ de (N_μ, \cdot) é o elemento identidade de (G, \cdot) .*
2. *Se $N_\lambda \neq \emptyset$, então N_λ é um subgrupo de (G, \cdot) e o elemento identidade de (N_λ, \cdot) é o elemento identidade esquerdo de (G, \cdot) .*
3. *Se $N_\rho \neq \emptyset$, então N_ρ é um subgrupo de (G, \cdot) e o elemento identidade de (N_ρ, \cdot) é o elemento identidade direito de (G, \cdot) .*

Exemplo 3.3. *Considere o quasegrupo $G = \{e, x, y\}$ com operação (\cdot) definida pela tabela de Cayley*

	ϱ	x	y
ϱ	ϱ	x	y
x	y	ϱ	x
y	x	y	ϱ

Temos que $N_\lambda = \{\varrho\}$ e ϱ é um elemento identidade esquerdo para (G, \cdot) , entretanto ϱ não é um elemento identidade para (G, \cdot) .

Assim, quando (G, \cdot) é um quasegrupo que não é um loop, não se pode esperar melhorar as declarações (1) e (2) do **Teorema 3.2**.

Se (G, \cdot) é um loop, sabemos que (G, \cdot) compartilha seu elemento identidade em cada um de seus subloops. Além de que, os núcleos N_λ, N_μ e N_ρ serão subgrupos de (G, \cdot) .

Definição 3.4. *Seja (G, \cdot) um grupóide com núcleo N . O centro Z de (G, \cdot) é definido por*

$$Z = \{a \in N \mid L(a) = R(a)\}.$$

Assim, um elemento $a \in G$ é um membro do centro Z , do grupóide (G, \cdot) , se, e somente se, para todo $x, y \in G$,

$$a \cdot (xy) = (ax) \cdot y, \quad (xa) \cdot y = x \cdot (ay), \quad (xy) \cdot a = x \cdot (ya) \quad \text{e} \quad a \cdot x = x \cdot a,$$

Quando (G, \cdot) é um grupóide com respectivos núcleo e centro N e Z , se N e Z são não vazios, eles herdam a estrutura de (G, \cdot) , isto é, eles são subgrupóides de (G, \cdot) .

Se assumirmos que (G, \cdot) é um quasegrupo e se N e Z são não vazios, eles têm uma estrutura a mais:

Teorema 3.5. *Seja (G, \cdot) um quasegrupo com núcleo N e centro Z . Se N é não vazio, então Z é não vazio e N e Z são subgrupos de (G, \cdot) , com Z sendo um subgrupo comutativo de (N, \cdot) .*

Teorema 3.6. *Seja (G, \cdot) um loop com núcleo N e centro Z . Então N e Z são subgrupos de (G, \cdot) com Z sendo um subgrupo comutativo de (N, \cdot) .*

Exemplo 3.7. *Seja $G = \{1, 2, 3, 4, 5\}$ com operação (\cdot) dada pela tabela de Cayley a seguir:*

\cdot	1	2	3	4	5
1	1	2	3	4	5
2	2	1	5	3	4
3	3	5	4	2	1
4	4	3	1	5	2
5	5	4	2	1	3

Temos que o $N_\lambda = \{1\}$, o $N_\rho = \{1\}$ e o $N_\mu = \{1\}$.

4 Considerações finais

Podemos perceber que, em um quasegrupo não associativo, a existência de um elemento nuclear à esquerda (à direita) implica na existência de um elemento identidade à esquerda (à direita). Além disso, vimos que um quasegrupo que possui elemento nuclear central necessariamente é um loop.

Agradecimentos

O autor Gabriel Simão Mucci é bolsista do Programa de Educação Tutorial PET-Matemática da Universidade Federal de Uberlândia e agradece o auxílio financeiro recebido durante o desenvolvimento deste trabalho.

Referências

- [1] PFLUGFELDER, H. O., **Quasigroups and Loops: Introduction**. Berlin: Sigma Series in Pure Mathematics, Heldermann Verlag, 1990.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

24 A 26 DE MAIO DE 2021

Uma definição de caos e sua redução

A definition of chaos and its reduction

Guilherme Daniel Nogueira

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
guilherme.daniel@ufu.br

Jean Venato Santos

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
jvenatos@ufu.br

Resumo

Palavras-chave

Sistemas dinâmicos
Caos
Aplicação shift
Redução do caos

Neste trabalho serão considerados sistemas dinâmicos discretos que consistem em funções definidas em espaços nos quais é possível iterar repetidamente e assim estudar os sucessivos pontos que surgirão a cada iteração. Mais especificamente são apresentados sistemas que apresentam comportamento caótico. Será explorada a definição de caos apresentada no popular texto do Devaney [1] que capta tal comportamento em três propriedades, a saber: densos pontos periódicos, existência de órbita densa e sensibilidade às condições iniciais. Em seguida será apresentada uma redução desta noção de caos, proposta por Banks et al [2], na qual a terceira condição é obtida como consequência das duas primeiras.

Abstract

Keywords

Dynamical systems
Chaos
Shift map
Chaos reduction

In this work, discrete dynamical systems will be considered, which consist of functions defined in spaces in which it is possible to iterate repeatedly and thus study the successive points that will appear at each iteration. More specifically, systems that exhibit chaotic behavior will be investigated. The definition of chaos presented in the popular Devaney's text [1] that captures such behavior in three properties will be explored, namely: dense periodic points, existence of dense orbit and sensitivity to initial conditions. Then, a reduction of this notion of chaos, proposed by Banks et al [2], will be presented, in which the third condition is obtained as a consequence of the first two.

1 Introdução

Em 19 de fevereiro de 1998, computadores do sistema de previsão de tempestades tropicais dos Estados Unidos diagnosticaram a formação de uma tempestade tropical sobre Louisiana em três dias. Sobre o Oceano Pacífico um meteorologista daquela agência descobriu que havia uma pequena diferença nas medições executadas, e que estas poderiam prever uma pequena diferença no deslocamento das massas de ar. A diferença foi detectada através de uma movimentação do ar em maior velocidade na região do Alasca. Em função das diferenças, houve uma realimentação de dados nos computadores, estes refazendo os cálculos previram que a formação da tempestade tropical em Louisiana não ocorreria, mas haveria sim a formação de um tornado de proporções gigantescas em Orlando, na Flórida, o que realmente ocorreu em 22 de fevereiro de 1998 ([3]).

Fenômenos como o descrito acima, nos quais erros mínimos nos dados utilizados podem gerar grandes discrepâncias são popularmente conhecidos como Efeito Borboleta. Este efeito foi analisado pela primeira vez em 1963 pelo meteorologista, matemático e filósofo estadunidense Edward Lorenz, que o descobriu quando estava trabalhando com previsões meteorológicas no MIT e verificou a influência ocasionada em sistemas dinâmicos quando são feitas alterações muito pequenas nos dados iniciais inseridos em computadores numéricos programados para fazerem cálculos em série ([3]).

Na Matemática, o estudo dos sistemas com sensibilidade em relação aos dados iniciais se insere no contexto da Teoria do Caos que por sua vez é um tópico dentro da grande área de sistemas dinâmicos. Um sistema dinâmico discreto consiste numa função em um certo espaço que podemos iterar repetidamente e assim obter valores de certa variável correspondentes a momentos sucessivos.

Em termos matemáticos, o que seria caos? Ou seja, quando se diz que um sistema dinâmico é caótico? Apesar de não haver uma definição matemática de caos universalmente adotada, no difundido livro texto [1] sobre o assunto, Devaney indica três propriedades que expressam a essência da manifestação do caos.

Estas propriedades são formuladas para funções $f: X \rightarrow X$ definidas em algum espaço métrico X . A primeira delas se refere ao conjunto dos pontos periódicos de f , ou seja, pontos $x \in X$ tais que $f^n(x) = x$, para algum natural n . A propriedade é que o conjunto dos pontos periódicos seja denso em X . Esta abundância de pontos periódicos indica, em certo sentido, um elemento de regularidade no sistema. A segunda característica de sistemas caóticos é a *transitividade*, a qual estabelece que para qualquer par de pontos x e y em X , e qualquer $\varepsilon > 0$, existe um ponto z a uma distância de até ε de x , e $f^n(z)$ dista no máximo ε de y para algum $n \in \mathbb{N}$. A transitividade é uma propriedade de irreduzibilidade, pois ela implica que o sistema não pode ser decomposto em subsistemas não triviais. A condição final é a *sensibilidade às condições iniciais* que ocorre se existir um $\beta > 0$ tal que para quaisquer x em X e $\varepsilon > 0$, há um y que está ε -próximo de x e um k natural tais que a distância entre $f^k(x)$ e $f^k(y)$ é pelo menos β . Esta propriedade captura o fato de que o efeito borboleta está presente em sistemas caóticos e portanto é amplamente aceita como a condição central da noção de caos.

Em resumo:

Definição 1.1. [Caos de Devaney [1]] *Seja X um espaço métrico. Uma função $f: X \rightarrow X$ é dita caótica em X se:*

1. *os pontos periódicos de f são densos em X ,*
2. *f é transitiva,*
3. *f depende sensivelmente das condições iniciais.*

2 Aplicação shift: um exemplo de função caótica

A fim de apresentar um exemplo de sistema caótico introduziremos a aplicação shift a qual é definida no **espaço de sequências** dado por $\Sigma = \{(s_0s_1s_2\dots) \mid s_j = 0 \text{ ou } 1\}$, ou seja, no conjunto de todas as sequências possíveis de 0's e 1's. Considere a função $d: \Sigma \times \Sigma \rightarrow \mathbb{R}$ dada por

$$d[s, t] = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}. \quad (1)$$

Note que o máximo desta função ocorre entre dois pontos de Σ satisfazendo $s_i \neq t_i$, para todo $i = 0, 1, 2, \dots$. Neste caso, $d[s, t] = 2^{-i}$. Isto mostra que a série em (1) é sempre convergente.

Proposição 2.1. *A função d , definida em (1), estabelece uma métrica no conjunto Σ .*

Demonstração. Sejam $s = (s_0s_1s_2\dots), t = (t_0t_1t_2\dots), u = (u_0u_1u_2\dots) \in \Sigma$. Como o termo geral da série que define d é dado por $\frac{|s_i - t_i|}{2^i} \geq 0$, segue que $d[s, t] \geq 0$, e $d[s, t] = 0$ se e somente se $s_i = t_i$, para todo $i = 0, 1, 2, \dots$, ou seja, $s = t$. Como $|s_i - t_i| = |t_i - s_i|$ temos que $d[s, t] = d[t, s]$. Por último, temos que para quaisquer 3 números reais s_i, t_i e u_i , vale a desigualdade triangular:

$$|s_i - u_i| \leq |s_i - t_i| + |t_i - u_i|,$$

desse fato, podemos ver que:

$$d[s, u] = \sum_{i=0}^{\infty} \frac{|s_i - u_i|}{2^i} \leq \sum_{i=0}^{\infty} \frac{|s_i - t_i| + |t_i - u_i|}{2^i} = d[s, t] + d[t, u].$$

Portanto, d satisfaz as três propriedades de métrica. ■

Uma interessante característica do conjunto Σ é a facilidade de se medir distâncias entre seus elementos, o que é fomentado pelo seguinte resultado:

Teorema 2.2 (Teorema da Proximidade). *Sejam $s, t \in \Sigma$ e vamos supor $s_i = t_i$ para $i = 0, 1, 2, \dots, n$. Então $d[s, t] \leq \frac{1}{2^n}$. Ademais, se $d[s, t] < \frac{1}{2^n}$, então $s_i = t_i$ para $i \leq n$.*

Demonstração. Se $s_i = t_i$ para $i \leq n$, então

$$d[s, t] = \sum_{i=0}^n \frac{|s_i - t_i|}{2^i} + \sum_{i=n+1}^{\infty} \frac{|s_i - t_i|}{2^i} \leq 0 + \sum_{i=n+1}^{\infty} \frac{1}{2^i} = \frac{1}{2^n}.$$

Se $s_j \neq t_j$ para algum $j \leq n$, então: $d[s, t] \geq \frac{1}{2^j} \geq \frac{1}{2^n}$. Logo, se $d[s, t] < \frac{1}{2^n}$, então $s_i = t_i$ para $i \leq n$. O que encerra a prova do teorema. ■

Definição 2.3. A aplicação shift $\sigma : \Sigma \rightarrow \Sigma$ é definida por: $\sigma(s_0s_1s_2\dots) = (s_1s_2s_3\dots)$.

Proposição 2.4. A aplicação $\sigma : \Sigma \rightarrow \Sigma$ é contínua.

Demonstração. Dados $\varepsilon > 0$ e $s = (s_0s_1s_2\dots) \in \Sigma$, existe natural n tal que $\frac{1}{2^n} < \varepsilon$, tome $\delta = \frac{1}{2^{n+1}}$. Dado r em Σ tal que $d[r, s] < \delta$, então, pelo Teorema da Proximidade, temos que $s_i = r_i$, para $i = 0, 1, 2, \dots, n + 1$, isto é, $r = (s_0s_1s_2\dots s_{n+1}r_{n+2}r_{n+3}r_{n+4}\dots)$. Logo, $\sigma(r) = (s_1s_2\dots s_{n+1}r_{n+2}r_{n+3}r_{n+4}\dots)$ tem dígitos correspondentes aos de $\sigma(s)$ para as primeiras $n + 1$ entradas. Do Teorema da Proximidade segue que: $d[\sigma(s), \sigma(r)] \leq \frac{1}{2^n} < \varepsilon$. ■

Teorema 2.5. A aplicação shift $\sigma : \Sigma \rightarrow \Sigma$ é um sistema dinâmico caótico.

Demonstração. (1) O conjunto dos pontos periódicos de σ é denso em Σ .

Mostremos que dada qualquer sequência de Σ na forma $s = (s_0s_1s_2\dots)$, existe um ponto periódico tão próximo quanto se queira de s . Dado um $\varepsilon > 0$, escolhamos um inteiro n , tal que $\frac{1}{2^n} < \varepsilon$ e consideramos a sequência: $r_n = (s_0s_1s_2\dots s_n \overline{s_0s_1s_2\dots s_n})$ onde as primeiras $n + 1$ entradas das sequências s e r_n são as mesmas. Como a barra na definição de r_n indica a repetição indefinida dos termos sob ela, segue que r_n é σ periódico. Do Teorema da Proximidade, $d[s, r_n] \leq \frac{1}{2^n} < \varepsilon$.

(2) Mostraremos agora a transitividade de σ .

Note que um sistema com uma órbita densa apresenta transitividade uma vez que tal órbita passa arbitrariamente próximo de quaisquer pontos. Mostremos então que σ possui uma órbita densa.

Considere o ponto $p = \{ \underbrace{01}_{\text{blocos 1}} \underbrace{00011011}_{\text{blocos 2}} \dots \}$ onde p é a sequência que consiste em todas as possibilidades de blocos de 0's e 1's, de i dígitos, com $i = 1, 2, \dots$. Para mostrar que p tem uma órbita densa, sejam $s = (s_0s_1s_2\dots) \in \Sigma$ e $\varepsilon > 0$ quaisquer. Tomamos n tal que $\frac{1}{2^n} < \varepsilon$ e mostraremos que a órbita de p está a uma distância de até $\frac{1}{2^n}$ de s . Sabemos que em p existe um bloco de $n + 1$ dígitos, cujos dígitos são $s_0s_1s_2\dots s_n$. Supondo que o s_0 deste bloco seja a k -ésima entrada da sequência p . Aplicando shift k vezes em p , temos $\sigma^k(p) = (s_0s_1s_2\dots s_n\dots)$, logo pelo Teorema da Proximidade:

$$d[\sigma^k(p), s] \leq \frac{1}{2^n} < \varepsilon.$$

Sendo assim, a órbita de p é densa, e logo σ apresenta transitividade.

Na próxima seção mostraremos que a densidade de órbitas periódicas aliada à transitividade implicam na sensibilidade às condições iniciais, completando assim a prova de que o shift é caótico. ■

3 Redução no caos de Devaney

Em [2] Banks et al mostraram a seguinte redução na Definição 1.1:

Teorema 3.1 (Banks et al [2]). *Seja X um espaço métrico. Se uma função contínua $f: X \rightarrow X$ é transitiva e possui densas órbitas periódicas então f depende sensivelmente das condições iniciais.*

Demonstração. Supondo que f seja transitiva e possua densos pontos periódicos, existe um $\delta_0 > 0$, tal que para todo $x \in X$ existe um ponto periódico $q \in X$ cuja órbita $O(q)$ dista pelo menos $\frac{\delta_0}{2}$ de x . Escolhendo dois pontos periódicos q_1 e q_2 , com órbitas $O(q_1)$ e $O(q_2)$ disjuntas, denotemos δ_0 a distância entre $O(q_1)$ e $O(q_2)$. Então, denotando por d a métrica em X , por desigualdade triangular,

$$d(O(q_1), O(q_2)) \leq d(O(q_1), x) + d(x, O(q_2)) \Rightarrow \delta_0 \leq d(O(q_1), x) + d(x, O(q_2)).$$

Logo, todo $x \in X$ está a uma distância de pelo menos $\frac{\delta_0}{2}$ de $O(q_1)$ ou $O(q_2)$.

Vamos mostrar que f é sensível às condições iniciais com constante de sensibilidade $\delta = \frac{\delta_0}{8}$.

Tomemos $x \in X$ e seja N alguma vizinhança de x . Como os pontos periódicos de f são densos, existe um ponto periódico p na intersecção $U = N \cap B_\delta(x)$ de N com a bola aberta $B_\delta(x)$ de raio δ centrado em x . Seja n o período de p . Como mostramos acima, existe um ponto periódico $q \in X$ cuja órbita $O(q)$ dista pelo menos 4δ de x . Considere $V = \bigcap_{i=0}^n f^{-i}(B_\delta(f^i(q)))$.

É visível que V é um conjunto aberto e não-vazio, pois $q \in V$. Assim, como f é transitiva, existe $y \in U$ e $k \in \mathbb{N}$ tal que $f^k(y) \in V$. Seja j a parte inteira de $\frac{k}{n} + 1$. Assim, como k e n são naturais, temos: $\frac{k}{n} < j \leq 1 + \frac{k}{n} = \frac{k+n}{n} \Rightarrow 1 \leq jn - k \leq n$.

Por construção, temos: $f^{nj}(y) = f^{nj-k}(f^k(y)) \in f^{nj-k}(V) \subseteq B_\delta(f^{nj-k}(q))$.

Como $f^{nj}(p) = p$, $d(f^{nj}(p), f^{nj}(y)) = d(p, f^{nj}(y))$. Assim, por desigualdade triangular,

$$d(x, f^{nj-k}(q)) \leq d(x, p) + d(p, f^{nj-k}(q)) \leq d(x, p) + d(p, f^{nj}(y)) + d(f^{nj}(y), f^{nj-k}(q)),$$

$$d(p, f^{nj}(y)) \geq d(x, f^{nj-k}(q)) - d(x, p) - d(f^{nj}(y), f^{nj-k}(q)).$$

Consequentemente, como $d(x, O(q)) \geq 4\delta$, $p \in B_\delta(x)$ e $f^{nj}(y) \in B_\delta(f^{nj-k}(q))$, segue que

$$d(p, f^{nj}(y)) > 4\delta - \delta - \delta = 2\delta.$$

Utilizando novamente a desigualdade triangular, temos que:

$$2\delta < d(f^{nj}(p), f^{nj}(y)) \leq d(f^{nj}(p), f^{nj}(x)) + d(f^{nj}(x), f^{nj}(y)),$$

logo, $d(f^{nj}(x), f^{nj}(y)) > \delta$ ou $d(f^{nj}(p), f^{nj}(x)) > \delta$, de qualquer forma, encontramos um ponto em N tal que sua nj -ésima iterada dista mais que δ de $f^{nj}(x)$. O que completa a prova. ■

4 Considerações finais

Mostramos nesse trabalho a importância dos sistemas dinâmicos, aplicados por exemplo em problemas meteorológicos. Com a definição de caos de Devaney [1], mostramos que a aplicação shift, de extrema importância para o estudo nessa área, apresenta comportamento caótico e ainda, apresentamos uma redução mostrada por Banks et al [2], que torna provar a presença de comportamento caótico de uma função mais simples.

Agradecimentos

Agradecimentos ao PET Matemática UFU, no qual participo na condição de bolsista, diretamente vinculado ao MEC e ao FNDE, que atuam no pagamento de bolsas de fomento à pesquisa.

Referências

- [1] DEVANEY, Robert L. **A first course in chaotic dynamical systems: theory and experiment.** Westview Press: Perseus Books, 1992. 300p.
- [2] BANKS, J., BROOKS, J., CAIRNS, G., DAVIS, G. e STACEY, P.. **On Devaney's definition of chaos,** The American Mathematical Monthly, v. 99, 4, 332-334, 1992.
- [3] EFEITO BORBOLETA. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2019. Disponível em: <https://pt.wikipedia.org/w/index.php?title=Efeito_borboleta&oldid=56343204>. Acesso em: 26 set. 2019.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT
UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Polinômios sobre anéis não comutativos
Polynomials over non commutative rings

Júlia Bernardes Coelho

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

julia.bernardes@ufu.br

Dylene Agda Souza de Barros

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

dylene@ufu.br

Palavras-chave

Polinômios sobre anéis de divisão.
Anéis de divisão.
Quatérnios.
.
.

Resumo

Nesse trabalho tratamos dos anéis de polinômios sobre anéis de divisão, em particular sobre os anel dos quatérnios reais.

Keywords

Polynomials over division rings.
Division rings.
Quaternions.
.
.

Abstract

In this work we deal with the ring of polynomials over division rings, in particular over the ring of real quaternions.

1 Introdução

É comum em cursos de graduação o estudo dos anéis de polinômios sobre corpos como um importante exemplo de anéis euclidianos. Neste trabalho abordamos os anéis de polinômios sobre anéis de divisão e como consequência da falta de comutatividade nos deparamos com os conceitos de *raiz a direita* e *pseudorraiz* de um polinômio.

2 Anéis de Divisão dos Quatérnios Reais

A palavra 'anel' será usada para um anel com elemento unidade 1. Um anel de divisão R é um anel no qual para todo elemento não nulo a de R existe b em R tal que $ab = ba = 1$.

Nesta seção apresentaremos o anel de divisão dos Quatérnios Reais. Mais detalhes podem ser encontrados em [1]. Sejam \mathbb{R} o corpo dos números reais e os símbolos $1, i, j, k$. Seja $\mathbb{H} = \{a_01 + a_1i + a_2j + a_3k \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$. Temos $a_01 + a_1i + a_2j + a_3k = b_01 + b_1i + b_2j + b_3k$ se, e somente se, $a_i = b_i$, para $i = 1, 2, 3, 4$.

A adição é dada em \mathbb{H} por:

$$(a_0 + a_1i + a_2j + a_3k) + (b_0 + b_1i + b_2j + b_3k) = (a_0 + b_0) + (a_1 + b_1)i + (a_2 + b_2)j + (a_3 + b_3)k.$$

Definimos a multiplicação dos quatérnios por:

$$\begin{aligned} &(a_0 + a_1i + a_2j + a_3k)(b_0 + b_1i + b_2j + b_3k) = \\ &(a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) + (a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2)i \\ &+ (a_0b_2 + a_2b_0 + a_3b_1 - a_1b_3)j + (a_0b_3 + a_3b_0 + a_1b_2 - a_2b_1)k. \end{aligned}$$

Esta fórmula do produto \mathbb{H} é obtida multiplicando as somas formais termo a termo, sujeitas as seguintes relações:

1. associatividade;
2. $ri = ir; rj = jr; rk = kr$ ($\forall r \in \mathbb{R}$);
3. $i^2 = j^2 = k^2 = -1$;
4. $ij = -ji = k; jk = -kj = i; ki = -ik = j$.

Com este produto, \mathbb{H} é um anel de divisão não comutativo no qual o inverso multiplicativo de $\alpha = a_0 + a_1i + a_2j + a_3k \neq 0$ é $\beta = \frac{a_0}{d} - \frac{a_1}{d}i - \frac{a_2}{d}j - \frac{a_3}{d}k$, onde $d = a_0^2 + a_1^2 + a_2^2 + a_3^2$.

3 Polinômios sobre Anéis Não Comutativos

Para qualquer anel R , $R[t]$ denota o anel dos polinômios em uma variável t sobre R , onde t é elemento central, isto é, t comuta com todos os elementos de R . Para um polinômio $f(t) = \sum_{i=0}^n a_i t^i \in R[t]$, e um elemento $r \in R$, definimos $f(r)$ como o elemento $\sum_{i=0}^n a_i r^i \in R$.

Notemos que embora $\sum_{i=0}^n a_i t^i = \sum_{i=0}^n t^i a_i \in R[t]$, os dois elementos $\sum_{i=0}^n a_i r^i$, $\sum_{i=0}^n r^i a_i$ sobre R podem ser diferentes.

Para avaliar $f(t)$ em um elemento $r \in R$, devemos primeiro expressar f na forma $\sum_{i=0}^n a_i t^i$, e então substituir t por r .

Exemplo 3.1. Seja $f_1(t) = (t - j)(t - i) \in \mathbb{H}[t]$. Desenvolvendo $f_1(t)$ temos que $f_1(t) = t^2 - (i + j)t + ji$. Assim, $f_1(i) = 0$ e $f_1(j) = -2ij$. Por outro lado, se escrevemos $f_1(t) = t^2 - t(i + j) + ji$, temos que $f_1(i) = -2ij$ e $f_1(j) = 0$.

Afirmamos que i é a única raiz de f_1 . Suponhamos que $f_1(z) = 0$. Daí, $f_1(z) = z^2 - (i + j)z + ji = 0$. Logo, $z^2 - (i + j)z = -ji$, assim $(z - i)z = j(z - i)$. Isto mostra que z é conjugado a j por $(z - i)$. Em particular $z^2 = j^2 = -1$. Temos então, $ji = (i + j)z - z^2 = (i + j)z + 1$. Resolvendo esta equação linear em z , obtemos que $z = i$. Logo, f_1 é um polinômio quadrático em $\mathbb{H}[t]$ que possui uma única raiz.

Exemplo 3.2. Analisaremos o seguinte polinômio de grau 3: $f(t) = (t - i)(t - j)(t - k)$. Temos,

$$\begin{aligned} f(t) &= (t^2 - (i + j)t + k)(t - k) \\ &= t^3 - kt^2 - (i + j)t^2 + (i + j)kt + kt - k^2 \\ &= t^3 - (i + j + k)t^2 + (-j + i + k)t + 1. \end{aligned}$$

Agora, iremos fazer $t = i, j, k$:

$$\begin{aligned} f(k) &= k^3 - (i + j + k)k^2 + (-j + i + k)k + 1 \\ &= -k + i + j - k - jk + ik + k^2 + 1 \\ &= +i + j - i - j - 1 + 1 = 0. \end{aligned}$$

$$\begin{aligned} f(i) &= i^3 - (i + j + k)i^2 + (-j + i + k)i + 1 \\ &= -i + i + j + k + (-ji) + i^2 + ki + 1 \\ &= j + 2k - 1 + j + 1 = 2j + 2k. \end{aligned}$$

$$\begin{aligned} f(j) &= j^3 - (i + j + k)j^2 + (-j + i + k)j + 1 \\ &= -j + i + j + k + 1 + ij + kj + 1 \\ &= i + k + 1 + k - i + 1 = 2 + 2k. \end{aligned}$$

Dessa forma, concluímos que k é raiz desse polinômio e i, j não são.

Definição 3.3. Um elemento $r \in R$ é dito uma raiz à direita de $f(t) \in R[t]$ se $f(r) = 0$.

Ao longo deste texto, o termo 'raiz' significa 'raiz à direita'.

Proposição 3.4 (Forma não-comutativa do Teorema do Resto). Um elemento $r \in R$ é uma raiz de um polinômio não nulo $f(t) \in R[t]$ se, e somente se, $(t - r)$ é um divisor à direita de $f(t)$ em $R[t]$. O conjunto de polinômios em $R[t]$ tendo r como uma raiz é o ideal à esquerda $R[t](t - r)$.

Demonstração. Se $f(t) = \sum c_i t^i (t - r) = \sum c_i t^{i+1} - \sum c_i r t^i$, então $f(r) = \sum c_i r^{i+1} - \sum c_i r r^i = 0$. Reciprocamente, assumamos $f(r) = 0$. Pelo algoritmo de Euclides, $f(t) = g(t)(t - r) + s$ para algum $g(t) \in R[t]$ e algum $s \in R$. A primeira parte mostra que r é uma raiz de $g(t)(t - r)$. Assim, $0 = f(r) = s$, isto é, $f(t) = g(t)(t - r)$. ■

Proposição 3.5. Seja D um anel de divisão e seja $f(t) = g(t)h(t) \in D[t]$. Seja $d \in D$ tal que $a := h(d) \neq 0$. Então $f(d) = g(ada^{-1})h(d)$. Em particular, se d é uma raiz de f mas não de h , então ada^{-1} é uma raiz de g .

Demonstração. Seja $g(t) = \sum b_i t^i$. Então $f(t) = \sum b_i h(t)t^i$ e $f(d) = \sum b_i h(d)d^i = \sum b_i ad^i a^{-1}a = \sum b_i (ada^{-1})^i h(d)$. A última conclusão segue do fato de D não ter divisores de zero. ■

Se D um anel de divisão, o centro de D é o conjunto $Z(D) = \{a \in D : ax = xa, \text{ para todo } x \in D\}$. Observe que o centro $Z(D)$ é um corpo.

Definição 3.6. A classe de conjugação A é algébrica sobre o centro de um anel de divisão, $F = Z(D)$, (e portanto todos) de seus elementos é algébrico sobre F . Neste caso, os elementos de A têm o mesmo polinômio minimal sobre F , que chamaremos polinômio minimal de A .

Teorema 3.7. Seja $f(t)$ o polinômio minimal de A . Um polinômio $h(t) \in D[t]$ se anula sobre A se, e somente se, $h(t) \in D[t]f(t)$.

Demonstração. Para qualquer $a \in A$, $f(a) = 0$ implica que $f(t) \in D[t](t - a)$, ou seja, $f(t) = g_1(t)(t - a)$, onde $g_1(t) \in D[t]$. Se $h(t) \in D[t]f(t)$, isto é, $h(t) = g_2(t)f(t)$, então também temos que $h(t) \in D[t](t - a)$ de forma que $h(A) = 0$.

Reciprocamente assumamos que $h(A) = 0$ onde $0 \neq h(t) \in D[t]$. Pelo Algoritmo de Euclides, podemos escrever $h(t) = g(t)f(t) + h_1(t)$, onde $h_1(t) = 0$ ou $\text{gr}h_1(t) < \text{gr}f(t)$. Mas $h(A) = 0$ implica que $h_1(A) = 0$, de modo que $h_1 = 0$, isto é, $h(t) = g(t)f(t)$, pois $f(t)$ é minimal. ■

A prova dos próximos três teoremas podem ser encontradas na Seção 1.2 da referência [2].

Teorema 3.8. Sejam a, b dois elementos em um anel de divisão D , ambos algébricos sobre $F = Z(D)$. Então a, b são conjugados em D se, e somente se, eles têm o mesmo polinômio minimal sobre F .

Teorema 3.9. *Seja D um anel de divisão com centro F , e seja A uma classe de conjugação de D que é algébrica sobre F com um polinômio minimal $f(t)$ e $F[t]$ de grau n . Então existem $a_1, \dots, a_n \in A$ tal que $f(t) = (t - a_n) \cdots (t - a_1) \in D[t]$. Também, $f(t)$ é o produto dos mesmos fatores lineares, permutados ciclicamente. O elemento $a_1 \in A$ aqui pode ser arbitrariamente escolhido.*

A fatoração do polinômio minimal $f(t)$ em $D[t]$ no Teorema 3.9 está muito longe de ser única, pois a_1 pode ser escolhido para ser qualquer elemento de A . É sabido que, se o elemento a_1 não é central, $f(t)$ tem infinitas fatorações distintas em fatores lineares em $D[t]$.

Teorema 3.10. *Sejam D um anel de divisão e c_1, \dots, c_n n elementos de D dois a dois não conjugados. Então existe um único polinômio $g(t) \in D[t]$, mônico de grau n , tal que $g(c_1) = \cdots = g(c_n) = 0$. Além disso, $g(t)$ tem as seguintes propriedades:*

- a) c_1, \dots, c_n são todas raízes de g em D .
- b) Se $h(t) \in D[t]$ se anula sobre todos os $c_i, 1 \leq i \leq n$, então $h(t) \in D[t]g(t)$.

Teorema 3.11. *Seja D um anel de divisão com centro F , e seja A uma classe de conjugação de D que tem polinômio minimal quadrático $\lambda(t) \in F[t]$. Se $f(t) \in D[t]$ tem duas raízes em A , então $f(t) \in D[t]\lambda(t)$ e $f(A) = 0$.*

Demonstração. Escreva $f(t) = q(t)\lambda(t) + (at + b)$, onde $q(t) \in D[t]$ e $a, b \in D$. Como $\lambda(A) = 0$ e $f(t) \in D[t]$ tem duas raízes em A , então $at + b$ tem duas raízes em A e, assim, $a = b = 0$. Logo, $f(t) = q(t)\lambda(t)$, que se anula identicamente sobre A . ■

Exemplo 3.12. *Seja $f_2(t) = (t - (j - i))(t - i) = t^2 - jt + (1 - k) \in \mathbb{H}[t]$. Com simples cálculos, vemos que $f_2(i) = f_2(i + j) = 0$. Também i e $i + j$ não são conjugadas em \mathbb{H} , pois $(i + j)^2 \neq 1$. O Teorema 3.10 implica que i e $i + j$ são as únicas raízes de f_2 em \mathbb{H} .*

4 Considerações finais

Ao final do trabalho esperamos que as definições tenham sido abordadas de forma coesa e dinâmica. Para um aprofundamento melhor desta pesquisa observem as referências. A continuação deste trabalho tem por finalidade um entendimento sobre quasedeterminantes e os teoremas de Bezout e Vieta sobre anéis não comutativos.

Agradecimentos

Eu, Julia Bernardes Coelho, sou grata por ser bolsista do Programa de Educação Tutorial PET-Matemática da Universidade Federal de Uberlândia e agradeço ao auxílio financeiro recebido durante o desenvolvimento deste trabalho.

Referências

- [1] HEIRTEIN, I. N. **Tópicos de Álgebra**. São Paulo: Edusp, Polígono, 1970.
- [2] PIRES, R. M. **Quase-determinantes e o Teorema de Vieta**. 2004. 69 f. Dissertação (Mestrado em Matemática) - Instituto de Matemática e Estatística, Universidade Federal de Goiás, Goiânia, 2004.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Operadores Ortogonais Orthogonal Operators

Laura Barbosa Goulart

UFU, FAMAT, Uberlândia, Minas Gerais, Brasil

laurag.net1@gmail.com

Adriana Rodrigues da Silva

UFU, FAMAT, Uberlândia, Minas Gerais, Brasil

adrianafamat@ufu.br

Palavras-chave

Operador ortogonal.
Transformação linear.
Ortogonalidade.
Álgebra linear.
.

Resumo

No seguinte trabalho serão abordados os conceitos de operadores lineares ortogonais e alguns exemplos. Operadores ortogonais preservam comprimentos e ângulos, que são propriedades necessárias para realizarmos um movimento rígido.

Keywords

Orthogonal operator.
Linear transformation.
Orthogonality.
Linear algebra.
.

Abstract

In the following text, the concepts of orthogonal linear operators and some examples will be discussed. Orthogonal operators preserve length and angles, which are necessary properties to make a rigid movement.

1 Introdução

Na Álgebra Linear, tem-se o conceito de Transformação Linear que nada mais é que um tipo especial de função cujo domínio e o contradomínio são espaços vetoriais onde as operações são mantidas, dessa forma tanto a variável independente, quanto a variável dependente são vetores.

Neste trabalho serão abordados os conceitos de operadores lineares ortogonais e suas propriedades. Além disso, são exemplificados alguns dos conceitos citados.

2 Conceitos gerais

Definição 2.1. *Um produto interno em um espaço vetorial real V é uma função de $V \times V$ em \mathbb{R} , em que todos os pares de vetores pertencentes em $V \times V$ são associados a um número real. O produto interno, de acordo com [1], é representado por $u \cdot v$ ou $\langle u, v \rangle$, deve satisfazer os seguintes axiomas, para todos $u, v \in V$:*

1. $u \cdot v = v \cdot u$;
2. $u \cdot (v + w) = u \cdot v + u \cdot w$;
3. $(\alpha u) \cdot v = \alpha \cdot (u \cdot v)$, $\forall \alpha \in \mathbb{R}$;
4. $u \cdot u \geq 0$ e $u \cdot u = 0$ se, e somente se, $u = 0$.

Observação: Sejam $u = (x_1, x_2, \dots, x_n)$, $v = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$. O produto interno canônico de \mathbb{R}^n é definido por:

$$u \cdot v = x_1 \cdot y_1 + x_2 \cdot y_2 + \dots + x_n \cdot y_n$$

Consequentemente, para obter o módulo de um vetor u basta fazer $\|u\| = \sqrt{u \cdot u}$.

Seja $B = \{v_1, v_2, \dots, v_n\}$ uma base de um espaço vetorial V , B é dita ortonormal se:

$$v_i \cdot v_j = \begin{cases} 0, & \text{se } i \neq j \\ 1, & \text{se } i = j. \end{cases}$$

Definição 2.2. *Sejam V e W espaços vetoriais reais. $T : V \rightarrow W$ é uma transformação linear de V em W , se para quaisquer $v, u \in V$ e $\alpha \in \mathbb{R}$, tem-se que:*

1. $T(v + u) = T(v) + T(u)$;
2. $T(\alpha u) = \alpha T(u)$.

Um operador linear é uma transformação linear de um espaço vetorial nele mesmo, ou seja, $V = W$.

Definição 2.3. *Seja V um espaço vetorial real com produto interno. Um operador linear ortogonal é um operador linear onde para todo $v \in V$ vale que:*

$$\|T(v)\| = \|v\|.$$

Observação: Qualquer produto interno, representado em base ortonormal se comporta como o produto interno canônico de \mathbb{R}^n .

De fato, seja B uma base ortonormal de V e sejam $u_B = (a_1, a_2, \dots, a_n)$ e $v_B = (b_1, b_2, \dots, b_n)$. Ao fazer o produto interno de u e v , tem-se:

$$u_B \cdot v_B = a_1 b_1 + \dots + a_n b_n$$

A partir de agora serão usadas somente bases ortonormais e a matriz de T em tal base será denotada por $[T]$.

Proposição 2.4. *Seja $T : V \rightarrow V$ um operador linear, tem-se que $[T]^{-1} = [T]^t$. Ou seja, a matriz do operador T é ortogonal.*

Demonstração. Observe que

$$\begin{aligned} \|v\| &= \|T(v)\| \implies \sqrt{v \cdot v} = \sqrt{T(v) \cdot T(v)} \\ \implies [v \cdot v] &= [T(v) \cdot T(v)] \implies [v]^t \cdot [v] = [T(v)]^t [T(v)]. \end{aligned}$$

Sabe-se que:

$$[T(v)]^t [T(v)] = ([T][v])^t [T][v] = [v]^t [T]^t [T][v].$$

Substituindo na equação anterior temos:

$$\begin{aligned} [v]^t [v] &= [v]^t [T]^t [T][v] \implies [T]^t [T] = I \\ \implies [T]^t &= [T]^{-1}. \end{aligned}$$

■

Proposição 2.5. *O determinante de uma matriz ortogonal é 1 ou -1.*

Demonstração. Se temos $[T]$ ortogonal então $[T]^t [T] = I$. Daí,

$$\det([T]^t [T]) = \det(I) \implies \det[T]^t \det[T] = 1.$$

Como

$$\det[T]^t = \det[T] \implies (\det[T])^2 = 1,$$

finalmente chegamos a

$$\det(T) = 1 \text{ ou } \det(T) = -1.$$

■

Corolário 2.6. *Todo operador linear ortogonal é inversível.*

Exemplo 2.7. *Seja A uma matriz ortogonal 2×2 . Então*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ e } \det A = ad - bc = \pm 1, \text{ pela proposição 2.5.}$$

Além disso,

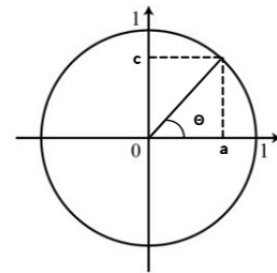
$$A^t = A^{-1} \Rightarrow \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \pm \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \text{ pela proposição 2.4}$$

Desta forma, temos que:

$$A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \text{ ou } A = \begin{pmatrix} a & c \\ c & -a \end{pmatrix}, \text{ ambas com } a^2 + c^2 = 1$$

• Colocando no círculo trigonométrico, temos que:

$$\begin{cases} a = \cos\theta \\ c = \operatorname{sen}\theta. \end{cases}$$



Portanto,

$$A = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} = R_\theta \text{ (Rotação do ângulo } \theta \text{)}$$

ou

$$A = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix} = R_\theta \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

onde $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ representa a reflexão em torno do eixo x

Portanto, temos que os operadores ortogonais em \mathbb{R}^2 são rotações, reflexões ou compostas destas.

Proposição 2.8. *Todo operador linear ortogonal $T : V \rightarrow V$ preserva o produto interno de vetores, isto é, para quaisquer $u, v \in V$, tem-se:*

$$u \cdot v = T(u) \cdot T(v).$$

Demonstração. Temos que

$$[T(u) \cdot T(v)] = [T(u)]^t [T(v)] = ([T][u])^t [T][v] = [u]^t [T]^t [T][v].$$

Como $[T]^t[T] = I$, tem-se:

$$[T(u) \cdot T(v)] = [u]^t[v] = [u \cdot v] \implies u \cdot v = T(u) \cdot T(v).$$

■

Proposição 2.9. *A composta de duas transformações ortogonais também é transformação ortogonal, ou, equivalentemente, o produto de duas matrizes ortogonais é uma matriz ortogonal também.*

Demonstração. Sejam T_1 e T_2 operadores ortogonais, dessa forma, temos:

$$[T_1 \circ T_2][T_1 \circ T_2]^t = ([T_1] \cdot [T_2]) \cdot ([T_1] \cdot [T_2])^t = [T_1] \cdot ([T_2] \cdot [T_2]^t) \cdot [T_1]^t = [T_1] \cdot [T_1]^t = I.$$

Portanto, $T_1 \circ T_2$ é ortogonal.

■

Proposição 2.10. *As colunas de uma matriz ortogonal são vetores ortonormais.*

Demonstração. Tomando [2] como referência, sejam $A = \{e_1, e_2, e_3, \dots, e_n\}$ base ortonormal de V , e $T : V \rightarrow V$ um operador ortogonal cuja matriz com relação a esta base é representada por:

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

Temos que:

$$\|e_1\| = \|e_2\| = \dots = \|e_n\| = 1, \quad e_i \cdot e_j = 0 \text{ se } i \neq j, \text{ e}$$

$$\begin{aligned} T(e_1) &= a_{11}e_1 + a_{21}e_2 + \dots + a_{n1}e_n \\ T(e_2) &= a_{12}e_1 + a_{22}e_2 + \dots + a_{n2}e_n \\ &\vdots \\ T(e_n) &= a_{1n}e_1 + a_{2n}e_2 + \dots + a_{nn}e_n \end{aligned}$$

Então

$$\begin{aligned} \|T(e_1)\|^2 &= T(e_1) \cdot T(e_1) = a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2 = 1 \\ \|T(e_2)\|^2 &= T(e_2) \cdot T(e_2) = a_{12}^2 + a_{22}^2 + \dots + a_{n2}^2 = 1 \\ &\vdots \\ \|T(e_n)\|^2 &= T(e_n) \cdot T(e_n) = a_{1n}^2 + a_{2n}^2 + \dots + a_{nn}^2 = 1 \end{aligned}$$

e, para $i \neq j$:

$$T(e_i) \cdot T(e_j) = a_{1i}a_{1j} + a_{2i}a_{2j} + \dots + a_{ni}a_{nj} = 0.$$

Dessa forma, as colunas:

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

são representações de vetores ortonormais de V .



Exemplo 2.11. *Seja a matriz*

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

Os vetores colunas são: $u_1 = (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$, $u_2 = (\frac{1}{2}, 0, \frac{1}{2})$ e $u_3 = (0, 1, 0)$. Então

$$\|u_1\| = \|u_2\| = \|u_3\| = 1.$$

Tem-se também que:

$$u_1 \cdot u_2 = u_1 \cdot u_3 = u_2 \cdot u_3 = 0.$$

Logo, o conjunto u_1, u_2, u_3 é base ortogonal do \mathbb{R}^3 .

3 Considerações finais

Neste trabalho foram estudados alguns conceitos em torno dos operadores lineares ortogonais. Operadores ortogonais preservam comprimentos e ângulos, que são propriedades necessárias para realizarmos um movimento rígido.

Agradecimentos

Primeiramente gostaria de agradecer ao PET, que possibilitou essa Iniciação científica de acontecer e à portaria SESu-MEC pelo auxílio financeiro, pois sem isso não conseguiria todo o amparo necessário para realizar este trabalho.

Referências

- [1] STEINBRUCH, A.; WINTERLE, P. **Álgebra Linear**. São Paulo: McGraw-Hill, 1987.
- [2] BOLDRINI, J. L.; COSTA, S. R. I.; FIGUEIREDO, V. L. Wetzler. **Álgebra Linear**. São Paulo: Harbra, 1984.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Dinâmica Unidimensional Unidimensional Dynamics

Leonardo Henrique Soria

UFU, FAMAT, Uberlândia, MG, Brasil

leonardohsoria@gmail.com

Marcus Augusto Bronzi

UFU, FAMAT, Uberlândia, MG, Brasil

mbronzi@ufu.br

Palavras-chave

Sistemas Dinâmicos.
Aplicação shift.
Família quadrática.
Conjugação topológica.
Rotação em S^1 .

Resumo

O trabalho tem como objetivo ilustrar por meio de exemplos alguns dos objetos estudados pela Teoria de Sistemas Dinâmicos. Serão abordados aspectos da teoria unidimensional, por meio de aplicações em S^1 e na reta real. Além disso traremos alguns elementos da dinâmica simbólica e como ela pode ser útil para descrever dinâmicas mais difíceis.

Keywords

Dynamical Systems.
Shift application.
Quadratic family.
Topological conjugation.
Rotation of S^1 .

Abstract

The objective of this work is to illustrate by means of examples some of the objects studied by the theory of dynamical systems. Aspects of unidimensional theory will be addressed, through applications in S^1 and in the real line. In addition we will bring some elements of symbolic dynamics and how it can be useful to describe more difficult dynamics.

1 Introdução

A expressão “Sistemas Dinâmicos” faz referência a sistemas que se modificam a partir de uma regra ligando os estados atuais, passados e futuros. Pode-se pensar, por exemplo, na evolução de um fenômeno biológico em cada instante de tempo $t = 0, 1, 2, \dots$, no comportamento de um pêndulo e muitos outros.

Na dinâmica unidimensional é possível identificar e estudar os principais fenômenos e patologias comuns em sistemas de dimensões maiores. Por isso, o estudo de aplicações no círculo ou na reta proporciona o entendimento das ferramentas que originaram o conceito de hiperbolicidade, estabilidade, conjuntos atratores, conjuntos estáveis e instáveis, perturbações dentre outros.

Nesse sentido, trataremos alguns aspectos da Teoria por meio de alguns exemplos em espaços como S^1 e \mathbb{R} . Inicialmente,

Definição 1.1. *Seja $f : X \rightarrow X$ uma função. Um ponto x é fixo para f se $f(x) = x$. Dizemos que x é periódico se para algum n , $f^n(x) = x$. Nesse caso, definimos o período de x como sendo o menor inteiro n tal que $f^n(x) = x$. Além disso, vamos denotar o conjunto dos pontos fixos de f por $Fix(f)$ e por $Per_n(f)$ o conjunto dos pontos de período até n . Definimos também a órbita positiva de x como sendo o conjunto dos pontos $x, f(x), f^2(x) \dots$ e denotamos por $O^+(x)$. Se f é um homeomorfismo podemos definir a órbita completa de x , $O(x)$, como o conjunto de pontos $f^n(x)$ para $n \in \mathbb{Z}$.*

2 Dinâmica de S^1

Exemplo 2.1 (Rotação em S^1). *Seja $\lambda \in \mathbb{R}$ e $f_\lambda : S^1 \rightarrow S^1$, tal que $f_\lambda(x) = x + 2\pi\lambda$. As aplicações f_λ se comportam de maneira bastante diferentes dependendo da racionalidade ou irracionalidade de λ . Se $\lambda = p/q$, com p, q inteiros, então $f_\lambda^q(x) = x + 2\pi p = x$ e portanto f_λ fixa todos os pontos na q -ésima iteração. Quando λ é irracional a situação é um tanto quanto diferente.*

O resultado a seguir é conhecido como Teorema de Jacob.

Teorema 2.2. *Toda órbita de f_λ é densa em S^1 se λ é irracional.*

Demonstração. A demonstração deste resultado pode ser vista em [1]. ■

Para auxiliar a compreensão da dinâmica no círculo vamos recorrer ao levantamento de S^1 para a reta. Mas antes, defina $\pi : \mathbb{R} \rightarrow S^1$ tal que

$$\pi(x) = \exp(2\pi ix) = \cos(2\pi x) + i \sin(2\pi x). \quad (1)$$

π desempenha a função de "enrolar" a reta ao redor de S^1 . Em [2] é possível verificar a existência

de um homeomorfismo $F : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$\pi \circ F = f \circ \pi, \quad (2)$$

para f aplicação em S^1 . E então chegamos à seguinte definição:

Definição 2.3. *Seja $F : \mathbb{R} \rightarrow \mathbb{R}$ um homeomorfismo crescente. F é chamada de levantamento de $f : S^1 \rightarrow S^1$ se satisfaz*

$$\pi \circ F = f \circ \pi. \quad (3)$$

Exemplo 2.4. *Seja f_λ como no Exemplo 2.1. O levantamento de f_λ é a função real*

$$F(x) = x + \lambda + k. \quad (4)$$

Existem muitas propriedades intrínsecas a levantamentos de uma aplicação em S^1 e algumas delas é resultado das seguintes proposições:

Proposição 2.5. *Quaisquer dois levantamentos de f diferem por um inteiro.*

Demonstração. A demonstração deste resultado pode ser vista em [2]. ■

Proposição 2.6. *Seja F um levantamento de f . Então $F(x + 1) = x + 1$.*

Demonstração. A demonstração deste resultado pode ser vista em [2]. ■

2.1 Número de rotação

Um dos principais invariantes associado a aplicações no círculo é o número de rotação. Essencialmente, esse número é um valor entre 0 e 1 que determina a média de pontos que são girados por uma iteração de uma aplicação. Antes de definir formalmente esse número vamos introduzir um conceito preliminar.

Definição 2.7. *Seja $f : S^1 \rightarrow S^1$ uma translação que preserva orientação e escolha qualquer levantamento F de f . Defina*

$$\rho_0(F) = \lim_{n \rightarrow \infty} \frac{F^n(x)}{n}. \quad (5)$$

Em [1] é possível verificar que ρ_0 existe e independe do valor de x escolhido.

Exemplo 2.8. *Seja $f_\lambda(\theta) = \theta + 2\pi\lambda$ como no Exemplo 2.1. Considere o levantamento $F_\lambda(x) = x + \lambda + k$. Temos*

$$\rho_0(F_\lambda) = \lim_{n \rightarrow \infty} \frac{x + n\lambda + nk}{n} = \lambda + k. \quad (6)$$

Isto é, diferentes levantamentos para f_λ produzem diferentes ρ_0 que diferem por um inteiro. Esse exemplo abre caminho para o fato mais geral. Sejam F_1 e F_2 levantamentos para f . Como vimos anteriormente, existe c inteiro tal que $F_2(x) = F_1(x) + c$. Da Proposição 2.6 segue facilmente que $F_2^n(x) = F_1^n(x) + nc$. E então $\rho_0(F_2) = \rho_0(F_1) + c$.

Finalmente,

Definição 2.9. O número de rotação de f , $\rho(f)$, é a parte fracionária de $\rho_0(F)$ para qualquer levantamento F de f . Isto é, $\rho(f)$ é o único número em $[0, 1)$ tal que $\rho_0(F) - \rho(f)$ é um inteiro. Note que, uma vez que $\rho_0(F)$ existe, o número de rotação $\rho(f)$ está bem definido.

Por fim, o resultado que permite visualizar o papel do número de rotação para aplicações em S^1 :

Teorema 2.10. $\rho(f)$ é irracional se, e somente se, f não possui ponto periódico.

Demonstração. A demonstração deste resultado pode ser vista em [1]. ■

3 Família quadrática

Nesta seção iremos focar o olhar sobre a família de funções quadráticas $F_\mu(x) = \mu x(1 - x)$. Em especial, trabalharemos nela com o parâmetro $\mu = 3,839$. O objetivo é visualizar como a dinâmica simbólica pode auxiliar na compreensão da dinâmica dessa função. Inicialmente,

Teorema 3.1. Seja $f : \mathbb{R} \rightarrow \mathbb{R}$ contínua. Suponha que f possua um ponto de período três. Então f possui pontos periódicos de todos os períodos.

Definição 3.2. Considere o conjunto

$$\Sigma_N = \{s = (s_0 s_1 s_2 \dots) \mid s_j \in \mathbb{N}, 1 \leq s_j \leq N\}$$

de todas as possíveis seqüências s de números naturais entre 1 e N . Σ_N é chamado de Espaço de Sequência de N símbolos. Elementos de Σ_N são seqüências infinitas como: $(121211\dots)$ e $(111\dots N\dots)$.

Podemos transformar Σ_N em um espaço métrico da seguinte maneira: para duas seqüências $s = (s_0 s_1 s_2 \dots)$ e $t = (t_0 t_1 t_2 \dots)$ definimos a distância entre elas por:

$$d[s, t] = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{N^i}.$$

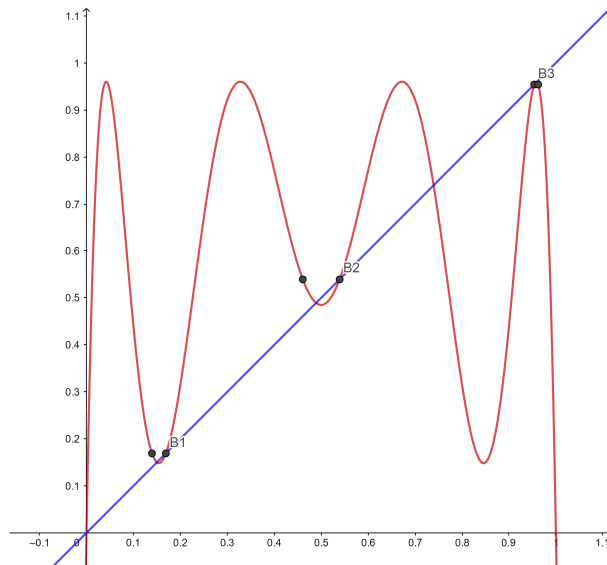
Definição 3.3. A função shift $\sigma : \Sigma_N \rightarrow \Sigma_N$ é definida por $\sigma(s_0 s_1 s_2 \dots) = (s_1 s_2 s_3 \dots)$.

O shift "esquece" a primeira entrada de uma seqüência e desloca todas as próximas para um lugar à esquerda. E mais, na métrica anterior o shift é contínuo.

Proposição 3.4. $\sigma : \Sigma_N \rightarrow \Sigma_N$ é contínuo.

Demonstração. A demonstração pode ser vista em [1]. ■

Retomando a família quadrática, analisando graficamente temos:

Figura 1: Gráfico de F^3 .

Note que cada ponto periódico e o respectivo ponto de mesma imagem definem um intervalo aberto, nos quais denotaremos por A_1 (com um dos extremos sendo B_1), A_2 (com um dos extremos sendo B_2) e A_3 (com um dos extremos sendo B_3). Os intervalos adjacentes chamaremos de I_0 , I_1 , I_2 e I_3 .

Proposição 3.5. *Todos os pontos periódicos de F pertencem à $I_1 \cup I_2$.*

Demonstração. A demonstração pode ser vista em [1]. ■

Vamos denotar por Λ o conjunto dos pontos no qual a órbita está inteiramente contida nesses dois intervalos. Observe ainda que este conjunto é invariante por F . Para compreender a dinâmica de F em Λ vamos invocar a dinâmica simbólica. Definiremos uma sequência associada à x pela regra

$$S(x) = (s_0 s_1 s_2 \dots)$$

tal que $s_j = 1$ se $F^j(x) \in I_1$ e $s_j = 2$ se $F^j(x) \in I_2$.

Definição 3.6. *Seja $f : A \rightarrow A$ e $g : B \rightarrow B$ funções. f e g são ditas topologicamente conjugadas se existe um homeomorfismo $h : A \rightarrow B$ tal que $h \circ f = g \circ h$. O homeomorfismo h é chamado de conjugação topológica.*

As funções topologicamente conjugadas são completamente equivalentes em termos de sua dinâmica. Por exemplo, se f é topologicamente conjugado a g via h e p é um ponto fixo de f , então $h(p)$ é fixo em g . De fato, $h(p) = h(f(p)) = g(h(p))$.

Teorema 3.7. *A restrição de F a Λ é topologicamente conjugada ao subshift do tipo finito σ_A em Σ_A .*

Demonstração. A demonstração pode ser vista em [1]. ■

4 Conclusões Finais

Alguns aspectos da dinâmica de S^1 puderam ser compreendidos a partir de funções em \mathbb{R} , que por vezes são mais fáceis de se trabalhar. Analogamente, estabelecido a correspondência entre família quadrática e o shift é possível então verificar a existência de pontos periódicos de todos os períodos para F por meio de sequências que repetem "blocos" de comprimento equivalente ao período que se deseja.

Referências

- [1] DEVANEY, Robert L. *An Introduction to Chaotic Dynamical Systems*. 2. ed. Redwood City: Addison-Wesley, 1989.
- [2] GELFERT, Katrin. Disponível em: <<http://www.im.ufrj.br/gelfert/cursos/2017-1-SisDin/2017-1-SisDin.pdf>>. Acesso em: 13 de setembro de 2020.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O Sistema de Lorenz: Uma Alusão ao Caos em Dinâmica Contínua The Lorenz System: An Allusion to Chaos in Continuous Dynamics

Matheus Gonçalves Silva de Castro

Universidade Federal de Uberlândia, Instituto de Física, Uberlândia, Minas Gerais, Brasil

matheusg.s.castro@ufu.br

Jean Venato Santos

Universidade Federal de Uberlândia, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

jvenatos@ufu.br

Resumo

Palavras-chave

Sistema de Lorenz.
Análise Não Linear.
Atratores Estranhos.
Fluxos Caóticos.
Equações Diferenciais.

O principal objetivo desse trabalho é apresentar o sistema de Lorenz e, para tanto, discorre-se acerca de relevantes fatos históricos que permitirão compreender o contexto científico, físico e matemático no qual teve início o estudo desse sistema. Antes de, efetivamente, estudar-se o sistema de Lorenz, apresentam-se algumas ferramentas de análise não linear essenciais a tal investigação; assim, enuncia-se o teorema de Hartman-Grobman, alguns critérios de estabilidade e define-se as funções de Lyapunov. Por fim, a mais importante característica do sistema de Lorenz, sua sensibilidade às condições iniciais (que traduz o fenômeno do caos), é formalmente enunciada e a trajetória que culminou nesse enunciado formal é sucintamente apresentada.

Abstract

Keywords

Lorenz System.
Nonlinear Analysis.
Strange Attractors.
Chaotic Flows.
Differential Equations.

The main objective of this work is to present the Lorenz system and, for that, it is discussed about relevant historical facts that will allow to understand the scientific, physical and mathematical context in which the study of this system began. Before actually studying the Lorenz system, some nonlinear analysis tools essential to this investigation are presented; thus, the Hartman-Grobman theorem, some stability criteria are enunciated and Lyapunov's functions are defined. Finally, the most important characteristic of the Lorenz system, the sensitivity to the initial conditions (which translates the chaos phenomena), is formally stated and the trajectory that culminated in this formal statement is succinctly presented.

1 Introdução

O determinismo é um princípio físico segundo o qual causas idênticas só podem originar efeitos idênticos; essa concepção imperou nas ciências naturais por muito tempo e só foi questionada com o advento da mecânica quântica. Equações diferenciais ordinárias, usadas amplamente como modelos para fenômenos físicos, são sistemas determinísticos, conforme estabelece o teorema de existência e de unicidade de soluções para equações diferenciais de classe C^1 . Todavia, o conceito de determinismo não implica, por exemplo, que causas similares originem fenômenos também semelhantes. De fato, sistemas deterministas podem apresentar comportamento oposto ao expresso nessa última sentença.

Em 1963, Edward Lorenz, que era meteorologista, ao estudar uma simplificação de um modelo matemático para processos convectivos que ocorrem na atmosfera terrestre (hoje conhecido como sistema de Lorenz), deparou-se com um interessantíssimo fenômeno. Lorenz notou que, ao modificar-se ligeiramente os parâmetros atmosféricos relevantes ao sistema que então investigava, as soluções obtidas divergiam rapidamente ao longo do tempo. Isso revela o desafio enfrentado por cientistas na tentativa de melhorar as previsões climáticas. Essa peculiaridade foi denominada sensibilidade às condições iniciais e constitui a essência do conceito caos.

2 Conceitos Prévios

Antes de iniciar a investigação do sistema de Lorenz, é imprescindível apresentar alguns importantes resultados de análise não linear, os quais permitem investigar muitos sistemas dinâmicos contínuos, os quais foram extraídos de [1], [2].

Definição 2.1 (Fluxos). *Um sistema dinâmico contínuo (ou fluxo) de classe C^k em \mathbb{R}^n é uma aplicação $\phi: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ de classe C^k tal que $\phi(t, X) = \phi_t(X)$, com $X \in \mathbb{R}^n, t \in \mathbb{R}$ (sendo, fisicamente, a variável t o tempo), satisfazendo as seguintes condições:*

1. $\phi_0: \mathbb{R}^n \rightarrow \mathbb{R}^n$ é a identidade, ou seja, $\phi_0(X) = X$;
2. dadas ϕ_a e ϕ_b , com $a, b \in \mathbb{R}$, vale que $\phi_a \circ \phi_b = \phi_{a+b}$.

Agora, é importante comentar que muitos sistemas de equações diferenciais ordinárias autônomos, de primeira ordem, pertencente à classe de diferenciabilidade C^1 e com “n” componentes encontram-se vinculados a um fluxo em \mathbb{R}^n . De fato, dada $X' = F(X)$ e uma condição inicial qualquer $X_0 \in \mathbb{R}^n$, a aplicação $\phi(t, X_0): \mathbb{R} \rightarrow \mathbb{R}^n$ verificando $\phi(0, X_0) = X_0$ e $\phi'(t, X_0) = F(\phi(t, X_0))$, para todo $t \in \mathbb{R}$, é um fluxo. Contudo, as soluções de um sistema de equações diferenciais ordinárias não estão necessariamente, definidas para qualquer instante de tempo e, por isso, nem sempre definem um fluxo.

Definição 2.2 (Equilíbrios Estável, Atratores Local e Global, Assintoticamente Estável). *Um equilíbrio X_0 é denominado: estável se, para todo $\epsilon > 0$, existir $\delta(\epsilon)$ tal que: $\|X - X_0\| < \delta(\epsilon) \Rightarrow \|\phi(t, X) - X_0\| < \epsilon$, para todo $t > 0$; atrator local se, para algum $\epsilon > 0$, $\|X - X_0\| < \epsilon \Rightarrow$*

$\phi(t, X) \rightarrow X_0$, quando $t \rightarrow \infty$; assintoticamente estável se for estável e se for atrator local; atrator global se, para todo $X \in \mathbb{R}^n$, $\phi(t, X) \rightarrow X_0$, quando $t \rightarrow \infty$.

Definição 2.3 (Funções Definida Positiva e de Lyapunov). *Dado $A \subset \mathbb{R}^n$ e sendo $V: \mathbb{R}^n \rightarrow \mathbb{R}$ de classe C^1 , essa função é dita ser definida positiva em $\mathbb{R}^n - A$ se $V(X) > 0$ para todo $X \in \mathbb{R}^n - A$ e se $V(X) \leq 0$ para todo $X \in A$. Ainda, V é denominada função de Lyapunov estrita vinculada ao equilíbrio X_0 do sistema $X' = F(X)$, se $\langle \nabla V(X), F(X) \rangle < 0$, para todo $X \in \mathbb{R}^n - X_0$ e é designada função de Lyapunov generalizada vinculada ao conjunto A e associada ao sistema $X' = F(X)$, se $\langle \nabla V(X), F(X) \rangle < c < 0$, para todo $X \in \mathbb{R}^n - A$ e para algum $c \in \mathbb{R}$.*

Teorema 2.4 (Critério de Estabilidade de Lyapunov [1]). *Seja $V: \mathbb{R}^n \rightarrow \mathbb{R}$ função de Lyapunov estrita vinculada ao equilíbrio X_0 do sistema $X' = F(X)$. Nesse caso X_0 é assintoticamente estável e X_0 é também um atrator global. Similarmente, seja $Y: \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de Lyapunov generalizada vinculada ao conjunto A e associada ao sistema $X' = F(X)$. Seja, ainda, ϕ um fluxo associado a esse mesmo sistema. Então, $\phi_t(A) \subset A$ para todo $t > 0$ e existe uma função $\tau: \mathbb{R}^n \rightarrow \mathbb{R}_+$ tal que $\phi_t(X) \in A$, para todo $t \geq \tau(X)$.*

Definição 2.5 (Matriz e Equilíbrio Hiperbólicos). *Uma matriz $A \in \mathbb{R}^{n \times n}$ é dita hiperbólica se a parte real de cada autovalor seu é não nula e é estável se todos os seus autovalores apresentam parte real negativa. O equilíbrio X_0 do sistema $X' = F(X)$ é chamado hiperbólico se $JF(X_0)$ é hiperbólica.*

Definição 2.6 (Conjugação Topológica Local). *Considere-se $\phi: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ e $\bar{\phi}: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ dois fluxos e suponha-se que exista $h: V \subset \mathbb{R}^n \rightarrow U \subset \mathbb{R}^n$ (com U e V abertos) um homeomorfismo tal que $(\phi_t \circ h)(X) = (h \circ \bar{\phi}_t)(X)$ sempre que $X, \phi_t(X) \in V$. Então, h é denominada uma conjugação topológica local e as aplicações $\phi_t|_U, \bar{\phi}_t|_V$ são ditas localmente conjugadas por h .*

Teorema 2.7 (Hartman–Grobman para Fluxos [1]). *Seja X_0 um equilíbrio hiperbólico do fluxo $\phi: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ associado ao sistema $X' = F(X)$ (onde F é de classe C^1) e seja $\bar{\phi}: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ o fluxo vinculado ao sistema linear $X' = JF(X_0)X$. Nessas condições, existem abertos $U, V \subset \mathbb{R}^n$ contendo, respectivamente, X_0 e $\vec{0}$ e existe um homeomorfismo $h: U \rightarrow V$ o qual é conjugação topológica local entre $\phi_t|_U$ e $\bar{\phi}_t|_V$.*

O corolário abaixo é consequência imediata da aplicação do Teorema de Hartman–Grobman para Fluxos; a verificação desse fato será deixada a cargo do leitor.

Corolário 2.8 (Estabilidade de Equilíbrios). *Se X_0 é um equilíbrio hiperbólico do fluxo suave associado ao sistema de equações $X' = F(X)$, então: $JF(X_0)$ é uma matriz estável se, e somente se, X_0 é um ponto fixo assintoticamente estável; não sendo $JF(X_0)$ uma matriz estável, X_0 é equilíbrio instável.*

3 O Sistema de Lorenz

O sistema de equações

$$\begin{cases} x'(t) = \sigma(y(t) - x(t)) \\ y'(t) = rx(t) - y(t) - x(t)z(t), \\ z'(t) = x(t)y(t) - bz(t) \end{cases} \quad (1)$$

onde σ , r e b são constantes positivas e onde $\sigma > b + 1$, define um fluxo. Esse fluxo é o sistema de Lorenz e as equações diferenciais acima serão denotadas, mais simplesmente, por $X' = \mathcal{L}(X)$. Os equilíbrios desse sistema (as soluções de $\mathcal{L}(X) = \vec{0}$), para $r > 1$, são: $\vec{0}$ e $Q_{\pm} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$. Se não for imposto que $r > 1$, a origem será o único ponto do sistema de Lorenz. Quando $0 \leq r < 1$, a matriz Jacobina do campo vetorial \mathcal{L} em $\vec{0}$ apresentará autovalores $-b$ e $\lambda_{\pm} = \left(-(\sigma+1) \pm \sqrt{(\sigma+1)^2 - 4\sigma(1-r)} \right) / 2$, os quais são todos negativos. Assim, pelo Teorema 2.8, a origem é um equilíbrio assintoticamente estável se $0 \leq r < 1$. Para ver que, além disso, a origem, nessas mesmas condições, é um atrator global, repare-se que $L(x, y, z) = x^2 + \sigma(y^2 + z^2)$ é uma função de Lyapunov estrita vinculada a $\vec{0}$. De fato, $L \in C^{\infty}$, $L(X) > 0$ e $\langle \nabla L(X), \mathcal{L}(X) \rangle = -2\sigma(x^2 + y^2 - (1+r)xy) - 2\sigma bz^2 < 0$, para todo $X \neq \vec{0}$.

Quando $r = 1$, a origem é um equilíbrio não hiperbólico; para $r > 1$ a origem é um equilíbrio hiperbólico instável e existem também os equilíbrios Q_{\pm} . Em particular, se $1 < r < r^* = \sigma(\sigma + b + 3)/(\sigma - b - 1)$, os equilíbrios Q_{\pm} são estáveis, pois os autovalores de $JF(Q_{\pm})$ são todos negativos. Para $r = r^*$, $JF(Q_{\pm})$ possui autovalores com parte real nula, de modo que os pontos fixos Q_{\pm} são não hiperbólicos; sendo que a determinação explícita de r^* pode ser feita ao impor-se $i\omega$ (com $\omega \in \mathbb{R}$) como raiz do polinômio característico de $JF(Q_+)$ ou de $JF(Q_-)$ e ao determinar-se a solução da equação resultante na incógnita r . Finalmente, para $r > r^*$, esses mesmos pontos são equilíbrios hiperbólicos instáveis. Tudo isso é percebido a partir da aplicação do Teorema 2.8.

Agora, será visto que existe um compacto A ao qual está vinculada a função de Lyapunov generalizada associada ao sistema $X' = \mathcal{L}(X)$ dada por $V(x, y, z) = rx^2 + \sigma y^2 + \sigma(z - 2r)^2 - \mu$, onde $\mu > br^2$ e $A = \{X \in \mathbb{R}^3 | V(X) \leq 0\}$. Realmente, $\langle \nabla V(X), \mathcal{L}(X) \rangle = -2\sigma(rx^2 + y^2 + b(z - 2r)^2 - br^2) < 0$ para todo $X \in A$ e, portanto, existe uma função $\tau: \mathbb{R}^3 \rightarrow \mathbb{R}_+$ tal que $\phi_t(X) \in A$, para todo $t \geq \tau(X)$.

Denote-se, então, por Λ , o conjunto dos pontos $X \in \mathbb{R}^3$ tais que $\phi_t(X) \in A$ para todo $t \in \mathbb{R}$. Claramente, esse conjunto é não vazio, uma vez que contém, ao menos, os três equilíbrios do sistema de Lorenz. Além disso, ele é invariante (ou seja, $\phi_t(\Lambda) = \Lambda$ para todo $t \in \mathbb{R}$), pois é união de órbitas as quais são conjuntos invariantes. Sobre Λ , também sabe-se que seu volume é zero. A fim de formalizar isso, será necessário definir o conceito de volume de um subconjunto de \mathbb{R}^3 , bem como apresentar um teorema o qual diz como evolui, por um fluxo, o volume de um conjunto, ao longo do tempo.

Na continuação do estudo do sistema de Lorenz, será mostrado que esse é dissipativo. Isso significa que se $B \subset \mathbb{R}^3$ e se $v(B)$ (o volume de $B \subset \mathbb{R}^3$, dado pela integral, em todo o \mathbb{R}^3 , da função f , onde $f(x) = 1$ se $x \in B$ e $f(x) = 0$ se $x \notin B$), é um número, então $v(\phi_t(B)) < v(B)$, para todo $t \in \mathbb{R}_+^*$.

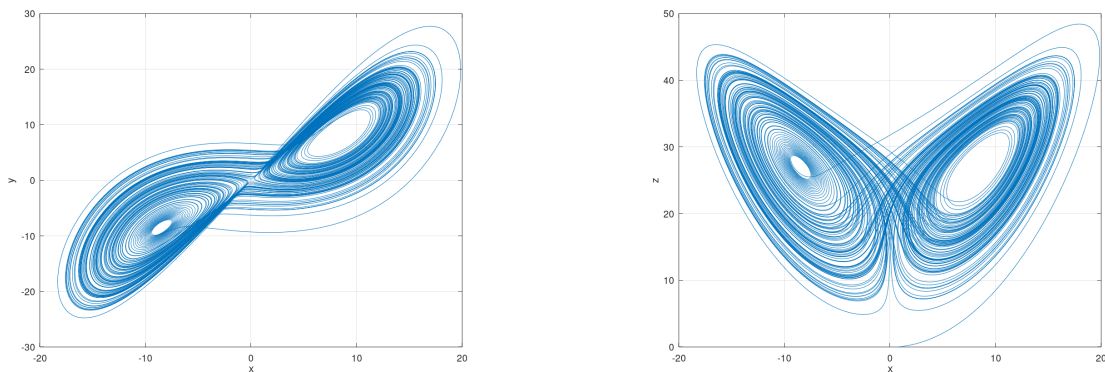


Figura 1: Projeções, respectivamente, nos planos xy e xz , da imagem do intervalo $[0; 150]$ pela solução do sistema de Lorenz (com parâmetros $\sigma = 10$, $r = 28$ e $b = 8/3$) com condição inicial $(1, 0, 0)$.

Para ver isso, será preciso utilizar o Teorema de Liouville (que é apresentado em [1]), segundo o qual,

$$\frac{dv}{dt}(\phi_t(B)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\phi_t(B)) \nabla \cdot \mathcal{L}(x, y, z) \, dx dy dz.$$

Calculando-se o divergente de \mathcal{L} , encontra-se o valor constante $-(\sigma+b+1)$ e, portanto, $v(\phi_t(B)) = e^{-(\sigma+b+1)t} v(\phi_0(B))$. Com isso, em decorrência da invariância de Λ , $v(\Lambda) = 0$. Um outro importante fato acerca de Λ (porém, de difícil prova) é que o conjunto Λ é um atrator do sistema de Lorenz. Para melhor estabelecer esse fato, apresenta-se, a seguir, a definição de conjunto atrator.

Definição 3.1 (Conjunto Atrator). *Um conjunto $\Lambda \subset \mathbb{R}^n$ é denominado um atrator do fluxo ϕ associado ao sistema $X' = F(X)$ se: Λ é compacto e invariante; existe $U \subset \mathbb{R}^n$ satisfazendo $\phi_{\mathbb{R}_+}(U) \subset U$ o qual é um aberto tal que $\Lambda \subset U$ e tal que $\bigcap_{t \geq 0} \phi_t(U) = \Lambda$; (transitividade topológica) se são dados $X_1, X_2 \in \Lambda$ e $U_1, U_2 \subset \mathbb{R}^n$ abertos quaisquer satisfazendo $X_i \in U_i$ para $i = 1, 2$, então existem $t_1, t_2 \in \mathbb{R}$ e uma solução $\varphi(t)$ de $X' = F(X)$ de modo que $\varphi(t_j) \in U_j$ para $j = 1, 2$.*

Falta agora apresentar, de maneira formal, a mais marcante característica do conjunto atrator Λ (isto é, a sensibilidade às condições iniciais constatadas por Lorenz). Antes, porém, será descrita, sucintamente e em perspectiva histórica, a sequência de avanços que culminou na comprovação dessa propriedade para o sistema (1) com os valores $\sigma = 10$, $r = 28$ e $b = 8/3$. Todos os fatos que serão listados a seguir foram extraídos de [3].

A fim de se compreender o contexto em que teve início a investigação do sistema de Lorenz, é crucial mencionar a introdução, por Stephen Smale, do conceito de Sistema Dinâmico Hiperbólico - designação atribuída aos sistemas que apresentam hiperbolicidade. Este conceito discrimina uma importante categoria de sistemas dinâmicos e merece menção por conta da importância que adquiriu. Todavia, o sistema de Lorenz não apresenta hiperbolicidade, mas possui robustez (consulte [3] para ver uma descrição dessas propriedades), sendo por isso também relevante.

Na sequência, é importante pontuar um importante avanço na investigação especificamente do sistema de Lorenz feito, independentemente, por dois grupos de colaboradores: de um lado, V. Afraimovich, V. Bykov, L. Shil'nikov e, por outro lado, J. Guckenheimer, R. Williams. Esses pesquisadores, inspirados no comportamento do sistema de Lorenz, introduziram o modelo geométrico de Lorenz, estabelecendo propriedades as quais, se satisfeitas por um fluxo, implicariam a existência de um atrator estranho (conceito definido abaixo). Um curioso fato é que o modelo geométrico de Lorenz preserva a similaridade entre o atrator do sistema de Lorenz e uma borboleta; semelhança evidenciada na projeção no plano xy vista na Figura 1.

Definição 3.2 (Atrator Estranho). *Sejam ϕ o fluxo associado a $X' = F(X)$, Λ um conjunto atrator de ϕ . Considere-se que, além disso, Λ é sensível às condições iniciais, isto é: que existe $\beta > 0$ tal que, para todos $X \in \Lambda$ e $\varepsilon > 0$, existem $Y \in \Lambda$ e $t \in \mathbb{R}$ satisfazendo $d[X, Y] < \varepsilon$, de modo que $d[\phi_t(X), \phi_t(Y)] > \beta$. Nesse caso, Λ é dito um atrator estranho e ϕ é um sistema tipo Lorenz.*

4 Considerações finais

Cabe enfatizar aqui que, apesar do modelo geométrico de Lorenz ter sido criado com inspiração no sistema de Lorenz, o sistema definido por (1) não foi provado ser do tipo Lorenz pelos autores que introduziram esse modelo. C. Morales, M. J. Pacífico e E. Pujals realizaram o próximo passo de grande relevância na investigação do sistema de Lorenz, ao mostrarem que todo sistema com um atrator robusto contendo algum ponto de equilíbrio é tipo Lorenz. Na sequência, W. Tucker, partindo desses significativos avanços, demonstrou que o sistema de Lorenz com os parâmetros $\sigma = 10$, $r = 28$ e $b = 8/3$ apresenta um atrator estranho. Finalmente, é com base na sensibilidade às condições iniciais verificada por atratores estranhos que o sistema de Lorenz é chamado caótico.

Agradecimentos

Agradecimentos ao CNPq pelo fomento financeiro via programa PICME.

Referências

- [1] M. W. Hirsch, S. Smale e R. L. Devaney, **Differential Equations, Dynamical Systems, and an Introduction to Chaos**, Elsevier Academic Press, 2ed, 2004.
- [2] M. Vidyasagar, **Nonlinear Systems Analysis**, Prentice Hall, 2ed, 19
- [3] M. Viana, **What's new on Lorenz strange attractors?**, Mathematical Intelligencer, v. 22(3), 6-19.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Construindo a distância de Hausdorff Building the Hausdorff distance

Paulo César da Silva Andreucci

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
paulo.andreucci@hotmail.com

Ana Paula Tremura Galves

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
ana.galves@ufu.br

Palavras-chave

Distância de Hausdorff.
Espaço de Hausdorff.
Espaço métrico.
Compacto.

Resumo

Neste trabalho vamos apresentar o Espaço de Hausdorff, e teremos como objetivo defini-lo como um espaço métrico. Para isso, faz-se necessário definir uma métrica a ele envolvida. A construção de tal métrica será detalhada no trabalho e é conhecida como distância de Hausdorff.

Keywords

Hausdorff distance.
Hausdorff's space.
Metric space.
Compact.

Abstract

In this paper, we will present the Hausdorff's space, and we will aim to define it as a metric space. For this, it is necessary to define a metric involved. The construction of such a metric will be detailed in the paper and is known as Hausdorff distance.

1 Introdução

As medidas de distâncias entre conjuntos podem ser abordadas de muitas formas, entre as mais convencionais, por exemplo, pela distância euclidiana entre as médias dos elementos dos conjuntos. Porém, em muitos casos, essas distâncias conferem uma noção de proximidade muito diversa daquela resultante da observação intuitiva humana e daquela que leva a interpretação adequada do processo em análise. A evolução desse conceito leva a consideração de outros melhor elaborados, como a *Distância de Hausdorff*, também conhecida como *Métrica de Hausdorff* que é apresentada como mais confiável na maioria dos casos.

Neste trabalho iremos apresentar os passos para a construção da distância de Hausdorff e utilizá-la para mostrar que o espaço de Hausdorff é metrizável. Conceitos preliminares de espaços métricos serão omitidos, mas podem ser encontrados em [1]. Já a ideia da construção da distância de Hausdorff é baseada em [2].

2 Distância de Hausdorff

Definição 2.1. *Seja (X, d) um espaço métrico completo. Chamaremos de espaço de Hausdorff o espaço formado por todos os subconjuntos não vazios compactos de X . Denotaremos por $\mathcal{H}(X)$, ou seja,*

$$\mathcal{H}(X) = \{K \subset X \mid K \text{ é compacto e } K \neq \emptyset\}.$$

A fim de trabalharmos com um espaço métrico é necessário definir uma métrica a ele envolvida. Para isso, definiremos a distância de um elemento do espaço métrico X a um elemento de $\mathcal{H}(X)$. A construção de tal métrica será detalhada a seguir e é conhecida como *distância de Hausdorff*.

Definição 2.2. *Sejam (X, d) um espaço métrico completo, $x \in X$ e $B \in \mathcal{H}(X)$. Definimos a distância do ponto x ao conjunto B como*

$$d(x, B) = \min \{d(x, y); y \in B\}.$$

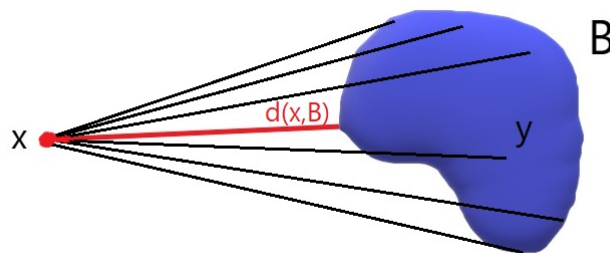


Figura 1: Distância de ponto a conjunto.

O conjunto de números reais $\{d(x, y); y \in B\}$ contém um mínimo valor uma vez que $B \in \mathcal{H}(X)$ é compacto e não vazio.

Definição 2.3. *Sejam (X, d) um espaço métrico completo e $A, B \in \mathcal{H}(X)$. A distância do conjunto A ao conjunto B é definida como*

$$d(A, B) = \max \{d(x, B); x \in A\}.$$

Em outras palavras, $d(A, B)$ é a maior das menores distâncias que separam os pontos de A a qualquer ponto de B .

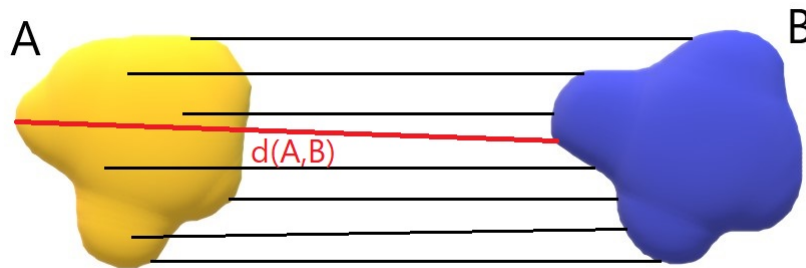


Figura 2: Distância entre dois conjuntos.

Da mesma forma que garantimos a existência de um mínimo, podemos afirmar a existência de um máximo para a definição anterior.

Para um melhor entendimento das definições anteriores, considere os exemplos a seguir.

Exemplo 2.4. *Considere $A = \left\{ \frac{1}{n} \mid n \in \mathbb{N}^* \right\} \cup \{0\}$ e $B = \left\{ \frac{1}{2} \right\}$. Dessa forma, A e B pertencem ao espaço $\mathcal{H}(X)$ e obtemos as distâncias a seguir*

$$\begin{aligned} d(A, B) &= \max \{d(x, B); x \in A\} \\ &= \max \{ \min \{d(x, y); y \in B\}; x \in A \} \\ &= \max \left\{ d \left(0, \frac{1}{2} \right), d \left(1, \frac{1}{2} \right), d \left(\frac{1}{2}, \frac{1}{2} \right), \dots, d \left(\frac{1}{n}, \frac{1}{2} \right), \dots \right\} \\ &= \max \left\{ \frac{1}{2}, 0, \dots, \frac{n-2}{2n}, \dots \right\} = \frac{1}{2}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} d(B, A) &= \max\{d(y, A); y \in B\} \\ &= \max\{\min\{d(y, x); x \in A\}; y \in B\} \\ &= \max\left\{\min\left\{d\left(\frac{1}{2}, 0\right), d\left(\frac{1}{2}, 1\right), d\left(\frac{1}{2}, \frac{1}{2}\right), \dots, d\left(\frac{1}{2}, \frac{1}{n}\right), \dots\right\}\right\} \\ &= \max\{0\} = 0 \end{aligned}$$

Portanto, $d(A, B) \neq d(B, A)$

Exemplo 2.5. Dados os conjuntos A e B dois retângulos, A com vértices $(100, 100)$, $(300, 100)$, $(300, 200)$ e $(100, 200)$ e B com vértices $(400, 100)$, $(800, 100)$, $(800, 200)$ e $(400, 200)$, com isso garantimos que são conjuntos fechados e limitados, o que com a métrica euclidiana garante que são compactos. Desse modo, conseguimos calcular a distância entre esses dois conjuntos, $d(A, B) = 300$ e $d(B, A) = 500$, ou seja $d(A, B) \neq d(B, A)$

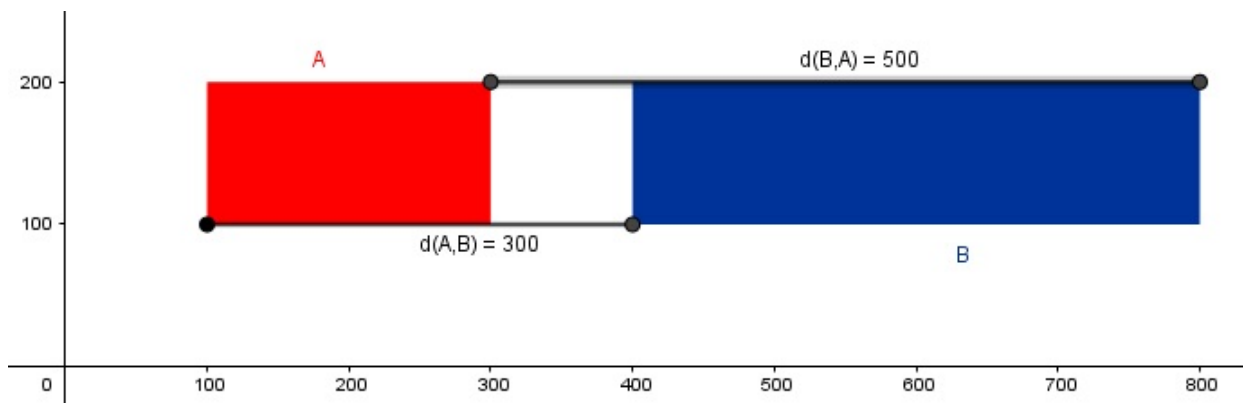


Figura 3: Distância entre dois conjuntos não necessariamente é comutativa.

Observação 2.6. A definição anterior de distância entre conjuntos não pode ser uma métrica, pois nem sempre $d(A, B) = d(B, A)$, como mostram os exemplos anteriores.

Definimos assim a distância de Hausdorff como a maior entre as distâncias $d(A, B)$ e $d(B, A)$. Formalmente temos.

Definição 2.7. Seja (X, d) um espaço métrico completo. A distância de Hausdorff entre os ponto A e B de $\mathcal{H}(X)$ é definida por

$$h(A, B) = d(A, B) \vee d(B, A),$$

onde o símbolo \vee representa o máximo entre as duas distâncias.

Proposição 2.8. Seja (X, d) um espaço métrico. Se $p \in X$ e $A \subset X$, então $d(p, A) = 0$ se, e somente se, $p \in \bar{A}$.

Demonstração. (\Rightarrow) Dado $\epsilon > 0$, como $d(p, A) = \min\{d(p, x); x \in A\} = 0$ existe $a \in A$, tal que $0 \leq d(p, a) < \epsilon$, caso contrário teríamos $0 < \epsilon \leq d(p, x)$, para todo $x \in A$, o que não é possível, pois, por hipótese $d(p, A) = 0$.

Então, $a \in B(p, \epsilon)$ e, portanto, $B(p, \epsilon) \cap A \neq \emptyset$. Logo, $p \in \bar{A}$.

(\Leftarrow) Suponhamos que $d(p, A) = \epsilon > 0$. Como por hipótese $B(p, \epsilon) \cap A \neq \emptyset$, então existe $a \in A$, tal que $d(a, p) < \epsilon$.

Assim, $\epsilon = d(p, A) \leq d(p, a) < \epsilon$, o que é um absurdo. Logo, $d(p, A) = 0$. ■

Teorema 2.9. A aplicação $h(A, B)$, distância de Hausdorff definida acima, é uma métrica no espaço $\mathcal{H}(X)$.

Demonstração. Tome $A, B, C \in \mathcal{H}(X)$. Iremos mostrar que a aplicação h satisfaz as três propriedades da definição de métrica .

(i) Inicialmente, suponhamos $h(A, B) = 0$.

Dessa forma, $d(A, B) \vee d(B, A) = 0$, ou seja, $\max\{d(A, B), d(B, A)\} = 0$, como distância sempre é maior ou igual a zero, temos $0 \leq d(A, B) \leq 0$ e $0 \leq d(B, A) \leq 0$.

Logo, $d(A, B) = d(B, A) = 0$.

Por outro lado, $d(A, B) = \max\{d(x, B); x \in A\}$ e $d(B, A) = \max\{d(y, A); y \in B\}$.

Dessa forma, para todo $x \in A$, temos $0 \leq d(x, B) \leq d(A, B) = 0$, ou seja, $d(x, B) = 0$.

Pela Proposição 2.8 , segue que $x \in \bar{B} = B$, pois B é compacto. Logo, $A \subset B$.

Analogamente, para todo $y \in B$, temos $d(y, A) = 0$ e assim, $y \in \bar{A} = A$. Logo, $B \subset A$.

Portanto, se $h(A, B) = 0$ temos que $A = B$.

Reciprocamente, suponhamos $A = B$. Então, $d(x, B) = d(y, A) = 0$, para quaisquer $x \in A$ e $y \in B$, obtendo assim

$$d(A, B) = \max\{d(x, B); x \in A\} = 0 = \max\{d(y, A); y \in B\} = d(B, A).$$

Portanto, $h(A, B) = d(A, B) \vee d(B, A) = 0$.

(ii) Por definição,

$$\begin{aligned} h(A, B) &= d(A, B) \vee d(B, A) \\ &= \max\{d(A, B), d(B, A)\} \\ &= d(B, A) \vee d(A, B) \\ &= h(B, A). \end{aligned}$$

(iii) Agora mostraremos a desigualdade triangular. Porém precisamos provar inicialmente que, $d(A, B) \leq d(A, C) + d(C, B)$.

De fato, para todo $x \in A$, temos $d(x, B) = \min \{d(x, y) ; y \in B\}$, Além disso , como d é uma métrica, para todo $z \in C$, $d(x, y) \leq d(x, z) + d(z, y)$. Logo,

$$\begin{aligned} d(x, B) &= \min\{d(x, y); y \in B\} \\ &\leq \min\{d(x, z) + d(z, y); y \in B\}, \forall z \in C \\ &= d(x, z) + \min\{d(z, y); y \in B\}, \forall z \in C \\ &= d(x, z) + d(z, B), \forall z \in C. \end{aligned}$$

Consequentemente, $d(x, B) \leq \min\{d(x, z) + d(z, B) ; z \in C\}$.

Note ainda que $d(z, B) \leq \max\{d(z, B) ; z \in C\} = d(C, B)$.

Então,

$$\begin{aligned} d(x, B) &\leq \min\{d(x, z) + d(z, B) ; z \in C\} \\ &\leq \min\{d(x, z) ; z \in C\} + d(C, B) \\ &= d(x, C) + d(C, B). \end{aligned}$$

Logo, $d(A, B) \leq d(A, C) + d(C, B)$.

Analogamente, podemos mostrar que $d(B, A) \leq d(B, C) + d(C, A)$.

Portanto,

$$h(A, B) = d(A, B) \vee d(B, A) \leq (d(A, C) \vee d(C, A)) + (d(C, B) \vee d(B, C)) = h(A, C) + h(C, B) .$$

Logo, por (i), (ii) e (iii) provamos que h é uma métrica no espaço $\mathcal{H}(X)$. ■

3 Considerações finais

A distância de Hausdorff é definida para medir o quanto dois conjuntos de pontos dispersos em um espaço cartesiano são similares. Tal distância é aplicada em geometria fractal e também em problemas de medida de similaridade entre geometrias computacionais.

Referências

- [1] DOMINGUES, H. H. **Espaços métricos e introdução à topologia**. São Paulo: Editora da Universidade de São Paulo, 1982.
- [2] BARNESLEY, M. **Fractals Everywhere**. San Diego: Academic Press, 1988.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Perturbações de sistemas lineares bidimensionais $x' = Ax$ Perturbations of two-dimensional linear systems $x' = Ax$

Rodrigo Lima Almeida

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil

rodrigo.almeida1@ufu.br

Hernán Roberto Montúfar López

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil

montufar@ufu.br

Resumo

Palavras-chave

Sistemas lineares.
Pontos críticos.
Retrato de fase.
Estabilidade.
Perturbações.

Este trabalho trata de uma análise geométrica do comportamento das soluções de sistemas lineares bidimensionais com coeficientes constantes. Essa análise é importante para entender a dinâmica dos fenômenos descritos pelos sistemas e na classificação dos chamados pontos críticos. Trataremos de sistemas cujo único ponto crítico encontra-se na origem. Veremos que depois de um longo período de tempo, cada trajetória individual exibe apenas um entre três tipos de comportamento. Quando $t \rightarrow \infty$, cada trajetória se aproxima do ponto crítico $x = 0$ ou percorre, repetidamente, uma curva fechada em torno do ponto crítico, ou torna-se ilimitada. Por último, analisamos quando o retrato de fase persiste sob pequenas perturbações dos sistemas lineares.

Abstract

Keywords

Linear systems.
Critical points.
Phase portrait.
Stability.
Perturbations.

This work deals with a geometric analysis of the behavior of the solutions of two-dimensional linear systems with constant coefficients. This analysis is important to understand the dynamics of the phenomena described by the systems and the classification of the so-called critical points. We will deal with systems whose only critical point is the origin. We will see that after a long period of time, each individual trajectory exhibits only one of three types of behavior. When $t \rightarrow \infty$, each trajectory approaches the critical point $x = 0$ or repeatedly travels a tight curve around the critical point, or becomes unlimited. Finally, we analyze when the phase portrait persists under minor perturbations of linear systems.

1 Introdução

Sistemas de equações diferenciais aparecem no estudo das leis que regem o comportamento do mundo físico. Como para muitos desses sistemas, encontrar uma solução de forma analítica não é fácil, nem muitas vezes conveniente, buscaremos, através de uma análise geométrica, estudar a estabilidade das soluções em pontos críticos. Esse tipo de abordagem faz parte da teoria qualitativa de equações diferenciais, criada por Henri Poincaré (1854-1912). Consideraremos sistemas com duas equações diferenciais simultâneas da forma

$$\mathbf{x}' = \mathbf{A}\mathbf{x}$$

em que \mathbf{A} é uma matriz constante 2×2 com $\det \mathbf{A} \neq 0$ e \mathbf{x} é um vetor 2×1 . Uma **solução** no intervalo aberto I é a função diferenciável $\mathbf{x} = \varphi(t)$ tal que $\varphi'(t) = \mathbf{A}\varphi(t)$ em todos os pontos do intervalo I .

Teorema 1.1. *Para todo $(t_0, x_0, y_0) \in I \times \mathbf{R}^2$ existe uma única solução $\varphi(t) = \varphi(t, t_0, x_0, y_0)$ de $\mathbf{x}' = \mathbf{A}\mathbf{x}$ em I tal que $\varphi(t_0) = (x_0, y_0)$.*

Os pontos em que $\mathbf{A}\mathbf{x} = 0$ são chamados de **pontos críticos** do sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$. Em tais pontos, temos $\mathbf{x}' = 0$, de modo que os pontos críticos correspondem a soluções constantes, ou de equilíbrio, do sistema de equações diferenciais. Vamos supor que \mathbf{A} é invertível, ou seja, que $\det \mathbf{A} \neq 0$. Segue que $\mathbf{x} = 0$ é o único ponto crítico do sistema linear.

Se procurarmos soluções da forma $\mathbf{x} = \xi e^{rt}$, então substituindo no sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$, obtemos

$$(\mathbf{A} - r\mathbf{I})\xi = 0.$$

Isso corresponde a que r seja um autovalor e ξ um autovetor associado da matriz de coeficientes \mathbf{A} . Os autovalores são as raízes da equação polinomial: $\det(\mathbf{A} - r\mathbf{I}) = 0 \Leftrightarrow r^2 - (\text{tr} \mathbf{A})r + \det(\mathbf{A}) = 0$. Logo, $r_1, r_2 = \frac{\text{tr} \mathbf{A} \pm \sqrt{(\text{tr} \mathbf{A})^2 - 4\det(\mathbf{A})}}{2}$. Onde $\text{tr} \mathbf{A} = a_{11} + a_{22}$ é o traço de \mathbf{A} . Temos os seguintes casos:

1. Os autovalores r_1, r_2 de \mathbf{A} são reais, distintos e não nulos.
2. Os autovalores são complexos conjugados, necessariamente com parte imaginária não nula.
3. Os autovalores são reais, iguais e não nulos.

A solução $\mathbf{x} = \varphi(t)$ pode ser considerada como uma representação paramétrica de uma curva no plano xy , o **plano de fase**. Um conjunto representativo de trajetórias é chamado de **retrato de fase**.

2 Estabilidade dos pontos críticos

Vamos determinar a configuração geométrica de todas as soluções de $\mathbf{x}' = \mathbf{A}\mathbf{x}$ e caracterizar a equação diferencial de acordo com o padrão geométrico formado por suas trajetórias, (veja [1]). Um ponto crítico \mathbf{x}_0 é dito **estável** se, dado qualquer $\epsilon > 0$, existe um $\delta > 0$ tal que toda solução $\mathbf{x} = \varphi(t)$

do sistema, que satisfaz em $t = 0$, $\|\varphi(0) - \mathbf{x}_0\| < \delta$, existe para todo $t > 0$ e satisfaz $\|\varphi(t) - \mathbf{x}_0\| < \epsilon$ para todo $t \geq 0$. Um ponto crítico que não é estável é dito **instável**. Um ponto crítico \mathbf{x}_0 é dito **assintoticamente estável** se é estável e se existe um δ_0 , com $0 < \delta_0 < \delta$, tal que se uma solução $\mathbf{x} = \varphi(t)$ satisfaz $\|\varphi(0) - \mathbf{x}_0\| < \delta_0$, então $\lim_{t \rightarrow \infty} \varphi(t) = \mathbf{x}_0$.

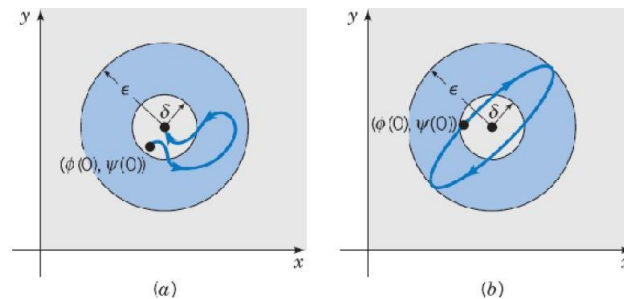


Figura 1: (a) ponto crítico estável e assintoticamente estável. (b) ponto crítico estável.

Ao analisar o sistema linear $\mathbf{x}' = \mathbf{Ax}$, precisamos considerar diversos casos diferentes, dependendo da natureza dos autovalores r_1 e r_2 de \mathbf{A} , (veja [2]).

Teorema 2.1. *O ponto crítico $\mathbf{x} = 0$ do sistema linear $\mathbf{x}' = \mathbf{Ax}$ será:*

1. *assintoticamente estável, se r_1 e r_2 são reais e negativos ou tem parte real negativa;*
2. *estável, mas não assintoticamente estável, se r_1 e r_2 são imaginários puros;*
3. *instável, se r_1 e r_2 são reais e um deles é positivo ou se ambos tem parte real positiva.*

Demonstração. Vamos analisar cada um dos casos possíveis:

- **CASO 1: Os autovalores r_1, r_2 de \mathbf{A} são reais, distintos e não nulos.** Sejam v_1, v_2 os autovetores associados aos autovalores r_1, r_2 . Denotamos por E_1 e E_2 as linhas geradas pelos autovetores. Toda solução de $\mathbf{x}' = \mathbf{Ax}$ pode ser escrita como $\varphi(t) = c_1 e^{r_1 t} v_1 + c_2 e^{r_2 t} v_2$. Portanto, temos: **nó atrator** ($r_2 < r_1 < 0$), **nó instável/fonte** ($r_2 > r_1 > 0$) e **sela** ($r_2 > 0 > r_1$).

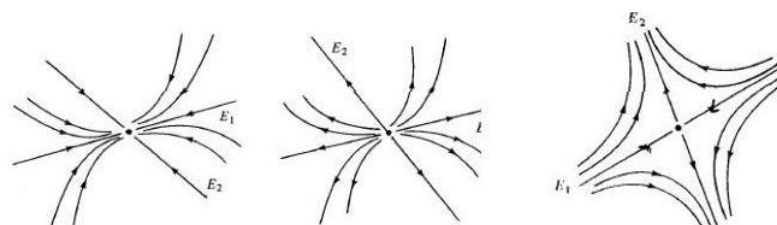


Figura 2: nó atrator (à esq.), nó instável (ao centro), sela (à dir.).

- **CASO 2: Os autovalores são complexos conjugados, necessariamente com parte imaginária não nula.** Toda solução de $\mathbf{x}' = \mathbf{Ax}$ pode ser escrita na forma $\varphi(t) = e^{\alpha t} \rho [\cos(\omega - \beta t)v_1 + \text{sen}(\omega - \beta t)v_2]$. Portanto, temos: **centro** ($\alpha = 0$), **foco atrator** ($\alpha < 0$) e **foco instável** ($\alpha > 0$).

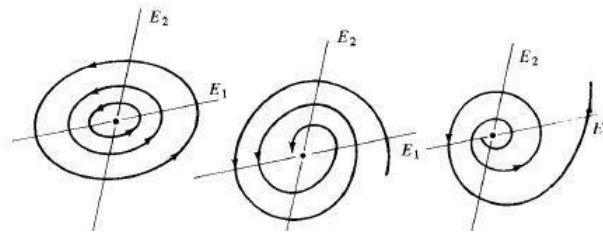


Figura 3: centro (à esq.), foco atrator (ao centro), foco instável (à dir.).

• **CASO 3: Os autovalores são reais, iguais e não nulos.**

- Se o núcleo de $\mathbf{A} - \lambda\mathbf{I}$ é bidimensional, a solução pode ser escrita como: $\varphi(t) = e^{\lambda t}(c_1v_1 + c_2v_2)$. Conforme o sinal do autovalor λ da matriz de coeficientes \mathbf{A} , tem-se as seguintes configurações.

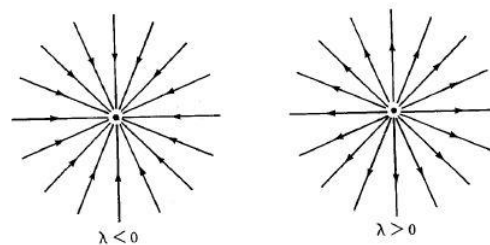


Figura 4: Autovalores iguais: dois autovetores independentes.

- Se o núcleo de $\mathbf{A} - \lambda\mathbf{I}$ é unidimensional, então temos um autovalor de multiplicidade algébrica 2 com apenas um autovetor independente associado v_1 . É possível demonstrar que $\varphi(t) = e^{\lambda t}[(c_1 + tc_2)v_1 + c_2v_2]$ é solução de $\mathbf{x}' = \mathbf{A}\mathbf{x}$, com $\varphi(0) = c_1v_1 + c_2v_2$ e $v_2 \nparallel v_1$. A configuração geométrica também vai se diferenciar conforme o sinal do autovalor. Se $\lambda < 0$, toda trajetória tende à origem conforme $t \rightarrow +\infty$. Se $\lambda > 0$, toda trajetória se afasta da origem conforme $t \rightarrow +\infty$.

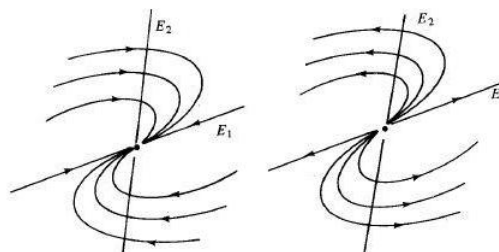


Figura 5: Autovalores iguais: um autovetor independente.



Segue-se do teorema que os autovalores r_1, r_2 da matriz \mathbf{A} determinam o tipo de ponto crítico em $\mathbf{x} = 0$ e suas características de estabilidade.

3 Análise de perturbações

Quando um sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$ aparece em algum campo aplicado, os coeficientes resultam, em geral, de medidas de determinadas quantidades físicas. Tais medidas estão sujeitas, muitas vezes, a pequenos erros, de modo que é de interesse investigar se pequenas mudanças (perturbações) nos coeficientes, afetam a estabilidade ou instabilidade de um ponto crítico e/ou o padrão das trajetórias.

Como os autovalores r_1, r_2 são as raízes da equação polinomial $\det(\mathbf{A} - r\mathbf{I}) = 0$, então pequenas perturbações em alguns dos coeficientes são refletidas em pequenas perturbações nos autovalores.

• **CASO 1: Os autovalores r_1, r_2 de \mathbf{A} são reais, distintos e não nulos.** Neste caso, perturbações suficientemente pequenas dos coeficientes não alteram a estabilidade ou instabilidade do sistema, nem o tipo de ponto crítico. Por exemplo, se r_1 e r_2 são reais, negativos e distintos, então uma mudança pequena nos coeficientes não vai alterar os sinais de r_1 e r_2 , nem vai permitir que eles se tornem iguais. Assim, o ponto crítico permanecerá um nó assintoticamente estável.

• **CASO 2: Os autovalores são complexos conjugados, necessariamente com parte imaginária não nula.** A situação sensível acontece quando $r_1 = i\mu$ e $r_2 = -i\mu$, ou seja, quando o ponto crítico é um centro. Uma ligeira mudança nos coeficientes, implica em novos autovalores $r'_1 = \lambda' + i\mu'$ e $r'_2 = \lambda' - i\mu'$, onde λ' é pequeno em valor absoluto e $\mu' \cong \mu$ (Figura 6). Se $\lambda' \neq 0$, então as trajetórias do sistema perturbado serão espirais, em vez de elipses. O sistema é assintoticamente estável se $\lambda' < 0$, mas é instável se $\lambda' > 0$.

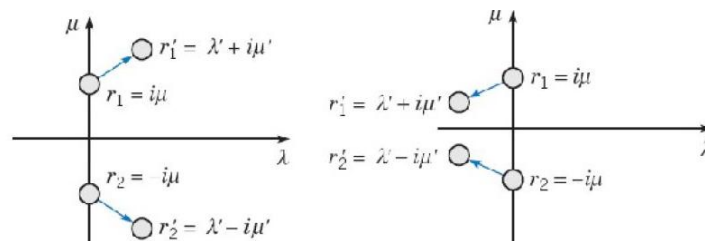


Figura 6: Perturbação esquemática de $r_1 = i\mu, r_2 = -i\mu$.

Por exemplo, sejam as matrizes $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} \epsilon & 1 \\ -1 & \epsilon \end{pmatrix}$ com $|\epsilon|$ arbitrariamente pequeno.

Para o sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$ temos $\det(\mathbf{A} - r\mathbf{I}) = 0 \Leftrightarrow r^2 + 1 = 0 \Leftrightarrow r = \pm i$. Logo $(0, 0)$ é um centro. Para o sistema $\mathbf{x}' = \mathbf{B}\mathbf{x}$ temos $\det(\mathbf{B} - r\mathbf{I}) = 0 \Leftrightarrow r^2 - 2\epsilon r + 1 + \epsilon^2 = 0 \Leftrightarrow r = \epsilon \pm i$. Logo $(0, 0)$ é um ponto espiral para $\epsilon \neq 0$. Se $\epsilon < 0$, o ponto espiral vai ser assintoticamente estável; se $\epsilon > 0$, o ponto espiral vai ser instável.

• **CASO 3: Os autovalores são reais, iguais e não nulos.** Neste caso, o ponto crítico é um nó. Pequenas perturbações nos coeficientes fazem com que as raízes iguais se separem (bifurquem). Se as raízes separadas forem reais, então o ponto crítico do sistema perturbado permanecerá um nó, mas se as raízes separadas forem complexas conjugadas, então o ponto crítico se transformará em um ponto espiral (Figura 7). Nesse caso, a estabilidade ou instabilidade do sistema não é afetada por pequenas

perturbações nos coeficientes, mas o tipo de ponto crítico pode mudar.

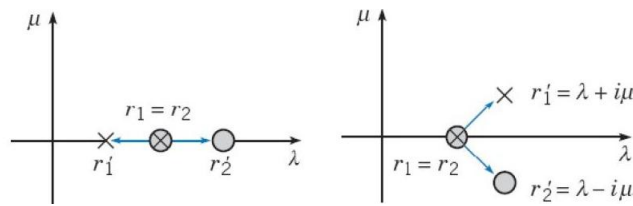


Figura 7: Perturbação esquemática de $r_1 = r_2$.

Por exemplo, sejam as matrizes $A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$ e $B = \begin{pmatrix} -1 & 1 \\ -\epsilon & -1 \end{pmatrix}$ com $|\epsilon|$ arbitrariamente pequeno. Para o sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$ temos $\det(\mathbf{A} - r\mathbf{I}) = 0 \Leftrightarrow (r + 1)^2 = 0 \Leftrightarrow r_1 = r_2 = -1$, de modo que $(0, 0)$ é um nó assintoticamente estável. Para o sistema $\mathbf{x}' = \mathbf{B}\mathbf{x}$ temos $\det(\mathbf{B} - r\mathbf{I}) = 0 \Leftrightarrow r^2 + 2r + 1 + \epsilon = 0 \Leftrightarrow r = -1 \pm \sqrt{-\epsilon}$. Se $\epsilon > 0$, então os autovalores são $-1 \pm i\sqrt{\epsilon}$, de modo que o nó assintoticamente estável se transformou em um ponto espiral assintoticamente estável. Se $\epsilon < 0$, então os autovalores são $-1 \pm \sqrt{|\epsilon|}$, e o ponto crítico permanece sendo um nó assintoticamente estável.

4 Considerações finais

- Pontos críticos correspondem a soluções de equilíbrio, ou seja, soluções constantes. Pode parecer razoável concluir que tais pontos não são muito interessantes, mas para sistemas lineares $\mathbf{x}' = \mathbf{A}\mathbf{x}$ a natureza do ponto crítico praticamente determina o comportamento das trajetórias.
- O comportamento das trajetórias do sistema $\mathbf{x}' = \mathbf{A}\mathbf{x}$ quanto $t \rightarrow +\infty$ são basicamente os seguintes: se aproximam do ponto crítico na origem, ou percorrem curvas fechadas em torno dele ou tornam-se ilimitadas.
- A dependência do comportamento global das soluções aos autovalores fica evidente ao perturbarmos o sistema, em dois dos casos essas perturbações alteram a natureza dos autovalores e isso influencia, em maior ou menor grau, o comportamento das trajetórias no plano de fase.

Referências

- [1] BOYCE, William E.; DIPRIMA, Richard C. **Equações diferenciais elementares e problemas de valores de contorno**. Guanabara Dois, 1985.
- [2] Sotomayor, J. **Lições de equações diferenciais ordinárias**. Instituto de Matemática Pura e Aplicada, 1979.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Propriedades notáveis de $\ell_p, 1 \leq p < \infty$ Notable properties of $\ell_p, 1 \leq p < \infty$

Stefânia Carvalho de Sousa

UFU, FAMAT/ Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil
stefaniacarvalho12@gmail.com

Fábio José Bertoloto

UFU, FAMAT/ Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil
bertoloto@ufu.br

Palavras-chave

Espaços Normados.
Espaços de Banach.
Separabilidade.

.

Resumo

O objetivo deste trabalho é mostrar que o conjunto $\ell_p, 1 \leq p < \infty$, que será posteriormente definido, é um espaço de Banach e, além disso, é separável. Para isso, serão necessários resultados preliminares, como definições de espaços métricos, de espaços normados, além da desigualdade de Minkowski.

Keywords

Normed Spaces.
Banach Spaces.
Separability.

.

Abstract

The purpose of this work is to show that the set $\ell_p, 1 \leq p < \infty$, which will be defined later, is a Banach space and, moreover, that is separable. For this, preliminary results will be necessary, such as definitions of metric spaces, normed spaces, besides the Minkowski's inequality

1 Introdução

Stefan Banach foi um matemático polonês que contribuiu de forma significativa aos estudos de Análise Funcional. Em 1932 publicou o seu trabalho mais importante, *Théorie des opérations linéaires* (Teoria dos operadores lineares). Dentre muitas de suas contribuições, os espaços normados completos, que receberam o nome de espaços de Banach devido a este matemático, serão utilizados neste trabalho.

O objetivo central é estudar o conjunto $\ell_p, 1 \leq p < \infty$. Definiremos uma norma sobre esse conjunto, de maneira que ele passará a ser um caso especial de espaço de Banach. Além disso, utilizando o conceito de conjunto enumerável, iremos mostrar que ℓ_p é separável.

Sempre trabalharemos com o corpo como sendo $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$.

2 Resultados Preliminares

2.1 Métrica e Norma

Definição 2.1. Dado um conjunto não vazio X , chamamos de métrica em X uma função $d: X \times X \rightarrow \mathbb{R}$ que associa a cada par ordenado $(x, y) \in X \times X$ um número real $d(x, y)$, chamado de distância de x a y , que satisfaz três condições:

- (1) $d(x, y) \geq 0$ e $d(x, y) = 0 \Leftrightarrow x = y$;
- (2) $d(x, y) = d(y, x)$ (Simetria);
- (3) $d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in X$ (Desigualdade Triangular).

Um Espaço Métrico é um par ordenado (X, d) em que X é um conjunto não vazio e d é uma métrica em X .

Exemplos usuais de métricas são as métricas euclidiana, do máximo e da soma (do táxi) em \mathbb{R}^n . Ver [1, p. 2].

Definição 2.2. Em cada espaço vetorial X , sobre um corpo \mathbb{K} , pode ser definida uma noção da distância de um elemento arbitrário à origem, ou seja, o vetor nulo do espaço. Logo, a distância de um elemento $x \in X$ à origem pode ser vista como o tamanho de x . Essa distância até a origem é chamada de norma e é denotada por $\|x\|$. Três condições são satisfeitas por uma norma:

- (1) $\|x\| \geq 0$ e $\|x\| = 0 \Leftrightarrow x = 0$;
- (2) $\|\alpha x\| = |\alpha| \|x\|$, para todo $\alpha \in \mathbb{K}$;
- (3) $\|x + y\| \leq \|x\| + \|y\|$ (Desigualdade triangular).

Um espaço normado pode sempre ser visto como um espaço métrico, em que a métrica é induzida pela norma da seguinte maneira: $d(x, y) = \|x - y\|, \forall x, y \in X$.

Não é difícil provar que d é uma métrica, chamada *métrica natural*.

2.2 Espaços de Banach

Definição 2.3. *Sejam X um espaço métrico com uma métrica d e (x_n) uma sequência em X . Dizemos que (x_n) é uma sequência de Cauchy se:*

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N} : m, n \geq n_0 \Rightarrow d(x_m, x_n) < \epsilon.$$

Definição 2.4. *Um espaço métrico (X, d) é dito completo se toda sequência de Cauchy em X for convergente em X .*

Definição 2.5. *Um espaço de Banach é um espaço vetorial normado que é completo na métrica induzida pela norma.*

O termo espaço de Banach foi cunhado pelo matemático francês M. Fréchet em 1928.

3 Espaços normados de sequências

Definição 3.1. *Dado $1 \leq p < \infty$, o conjunto ℓ_p é definido por:*

$$\ell_p = \left\{ (\xi_n)_{n=1}^{\infty} : \xi_j \in \mathbb{K} \text{ para todo } j \in \mathbb{N} \text{ e } \sum_{j=1}^{\infty} |\xi_j|^p < \infty \right\}.$$

Lema 3.2. *Sejam $a, b, \alpha, \beta > 0$, com $\alpha + \beta = 1$. Então, $a^\alpha b^\beta \leq a\alpha + b\beta$.*

Demonstração: Tal demonstração pode ser encontrada em [2, p. 6] ou [3, p. 5]. ■

Teorema 3.3 (Desigualdade de Hölder para somas). *Sejam $1 < p, q < \infty$, com $\frac{1}{p} + \frac{1}{q} = 1$, e sejam $(\xi_1, \dots, \xi_n), (\eta_1, \dots, \eta_n) \in \mathbb{K}^n$. Então:*

$$\sum_{j=1}^n |\xi_j \eta_j| \leq \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\eta_j|^q \right)^{\frac{1}{q}}.$$

Demonstração. Sejam $a_j = \frac{|\xi_j|^p}{\sum_{j=1}^n |\xi_j|^p}$, $b_j = \frac{|\eta_j|^q}{\sum_{j=1}^n |\eta_j|^q}$, $\alpha = \frac{1}{p}$ e $\beta = \frac{1}{q}$.

Segue do Lema 3.2 que:

$$\frac{|\xi_j \eta_j|}{\left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\eta_j|^q \right)^{\frac{1}{q}}} \leq \frac{a_j}{p} + \frac{b_j}{q}.$$

Agora,

$$\frac{\sum_{j=1}^n |\xi_j \eta_j|}{\left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\eta_j|^q \right)^{\frac{1}{q}}} \leq \frac{1}{p} \sum_{j=1}^n a_j + \frac{1}{q} \sum_{j=1}^n b_j = \frac{1}{p} + \frac{1}{q} = 1.$$

Portanto,

$$\sum_{j=1}^n |\xi_j \eta_j| \leq \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\eta_j|^q \right)^{\frac{1}{q}}.$$

■

Proposição 3.4 (Desigualdade de Minkowski). *Sejam $1 \leq p < \infty$ e $(\xi_j), (\eta_j) \in \ell_p$. Então,*

$$\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{\frac{1}{p}}. \quad (1)$$

Disso, concluímos que ℓ_p é um espaço vetorial.

Demonstração. Começemos com somas finitas variando até n , em que n é um natural dado. Para $p = 1$ a desigualdade é clara. Seja $p > 1$. Segue que:

$$\sum_{j=1}^n |\xi_j + \eta_j|^p = \sum_{j=1}^n |\xi_j + \eta_j| |\xi_j + \eta_j|^{p-1} \leq \sum_{j=1}^n |\xi_j| |\xi_j + \eta_j|^{p-1} + \sum_{j=1}^n |\eta_j| |\xi_j + \eta_j|^{p-1}.$$

Como $(p-1)q = p$, segue do Teorema 3.3 que:

$$\sum_{j=1}^n |\xi_j| |\xi_j + \eta_j|^{p-1} \leq \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{\frac{1}{q}}.$$

e

$$\sum_{j=1}^n |\eta_j| |\xi_j + \eta_j|^{p-1} \leq \left(\sum_{j=1}^n |\eta_j|^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{\frac{1}{q}}.$$

Logo,

$$\sum_{j=1}^n |\xi_j + \eta_j|^p \leq \left\{ \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^n |\eta_j|^p \right)^{\frac{1}{p}} \right\} \left(\sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{\frac{1}{q}}.$$

Como $1 - \frac{1}{q} = \frac{1}{p}$, segue que

$$\left(\sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^n |\eta_j|^p \right)^{\frac{1}{p}}.$$

Passando para séries, temos para todo $n \in \mathbb{N}$ que:

$$\left(\sum_{j=1}^n |\xi_j + \eta_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^n |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^n |\eta_j|^p \right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |\xi_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^{\infty} |\eta_j|^p \right)^{\frac{1}{p}}.$$

Assim, como é limitada e monótona não decrescente a sequência das somas parciais, concluímos que converge a série $\left(\sum_{j=1}^{\infty} |\xi_j + \eta_j|^p\right)^{\frac{1}{p}}$ obtendo, portanto, a equação (1). Segue, diretamente que $(\xi_j + \eta_j) \in \ell_p$, o que implica no fato de ℓ_p ser um espaço vetorial. ■

Proposição 3.5. ℓ_p é um espaço de Banach com a norma $\|\cdot\|_p$ dada por $\|(\xi_j)\|_p = \left(\sum_{j=1}^{\infty} |\xi_j|^p\right)^{\frac{1}{p}}$, para cada $(\xi_j) \in \ell_p$.

Demonstração. Segue da Proposição 3.4 que ℓ_p é um espaço vetorial e que $\|\cdot\|_p$ é uma norma em ℓ_p . Provemos que ℓ_p é completo.

Seja (x_n) uma sequência de Cauchy em ℓ_p , com $x_n = (\xi_{nj})_{j=1}^{\infty}$, para todo $n \in \mathbb{N}$. Então, dado $\epsilon > 0$, existe n_0 tal que

$$\|x_n - x_m\|_p = \left(\sum_{j=1}^{\infty} |\xi_{nj} - \xi_{mj}|^p\right)^{\frac{1}{p}} \leq \epsilon, \forall n, m \geq n_0.$$

Em particular, $|\xi_{nj} - \xi_{mj}| \leq \|x_n - x_m\|_p \leq \epsilon$, para todo $m, n \geq n_0$ e todo $j \in \mathbb{N}$. Assim, $(\xi_{nj})_{n=1}^{\infty}$ é de Cauchy em \mathbb{K} . Como \mathbb{K} é completo, do fato de $(\xi_{nj})_{n=1}^{\infty}$ ser uma sequência de Cauchy em \mathbb{K} , para cada $j \in \mathbb{N}$, existe $\xi_j = \lim_n \xi_{nj}$. Escrevamos $x = (\xi_j)_{j=1}^{\infty}$.

Temos que,

$$\left(\sum_{j=1}^k |\xi_{nj} - \xi_{mj}|^p\right)^{\frac{1}{p}} \leq \left(\sum_{j=1}^{\infty} |\xi_{nj} - \xi_{mj}|^p\right)^{\frac{1}{p}} \leq \epsilon, \forall n, m \geq n_0, k \in \mathbb{N}.$$

Como $\left(\sum_{j=1}^k |\xi_{nj} - \xi_{mj}|^p\right)^{\frac{1}{p}} \leq \epsilon$, fazendo $m \rightarrow \infty$, segue, para $k \in \mathbb{N}$, que:

$$\left(\sum_{j=1}^k |\xi_{nj} - \xi_j|^p\right)^{\frac{1}{p}} \leq \epsilon, \forall n \geq n_0.$$

Logo, $\left(\sum_{j=1}^{\infty} |\xi_{nj} - \xi_j|^p\right)^{\frac{1}{p}} \leq \epsilon$, para todo $n \geq n_0$. Assim, $x_n - x \in \ell_p$ e $\|x_n - x\|_p \leq \epsilon$, para todo $n \geq n_0$. Disso, $x = (x - x_n) + x_n \in \ell_p$ e $\|x_n - x\|_p \rightarrow 0$. Portanto, ℓ_p é um espaço de Banach. ■

Definição 3.6. Um espaço métrico X é dito separável se existir um subconjunto enumerável $D \subset X$ que é denso em X , ou seja, $\overline{D} = X$, em que \overline{D} é o fecho de D .

Proposição 3.7. ℓ_p , com a norma $\|\cdot\|_p$, é separável para cada $1 \leq p < \infty$.

Demonstração. Sejam

$$c_{00} = \{(\xi_j)_{j=1}^{\infty} : \xi_j \in \mathbb{K} \text{ para todo } n \in \mathbb{N} \text{ e existe } k \in \mathbb{N} \text{ tal que } \xi_j = 0 \text{ para todo } j \geq k\} \text{ e}$$

$$D = \{(\xi_j)_{j=1}^{\infty} \in c_{00} : \xi_j \in \mathbb{Q}\}.$$

De [4, p. 8], segue que D é enumerável. De fato, podemos escrever

$$D = \bigcup_{n=1}^{\infty} \mathbb{Q}^n \text{ onde } \mathbb{Q}^n = \mathbb{Q} \times \mathbb{Q} \times \cdots \times \mathbb{Q} \text{ (} n\text{-vezes)}.$$

Provemos, então, que D é denso em ℓ_p .

Sejam $x = (\xi_j) \in \ell_p$ e $\epsilon > 0$. Como $\sum_{j=1}^{\infty} |\xi_j|^p < \infty$, existe $n \in \mathbb{N}$ tal que $\sum_{j=n+1}^{\infty} |\xi_j|^p < \epsilon^p$. Do fato de \mathbb{Q} ser denso em \mathbb{R} , podemos tomar $y = (\xi_1, \dots, \xi_n, 0, 0, 0, \dots)$ e $z = (\zeta_1, \dots, \zeta_n, 0, 0, 0, \dots)$, com ζ_1, \dots, ζ_n racionais, tais que:

$$\sum_{j=1}^n |\xi_j - \zeta_j|^p < \epsilon^p.$$

Então, $y \in c_{00}$, $z \in D$ e

$$\|x - z\|_p \leq \|x - y\|_p + \|y - z\|_p < 2\epsilon.$$

Portanto, D é denso em ℓ_p , sendo este separável. ■

4 Considerações finais

O conteúdo deste trabalho é parte de um projeto de iniciação científica. Por exemplo, a desigualdade de Hölder, aqui utilizada, foi estudada em detalhes, sendo visto também que ela é válida para séries, apesar de aqui ter sido apresentada apenas para somas finitas.

Os resultados apresentados para ℓ_p , $1 \leq p < \infty$, também podem ser relacionados com $p = \infty$, sendo

$$\ell_{\infty} = \left\{ (\xi_j)_{j=1}^{\infty} : \xi_j \in \mathbb{K} \text{ para todo } j \in \mathbb{N} \text{ e } \sup_{j \in \mathbb{N}} |\xi_j| < \infty \right\}.$$

No caso, ℓ_{∞} é um espaço de Banach com a norma do supremo dada por $\|(\xi_j)_{j=1}^{\infty}\|_{\infty} = \sup_{j \in \mathbb{N}} \{|\xi_j|\}$, mas não é separável (ver [2, p. 17]). Não foram feitas as demonstrações deste caso, por não fazerem parte do objetivo deste trabalho.

Referências

- [1] LIMA, E. L. **Espaços Métricos**. 2ª Edição. Rio de Janeiro: IMPA, 1983.
- [2] BOTELHO, G.; PELLEGRINO, D.; TEIXEIRA, E. **Fundamentos de Análise Funcional**. Rio de Janeiro: SBM, 2015.
- [3] MUJICA, J. **Notas de Aula de Análise Funcional**. UNICAMP, 2003.
- [4] LIMA, E. L. **Análise Real**. 10ª Edição, vol. 1. Rio de Janeiro: IMPA, 2009.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

De Grupos até Subgrupos Normais: Qual a Construção? From Groups to Normal Subgroups: What is the Construction?

Victor Cruz Borges

UFU, FAMAT, Uberlândia, Minas Gerais, Brasil

victor.cruz@ufu.br

Cícero Fernandes de Carvalho

UFU, FAMAT, Uberlândia, Minas Gerais, Brasil

cicero@ufu.br

Resumo

Palavras-chave

Relação de equivalência.
Conjunto.
Grupo.
Classes.
Inteiros.

A teoria de grupos é objeto fundamental para introduzir os conceitos da álgebra abstrata. Nessa área são estudados conjuntos, suas relações com elementos deles ou até mesmo de outros a partir de relações de equivalência, de ordem e funções. Nesse presente trabalho serão mostrados conceitos construídos a partir de relações de equivalência em conjuntos não vazios com uma estrutura algébrica operacional definida neles, isto é, em grupos. E nesse caminho há a menção do que são subgrupos, subgrupos normais e a definição do grupo quociente junto com um teorema que relaciona todos estes conceitos.

Abstract

Keywords

Equivalence Relation.
Set.
Group.
Classes.
Integers.

Group theory is a fundamental object to introduce the concepts of abstract algebra. In this area, sets are studied, their relations with elements of them or even others based on relations of equivalence, order, functions. In this work, the concepts constructed from equivalence relations in non-empty sets with an algebraic operational structure defined in them, that is, in groups, will be shown. And in this way there is a mention of what are subgroups, normal subgroups and the definition of the quotient group together with a theorem that relates all these concepts.

1 Introdução

A base inicial da teoria de Grupos vem da noção primitiva e intuitiva de conjuntos e seus elementos. Nesse âmbito, para começar a relacioná-los, é necessário definir as

2 Relações de Equivalência

Definição 2.1. *Seja X um conjunto não vazio. Uma relação de equivalência em X é um subconjunto de $X \times X$.*

Definição 2.2. *Seja X um conjunto não vazio. A relação R em X é chamada de relação de equivalência se*

1. $(x, x) \in R$ para todo $x \in X$ (R é reflexiva),
2. Se $(x, y) \in R$ então $(y, x) \in R$ (R é simétrica),
3. Se $(x, y), (y, z) \in R$, então $(x, z) \in R$ (R é transitiva).

Se R é uma relação em um conjunto X , então a **classe de equivalência** de um elemento $x \in R$ é definida como o conjunto de todos os elementos de X que formam pares ordenados com x pela relação. E escrevemos [1]

$$[x] = \{y \in X : (x, y) \in R\}.$$

Exemplo 2.3. *Seja R uma relação em \mathbb{Z} definida por $(x, y) \in R$ se $y - x$ é divisível por 6. Então $(x, x) \in R$ visto que $x - x = 0 = 6 \cdot 0$, e claramente se $y - x$ é divisível por 6, então $x - y$ também é, então R é simétrica. Finalmente, se $y - x = 6k_1$ e $z - y = 6k_2$ então $z - y + y - x = z - x = 6(k_1 + k_2)$, então R é transitiva. Logo R é uma relação de equivalência em \mathbb{Z} .*

A classe de equivalência do 7 em R é $[7] = \{y \in \mathbb{Z} : 6|(y - 7)\}$. Observe que $y - 7 = 6k$ para algum $k \in \mathbb{Z}$ se $y = 7 + 6k = 1 + 6(k + 1)$, que é se $y - 1 = 6(k + 1)$. Em particular, y é equivalente a 7 se y for equivalente a 1. Então $[7] = [1]$ em R .

Observe que 1 e 7 têm a mesma classe de equivalência, e ambos possuem o mesmo resto na divisão por 6. Isso não é coincidência. Na verdade, em geral, se $(x, y) \in R$ então há algum k tal que $y = x + 6k$, então x e y têm o mesmo resto em $\{0, 1, 2, 3, 4, 5\}$. Isso dá outra maneira de descrever a classe de equivalência de x em R :

$$[x] = \{x + 6k : k \in \mathbb{Z}\}$$

Vamos listar as classes de equivalência distintas de R . Observe que, uma vez que existem apenas seis restos distintos por 6, deve haver exatamente 6 classes de equivalência distintas $[x]$, pois x percorre \mathbb{Z} :

$$[0] = \{6k : k \in \mathbb{Z}\}$$

$$[1] = \{6k + 1 : k \in \mathbb{Z}\}$$

$$[2] = \{6k + 2 : k \in \mathbb{Z}\}$$

$$[3] = \{6k + 3 : k \in \mathbb{Z}\}$$

$$[4] = \{6k + 4 : k \in \mathbb{Z}\}$$

$$[5] = \{6k + 5 : k \in \mathbb{Z}\}$$

Observe que todo inteiro x possui algum resto em $\{0, \dots, 6\}$, então esta é a lista completa de classes de equivalência de \mathbb{Z} . Observe também que o resto positivo produzido na aplicação do algoritmo de divisão para x e 6 é único, então cada par de classes de equivalência distintas tem interseção vazia. Dizemos que o conjunto de classes de equivalência de \mathbb{Z} formam uma partição de \mathbb{Z} .

Definição 2.4. Seja X um conjunto não vazio. Uma partição de X é uma coleção \mathcal{P} de subconjuntos de X que:

1. X é a união dos subconjuntos $P \in \mathcal{P}$,
2. $P \cap Q = \emptyset$ a menos que $P = Q$ para qualquer par de subconjuntos de X contido em \mathcal{P} .

Em outras palavras, \mathcal{P} é uma partição de X se X for a união disjunta dos subconjuntos de X contidos em \mathcal{P} .

3 Teoria de Grupos

Definição 3.1. Diz-se que um conjunto $G \neq \emptyset$ forma um grupo se em G está definida uma operação binária, denominada multiplicação e indicada por \cdot tal que [2]

1. $a, b \in G \Rightarrow a \cdot b \in G$ (fechamento)
2. $a, b, c \in G \Rightarrow a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (lei associativa)
3. $\exists e \in G$ tal que $a \cdot e = e \cdot a = a, \forall a \in G$ (existência de um elemento unidade em G)
4. $\forall a \in G, \exists a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = e$ (existência do elemento inverso em G)

Definição 3.2. Um grupo G é dito abeliano (ou comutativo) se para todo $a, b \in G, a \cdot b = b \cdot a$

Outra característica natural de um grupo é o número de elementos que ele contém. Chamamos isso de ordem de G e indicamos por $\circ(G)$. Esse número é evidentemente mais interessante quando G é finito. Nesse caso, dizemos que G é um grupo finito.

Exemplo 3.3. Seja n um inteiro qualquer. Construimos um grupo de ordem n da seguinte maneira: G consistirá de todos os símbolos $a^i, i = 0, 1, 2, \dots, n - 1$ onde exigimos que $a^0 = a^n = e, a^i \cdot a^j = a^{i+j}$ se $i + j \leq n$ e $a^i \cdot a^j = a^{i+j-n}$ se $i + j > n$. É denominado grupo cíclico de ordem n .

1. Se $b, c \in G$. então $b = a^k$ e $c = a^p$, $k, p \in \mathbb{N}$. Pela definição G temos que $b.c \in G$.
2. Se $b, c, d \in G$ então $b = a^k$, $c = a^m$ e $d = a^p$, $k, m, p \in \mathbb{N}$. Temos que $b.(c.d) = a^k.(a^m.a^p) = a^k.a^{m+p} = a^{k+m+p}$ e $(b.c).d = (a^k.a^m).a^p = a^{k+m}.a^p = a^{k+m+p}$ E chegamos que $b.(c.d) = (b.c).d$.
3. a^0 é o elemento do qual temos, para todo $a^i \in G$: $a^i.a^0 = a^i = a^0.a^i$.
4. Seja $a^j \in G$ temos que a^{-j} é o elemento que $a^j.a^{-j} = a^0 = a^{-j}.a^j$.

Antes de prosseguir com o presente trabalho, vamos omitir o sinal ”.” no produto de dois elementos $a, b \in G$ e a partir daqui denotaremos por ab o produto de a por b .

3.1 Subgrupos

Em geral, não estaremos interessados em subconjuntos arbitrários de um grupo G , pois eles não refletem o fato de que G tem uma estrutura algébrica imposta sobre si.

Definição 3.4. Um subconjunto H de um grupo G é dito um subgrupo de G se, com relação ao produto em G , o próprio H forma um grupo.

Observação 3.5. Se H é um subgrupo de G e K é um subgrupo de H , então K é um subgrupo de G .

No objetivo de introduzir um critério para decidir se um subconjunto de um grupo é um subgrupo, seguem-se as duas proposições:

Proposição 3.6. Seja $H \neq \emptyset$ um subconjunto de um grupo G . Dizemos que H é um subgrupo de G se, e somente se,

1. $a, b \in H \Rightarrow ab \in H$,
2. $a \in H \Rightarrow a^{-1} \in H$.

Demonstração. Se H é um subgrupo de G , então é óbvio que os dois itens acima são válidos. Suponhamos agora que H seja um subconjunto de G para o qual valem os itens 1 e 2. A fim de estabelecer que H é um subgrupo, basta verificar que $e \in H$ e que a lei associativa vale para G . Se $a \in H$, por 2, $a^{-1} \in H$ e portanto, $e = aa^{-1} \in H$ ■

Proposição 3.7. Se H é um subconjunto finito não vazio de um grupo G e H é fechado com relação à multiplicação, então H é um subgrupo de G .

Demonstração. Como já se tem metade do que é necessário à luz da última proposição, então basta provar que $\forall a \in H, a^{-1} \in H$. Suponhamos então que $a \in H$; assim $a^2 = aa \in H, a^3 = aaa \in H, a^m \in H, \dots$ pois H é fechado. Assim, a coleção infinita de elementos $a, a^2, \dots, a^m, \dots$ está toda em H , que é um subconjunto finito de G . Assim, concluímos que há repetições nessa coleção de elementos,

isto é, para certos inteiros r e s , com $r > s > 0$, $a^r = a^s$ Pelo cancelamento em G , $a^{r-s} = e$ (sendo e o elemento neutro de H); como $r - s - 1 \geq 0$ e $a^{-1} = a^{r-s-1}$, pois $aa^{r-s-1} = a^{r-s} = e$. Assim $a^{-1} \in H$, e a proposição está demonstrada. ■

Exemplo 3.8. *Seja G o grupo dos inteiros com a operação de adição usual e H o subconjunto de todos os múltiplos de 5. H é um subgrupo de G pois, dado $a, b \in H$, temos que $a = 5k$ e $b = 5q$, $k, q \in \mathbb{Z}$.*

Definição 3.9. *Se H é um subgrupo de G , $a \in G$, então $Ha = \{Ha | h \in H\}$. Ha é denominada classe lateral à direita de H em G .*

De modo análogo é definida a classe lateral à esquerda de H em G .

3.1.1 Subgrupos Normais e Grupo Quociente

Definição 3.10. *Um subgrupo N de G é dito um subgrupo normal de G se para todo $g \in G$ e $n \in N$, $gng^{-1} \in N$.*

Proposição 3.11. *N é um subgrupo normal de $G \Leftrightarrow gNg^{-1} = N, \forall g \in G$.*

Demonstração. Se $gNg^{-1} = N \forall g \in G$, certamente $gNg^{-1} \subset N$, portanto N é normal em G . Agora suponhamos que N seja normal em G . Assim, por definição, para $g \in G$, $gNg^{-1} \subset N$ e $g^{-1}Ng = g^{-1}N(g^{-1})^{-1} \subset N$. Agora, como $g^{-1}Ng \subset N$, $N = eNe = (gg^{-1})N(g^{-1}g) = g(g^{-1}Ng)g^{-1} \subset gNg^{-1}$, portanto $N = gNg^{-1}$. ■

Indiquemos por G/N a coleção das classes laterais à direita de N em G (isto é, os elementos de G/N são certos subconjuntos de G) e usamos o produto de subconjuntos de G para obtermos o produto em G/N . Para esse produto, afirmamos:

1. $X, Y \in G/N \Rightarrow XY \in G/N$; pois $X = Na, Y = Nb$ para certos $a, b \in G$ e $XY = NaNb = Nab \in G/N$.
2. $X, Y, Z \in G/N$, então $X = Na, Y = Nb, Z = Nc$ com $a, b, c \in G$ e então $(XY)Z = (NaNb)Nc = N(ab)Nc = N(ab)c = Na(bc) = Na(Nbc) = Na(NbNc) = X(YZ)$. Assim o produto em G/N satisfaz a lei associativa.
3. Consideremos o elemento $N = Ne \in G/N$. Se $Na = X \in G/N$, $a \in G$, então $XN = NaNe = Nae = Na = X$, e analogamente $NX = X$. Consequentemente, N é um elemento unidade em G/N .
4. Suponhamos $X = Na \in G/N$ (onde $a \in G$); assim $Na^{-1} \in G/N$ e $NaN^{-1} = Naa^{-1} = Ne$. Analogamente, $N^{-1}Na = Ne$. Portanto, N^{-1} é o inverso de Na em G/N .

Mas um sistema que satisfaz 1, 2, 3 e 4 é exatamente o que denominamos um grupo. Isto é,

Teorema 3.12. *Se G é um grupo, N um subgrupo normal de G , então G/N também é um grupo. É denominado o grupo quociente ou o grupo fator G por N .*

4 Conclusão

É claro que a Teoria de Grupos não para por aqui. Todavia, a base introdutória do conteúdo de álgebra pura foi não só apresentada mas também salientada pois, por mais que pesquisadores dessa área avancem e se aprofundem nela, tais aspectos cruciais podem ser às vezes esquecido ou ignorado por um tempo até que precise ser lembrado.

Agradecimentos

Em primeira instância, gostaria de agradecer à minha família, em especial minha mãe, que foi a minha base para chegar até aqui. Além disso, sou muito grato ao Programa de Educação Tutorial que, por meio do financiamento da portaria SESu-MEC me ajudou a não só me introduzir, mas também me manter nesse âmbito acadêmico. Por último, mas não menos importante, direciono meus agradecimentos ao meu orientador Cícero, que abriu de maneira expressiva minha noção da realidade de um pesquisador matemático e da área de álgebra pura.

Referências

[1] FRALEIGH, John B. *A first course in abstract algebra*. Pearson Education India, 2003

[2] HERSTEIN, Israel N. *Topics in algebra*. John Wiley & Sons, 2006.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

O determinante de um nó The determinant of a knot

Victor Patrick Sena Barbosa Lima

Universidade Federal de Uberlândia (UFU), Faculdade de Matemática, Uberlândia, MG, Brasil
victorpsena2015@gmail.com

Taciana Oliveira Souza

Universidade Federal de Uberlândia (UFU), Faculdade de Matemática, Uberlândia, MG, Brasil
tacioli@ufu.br

Palavras-chave

Topologia .
Teoria dos Nós .
O determinante de um nó.
.
.

Resumo

Um nó clássico é um mergulho de uma esfera unidimensional S^1 em um ambiente tridimensional real, geralmente \mathbb{R}^3 . Na Teoria dos Nós estudam-se invariantes algébricos extraídos do complementar do mergulho. Neste trabalho, apresentamos o invariante *determinante de um nó*, que é um número associado ao nó, obtido por meio de um processo algorítmico.

Keywords

Topology.
Knot Theory.
The determinant of a knot.
.
.

Abstract

A classic knot is an embedding of an one-dimensional sphere S^1 in a real three-dimensional environment, usually \mathbb{R}^3 . In Knot Theory we study algebraic invariants extracted from the complement of the embedding. In this work we present the invariant *determinant of a knot*, which is a number associated to the knot, obtained through an algorithmic process.

1 Introdução

A Topologia é um dos mais importantes campos da matemática. Essa área estuda propriedades de objetos geométricos que permanecem inalteradas quando são aplicados movimentos contínuos no próprio objeto ou no ambiente no qual ele está contido.

Dentro da grande área da Topologia existe uma subárea chamada Teoria dos Nós. Essa teoria (a Clássica) pode ser compreendida como o estudo das curvas fechadas, sem autointerseções, em três dimensões. De modo intuitivo, um nó é construído torcendo e entrelaçando um fio e unindo suas extremidades como mostra a Figura 1.

Figura 1: Construindo um nó



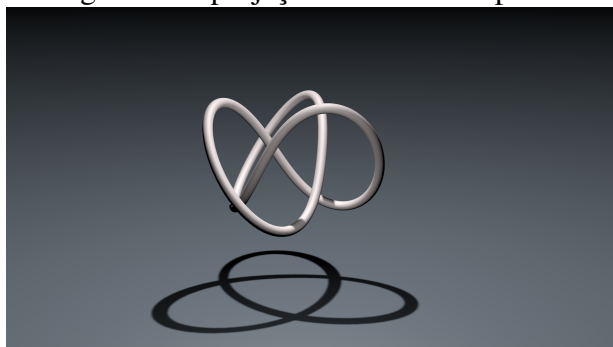
Fonte: Elaboração dos autores (2021).

O início do conceito matemático de “nó” aparece nos anos de 1830 com pesquisas de Carl Friedrich Gauss (1777-1855), [1]. Seu interesse, na época, era aplicar esse conceito na área de eletrodinâmica. Nesse mesmo século, décadas depois, outro grande pesquisador, William Thomson (Lord Kelvin) (1824-1907), se interessou pelo assunto, pois acreditava que os “nós” seriam a chave para a compreensão das substâncias químicas, [2], que, de acordo com suas crenças, seriam descritas pelas formas dos nós. Apesar de tal crença não condizer com a verdade, a Teoria dos Nós continuou a ser matematicamente estudada ao longo dos anos e aplicações foram sendo descobertas em outras áreas do conhecimento como: na Física, no estudo do DNA na Biologia Molecular e no estudo de estruturas tridimensionais de moléculas na Química.

Para estudarmos um nó, tendo em vista a dificuldade de desenharmos objetos tridimensionais, fixamos um plano do espaço e utilizamos sua projeção nesse plano. Sempre é possível considerar projeções que possuam apenas cruzamentos duplos, como ilustra a Figura 2. Nessa projeção, desenhamos pequenas interrupções próximas a cada cruzamento para indicarmos que um arco passa sobre o outro. Assim, consideramos o que chamamos de diagrama de um nó. Como ilustração temos, na Figura 3, o diagrama do nó presente na Figura 2.

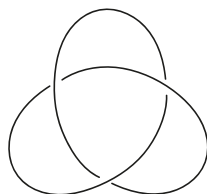
Para a Teoria dos Nós, quando um nó pode ser transformado continuamente em outro que, a princípio, aparenta ser diferente, diz-se que os dois nós são equivalentes. Mas se para registrarmos um nó utilizarmos um diagrama (uma projeção), é necessário investigar se dois diagramas que, a princípio, são aparentemente diferentes, podem representar o mesmo nó. Esse problema foi resolvido por Kurt Reidemeister (1883-1971) em 1920. Ele propôs que se considerasse três movimentos que poderiam ser

Figura 2: A projeção de um nó no plano.



Fonte: Elaboração dos autores (2021).

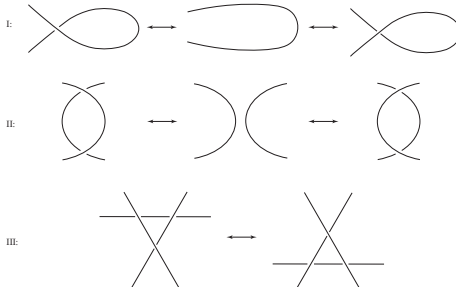
Figura 3: O diagrama de um nó.



Fonte: Elaboração dos autores (2021).

realizados em diagramas de nós. Hoje esses movimentos são chamados *movimentos de Reidemeister*. Eles estão ilustrados na Figura 4.

Figura 4: Os três movimentos de Reidemeister.



Fonte: Elaboração dos autores (2021).

Quando um diagrama de um nó pode ser transformado em outro diagrama mediante aplicação de uma sequência finita de movimentos de Reidemeister, diz-se que esses diagramas são equivalentes. Com essa consideração, Reidemeister provou que “dois nós são equivalentes se, e somente se, os respectivos diagramas são equivalentes”. Deste modo, pode-se estudar equivalência dos nós estudando equivalência dos diagramas dos nós. Apesar do excepcional resultado de Reidemeister, é geralmente difícil determinar quando um nó é equivalente a outro. Uma forma de transpor essa dificuldade é associar ao diagrama de um nó um objeto algébrico ou aritmético que permaneça o mesmo quando se efetue uma sequência finita de movimentos de Reidemeister. Assim, esse objeto pode auxiliar na detecção de nós que não são equivalentes. Chamamos tais objetos matemáticos de invariantes do nó.

Neste trabalho, apresentamos o invariante *determinate de um nó*, que é um número associado ao nó, obtido por meio de um processo algorítmico.

2 A teoria dos nós clássica

O que chamamos de nó (clássico) em \mathbb{R}^3 é a imagem de uma aplicação contínua injetiva $f : S^1 \rightarrow \mathbb{R}^3$, (chamamos isso de mergulho). Por padrão, identificamos um nó com a imagem de uma aplicação contínua injetiva e o denotamos $K = f(S^1)$. Assim quando consideramos, por exemplo, dois nós, estamos considerando dois mergulhos $f(S^1) \subset \mathbb{R}^3$ e $g(S^1) \subset \mathbb{R}^3$. Então, a diferença entre os diversos nós não está no espaço topológico que é a imagem do mergulho, está no espaço que circunda a imagem do mergulho, está no complementar $\mathbb{R}^3 \setminus f(S^1)$. Nossa noção intuitiva de deformação de um nó K_1 em outro nó K_2 é, na verdade, a noção de um movimento contínuo no espaço ambiente \mathbb{R}^3 que leva continuamente o subespaço topológico K_1 sobre o subespaço topológico K_2 . Essa noção é a associada à definição de isotopia ambiente.

Definição 2.1. *Considere um espaço topológico X , uma isotopia ambiente H em X é uma aplicação contínua $H : X \times [0, 1] \rightarrow X$ tal que $H_0 : X \times \{0\} \rightarrow X$ é a aplicação identidade e $H_t : X \times \{t\} \rightarrow X$ é um homeomorfismo para todo $t \in [0, 1]$.*

Definição 2.2. *Dizemos que dois nós K_1 e K_2 têm o mesmo tipo (são equivalentes) se existe uma isotopia ambiente $H : \mathbb{R}^3 \times [0, 1] \rightarrow \mathbb{R}^3$ tal que $H_1(K_1) = K_2$.*

2.1 O determinate de um nó

O determinante de um nó é um número real não negativo associado ao nó. Este número é um invariante do nó, isto é, quaisquer dois nós equivalentes possuem o mesmo determinante. Desse modo, dois nós com determinantes diferentes não são equivalentes.

O determinante de um nó é calculado de forma algorítmica, como segue:

1. A cada arco do diagrama do nó associa-se uma variável. Em seguida, a cada cruzamento associa-se uma equação da forma: $a + b - 2d = 0$. Veja a Figura 5

Figura 5: Cruzamento.



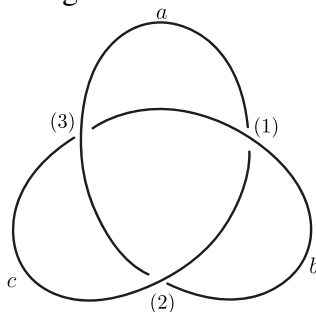
Fonte: Elaboração dos autores (2021).

2. Coloca-se uma variável qualquer igual a zero.
3. Descarta-se uma equação qualquer.

4. Assim, obtém-se um sistema de $n - 1$ equações e $n - 1$ variáveis, sendo n o número de cruzamentos do diagrama do nó. Calcula-se o determinante desse sistema e tomando seu módulo obtém-se o determinante do nó.

Exemplo 2.3. Para exemplificação, faremos o cálculo do determinante do nó Trevo, cujo diagrama está ilustrado na Figura 6.

Figura 6: Nó Trevo.



Fonte: Elaboração dos autores (2021).

Foram associadas aos arcos do diagrama as variáveis a , b e c . Assim, obtemos as seguintes equações:

$$(1) \quad a + c - 2b = 0$$

$$(2) \quad a + b - 2c = 0$$

$$(3) \quad c + b - 2a = 0$$

Colocando $b = 0$ e descartando a primeira equação, obtemos o seguinte sistema de duas equações e duas variáveis:

$$\begin{cases} a - 2c = 0 \\ -2a + c = 0 \end{cases}$$

cujo determinante é

$$\begin{vmatrix} 1 & -2 \\ -2 & 1 \end{vmatrix} = -3.$$

Assim, o determinante do nó Trevo é $D = |-3| = 3$.

3 Procedimentos Metodológicos

O método utilizado no desenvolvimento deste trabalho consistiu no estudo de materiais relacionados ao tema proposto. Para compreensão da construção e efetividade do invariante *determinante de um nó* e fundamentação teórica deste trabalho foram necessários estudos de Topologia Geral e Teoria dos Nós, realizados em [3], [4], [5] e [6].

4 Considerações finais

A Teoria dos Nós é intuitiva, de fácil percepção geométrica e possui diversas aplicações. Destacamos suas aplicações em outras áreas do conhecimento tais como Biologia, Física e Química, o que torna seu estudo ainda mais relevante. Esperamos que este trabalho possa contribuir para a difusão do tema entre estudantes e professores da Matemática e de outras áreas.

Agradecimentos

Agradecemos o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, por meio do Programa Institucional de Bolsas de Iniciação Científica PIBIC-UFU.

Referências

- [1] Gauss, K. F. Zur mathematischen theorie der elektro-dynamischen wirkungen. Werke Konigl. Gessell. Wiss. Gottingen, 1833
- [2] Enciclopedia Britânica. Disponível em: <<https://www.britannica.com/science/knot-theory>>. Acesso em: 20 abr. 2020.
- [3] Lima, E. L. **Elementos de Topologia Geral**. 3ed, SBM, Textos Universitários: Rio de Janeiro, 2014.
- [4] Munsignatti, M. J. **O Polinômio de Alexander e o Determinante de um Nó**. 2013, 51 f. Dissertação (Mestrado Profissional) - Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, 2013.
- [5] Murasugi, K. **Knot theory and its applications**. Birkhauser Boston, 1996.
- [6] Rolfsen, D. **Knots and Links**. Publish or Perish: Berkeley, CA, 1976.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Classificação dos grupos de ordem ≤ 8

Classification of groups of order ≤ 8

Victor Rodrigues Silva

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

victor.silval@ufu.com

Adriana Rodrigues da Silva

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil

adrianafamat@ufu.com

Palavras-chave

Grupos.
Classificação.
Ordem.
Isomorfismos.

Resumo

Neste trabalho, serão introduzidos alguns resultados gerais da Teoria de Grupos, como: a relação entre a ordem de um grupo e a ordem de seus subgrupos, a identificação de um grupo como grupo cíclico, o isomorfismo de grupos, a relação entre a ordem dos elementos de um grupo e sua caracterização, e propriedades de grupos gerados por dois elementos. Com esses resultados, conseguiremos fazer a classificação dos grupos finitos de ordem ≤ 8 .

Keywords

Groups.
Classification.
Order.
Isomorphisms.

Abstract

In this text, some general results of the Group Theory will be introduced, such as: the relation between the order of a group and the order of its subgroups, the identification of a group as a cyclic group, the isomorphism of groups, the relation between the order of the elements of a group and its characterization, and properties of groups generated by two elements. With these results, we will be able to make a classification of the finite groups of order ≤ 8 .

1 Introdução

No estudo de Teoria de Grupos é importante conhecer a estrutura dos grupos e suas propriedades. Os grupos finitos gerados por um elemento são cíclicos e podem ser facilmente classificados. Porém, classificar os grupos finitos gerados por dois ou três elementos pode ser extremamente complicado. Neste trabalho, aplicando alguns resultados preliminares, conseguiremos classificar os grupos de ordem ≤ 8 , a menos de isomorfismos.

2 Resultados Gerais

Teorema 2.1. (Teorema de Lagrange). *Sejam G um grupo finito e H um subgrupo de G . Então $|G| = |H|(G : H)$. Em particular, a ordem e o índice de H dividem a ordem de G .*

Demonstração. [1], página 119. ■

Corolário 2.2. *Seja G um grupo de ordem p , onde p é um número primo. Então G é cíclico.*

Demonstração. [2], página 134. ■

Proposição 2.3. *Seja G um grupo cíclico finito de ordem n . Então $G \simeq \mathbb{Z}_n$.*

Demonstração. [1], página 112. ■

Proposição 2.4. *Seja G um grupo e seja $\alpha \in G$. Se $\mathcal{O}(\alpha) = 2, \forall \alpha \neq e$, então G é um grupo abeliano.*

Demonstração. Sejam $a, b \in G$. Então temos que $ab \in G$. Logo, $(ab)(ab) = e \Rightarrow (ab)(ab)b = eb \Rightarrow (ab)a(bb) = b \Rightarrow (ab)a = b \Rightarrow (ab)aa = ba \Rightarrow ab = ba$. Portanto G é abeliano. ■

Teorema 2.5. *Sejam n, m, s, u inteiros ≥ 0 .*

(a) *Se G é um grupo de ordem nm que possui elementos a, b tais que*

$$G = \langle a, b \rangle; \quad a^n = e; \quad b^m = a^u; \quad ba = a^s b. \quad (1)$$

Então, $s^m \equiv 1 \pmod n$ e $u(s-1) \equiv 0 \pmod n$.

(b) *Quando existir um grupo de ordem nm satisfazendo as condições do sistema (1), ele é único a menos de isomorfismos.*

Demonstração. [3], página 185. ■

3 Classificação

3.1 Grupo de ordem 1

O único grupo de ordem 1 é o grupo $\{e\}$, onde e é o elemento neutro de G .

3.2 Grupos de ordem p , onde p é um número primo

Seja G um grupo de ordem p , com p primo. Temos, pelo Corolário 2.2, que G é cíclico. Assim, pela Proposição 2.3, G é um grupo cíclico com p elementos e portanto temos $G \simeq \mathbb{Z}_p$.

Portanto temos que, a menos de isomorfismos, os únicos grupos de ordem 2, 3, 5 e 7 são, respectivamente: o grupo \mathbb{Z}_2 , o grupo \mathbb{Z}_3 , o grupo \mathbb{Z}_5 e o grupo \mathbb{Z}_7 .

3.3 Grupos de ordem 4

Temos que os grupos \mathbb{Z}_4 e $\mathbb{Z}_2 \times \mathbb{Z}_2$ são grupos de ordem 4. Eles não são isomorfos pois \mathbb{Z}_4 possui elementos de ordem 4 enquanto $\mathbb{Z}_2 \times \mathbb{Z}_2$ não possui tais elementos. Vamos mostrar que, a menos de isomorfismos, estes dois grupos são os únicos de ordem 4.

Seja G um grupo de ordem 4. Se G possui um elemento de ordem 4, então $G \simeq \mathbb{Z}_4$. Se G não possui um elemento de ordem 4, pelo Teorema 2.1, todos os seus elementos $\neq e$ são de ordem 2 e, pela Proposição 2.4, temos que G é um grupo abeliano. Escrevendo $G = \{e, a, b, c\}$, vamos procurar sua tabela de multiplicação.

Primeiramente, vamos encontrar o resultado da multiplicação ab . Temos que $ab \neq e$, pois caso contrário, teríamos $a = b^{-1}$, o que é absurdo, pois b sendo de ordem 2 temos $b^{-1} = b$. Da mesma forma, temos que $ab \neq a$ pois caso contrário, teríamos $b = e$, o que é absurdo. E por último, temos $ab \neq b$, pois caso contrário, teríamos $a = e$, o que é absurdo.

Portanto, $ab = c$ e também $ba = c$, pois o grupo é abeliano. Similarmente, temos $ac = b = ca$ e $bc = a = cb$. Agora vamos verificar que a função abaixo é um isomorfismo:

$$\varphi: \mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow G; \quad (\bar{0}, \bar{0}) \mapsto e; \quad (\bar{1}, \bar{0}) \mapsto a; \quad (\bar{0}, \bar{1}) \mapsto b; \quad (\bar{1}, \bar{1}) \mapsto c.$$

A função é injetora pois elementos distintos do domínio são levados em elementos distintos do contra-domínio. Ela é também sobrejetora pois todo elemento do contra-domínio é imagem de alguém do domínio. Portanto a função é bijetora. Falta verificar que a função é um homomorfismo.

$$\underbrace{\varphi(\bar{1}, \bar{0}) \cdot \varphi(\bar{0}, \bar{1})}_{ab} = \underbrace{\varphi(\bar{1}, \bar{1})}_c; \quad \underbrace{\varphi(\bar{1}, \bar{0}) \cdot \varphi(\bar{1}, \bar{1})}_{ac} = \underbrace{\varphi(\bar{0}, \bar{1})}_b; \quad \underbrace{\varphi(\bar{0}, \bar{1}) \cdot \varphi(\bar{1}, \bar{1})}_{bc} = \underbrace{\varphi(\bar{1}, \bar{0})}_a.$$

Os outros produtos seguem do fato de G ser abeliano e de que $\varphi(\bar{0}, \bar{0}) = e$.

Assim, a menos de isomorfismos, \mathbb{Z}_4 e $\mathbb{Z}_2 \times \mathbb{Z}_2$ são os únicos grupos de ordem 4.

3.4 Grupos de ordem $2p$, com p primo ímpar

Seja G um grupo de ordem $2p$, com p primo ímpar. Pelo Teorema 2.1, os elementos $\alpha \in G, \alpha \neq e$, tem ordem 2, p ou $2p$.

Se G tiver um elemento de ordem $2p$, então ele é cíclico e isomorfo a \mathbb{Z}_{2p} .

Se G não for cíclico, então ele possui um elemento de ordem p . De fato, se $\mathcal{O}(\alpha) = 2, \forall \alpha \in G, \alpha \neq e$, temos, pela Proposição 2.4, que G é abeliano. Tomemos $\alpha, \beta \in G \setminus \{e\}$, onde $\beta \in G \setminus \langle \alpha \rangle$, então $H = \{e, \alpha, \beta, \alpha\beta\}$ seria um subgrupo de G com ordem 4. Mas $4 \nmid 2p$. Absurdo!

Se $\beta \in G$ tal que $\beta \in G \setminus \langle \alpha \rangle$ então $\mathcal{O}(\beta) = 2$. De fato, se $\mathcal{O}(\beta) = p$, então $e, \alpha, \alpha^2, \dots, \alpha^{p-1}, \beta, \beta^2, \dots, \beta^{p-1}$ são $2p - 1$ elementos distintos de G . Mas $\alpha\beta$ e $\alpha^2\beta$ são elementos de G distintos dos listados acima, então G teria pelo menos $2p + 1$ elementos. Absurdo!

Assim, $\exists \alpha \in G$ tal que $\mathcal{O}(\alpha) = p$ e $\exists \beta \in G$ tal que $\mathcal{O}(\beta) = 2$.

Logo, $G = \{e, \alpha, \alpha^2, \dots, \alpha^{p-1}, \beta, \alpha\beta, \alpha^2\beta, \dots, \alpha^{p-1}\beta\}$.

O elemento $\beta\alpha \notin \{e, \alpha, \alpha^2, \dots, \alpha^{p-1}, \beta\}$, pois $\beta \notin \langle \alpha \rangle$. Logo, $\beta\alpha \in \{\alpha^2\beta, \dots, \alpha^{p-1}\beta\}$. Portanto, G é um grupo de ordem $2p$ tal que

$$G = \langle \alpha, \beta \rangle; \quad \alpha^p = e; \quad \beta^2 = \alpha^s; \quad \beta\alpha = \alpha^s\beta.$$

Pelo item a) do Teorema 2.5, temos que $s^2 \equiv 1 \pmod p$. Logo $s = 1$ ou $s = p - 1$. Então temos duas possibilidades para G :

$$|G| = 2p; \quad G = \langle \alpha, \beta \rangle; \quad \alpha^p = e; \quad \beta^2 = e; \quad \beta\alpha = \alpha\beta. \quad (2)$$

$$|G| = 2p; \quad G = \langle \alpha, \beta \rangle; \quad \alpha^p = e; \quad \beta^2 = e; \quad \beta\alpha = \alpha^{p-1}\beta. \quad (3)$$

Pelo item b) do Teorema 2.5, existe um único grupo para cada uma das possibilidades, a menos de isomorfismos. No caso (2) o grupo é abeliano. Logo, é isomorfo a \mathbb{Z}_{2p} , com o isomorfismo que leva $\alpha \mapsto \bar{2}$ e $\beta \mapsto \bar{p}$. O grupo D_n das simetrias do polígono regular de n lados satisfaz as condições de (3) com $n = p$, logo no caso (3) temos $G \simeq D_p$.

Portanto, os grupos de ordem 6 são iguais ou isomorfos a \mathbb{Z}_6 ou a D_3 , onde D_3 é o grupo das simetrias do triângulo equilátero.

3.5 Grupos de ordem 8

Temos que $\mathbb{Z}_8, \mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ e D_4 são 4 grupos de ordem 8. Eles não são isomorfismos entre si, pois cada um deles possui características que os outros não têm. \mathbb{Z}_8 possui elementos de ordem 8, $\mathbb{Z}_4 \times \mathbb{Z}_2$ possui 4 elementos de ordem 4, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ possui todos os elementos $\neq e$ com ordem 2 e D_4 possui apenas 2 elementos de ordem 4.

Além desses 4 grupos citados, o grupo Q_8 dos quatérnios também é um grupo de ordem 8, cujos elementos são $\left\{ \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \pm \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \right\}$.

Tomando $A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ temos que:

$$|Q_3| = 8; \quad Q_3 = \langle A, B \rangle; \quad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad B^2 = A^2; \quad BA = A^3B.$$

Pela parte b) do Teorema 2.5, esse grupo é único a menos de isomorfismos. Note que Q_3 não é isomorfo a \mathbb{Z}_8 , $\mathbb{Z}_4 \times \mathbb{Z}_2$ ou $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, pois estes são abelianos e Q_3 não é. Além disso, Q_3 também não é isomorfo a D_4 , pois D_4 possui 5 elementos de ordem 2 (a saber as suas 4 reflexões e a rotação π), enquanto Q_3 possui somente um elemento de ordem 2 (a saber o elemento -id).

Assim, temos que \mathbb{Z}_8 , $\mathbb{Z}_4 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, D_4 e Q_3 são 5 grupos de ordem 8 não isomorfos entre si. Resta verificar que eles são os únicos, a menos de isomorfismos.

Seja G um grupo de ordem 8. Pelo Teorema 2.1, as possíveis ordens dos elementos de $G \setminus \{e\}$ são 2, 4 e 8. Temos 2 casos:

- **Caso 1: G possui um elemento de ordem 8.** Seja $\gamma \in G$ um elemento de ordem 8, então $G = \langle \gamma \rangle$ e $G \simeq \mathbb{Z}_8$.
- **Caso 2: G não possui elemento de ordem 8.** Então as possíveis ordens dos elementos $\neq e$ são 2 e 4. Vamos dividir em 2 subcasos:
 - **Caso 2.1: G não possui nenhum elemento de ordem 4.** Nesse caso, todos os elementos $\neq e$ de G são de ordem 2 e, pela Proposição 2.4, o grupo G é abeliano. Seja $a \neq e$. Como $\mathcal{O}(a) = 2$, temos que $H = \{e, a\}$ é um subgrupo de G . Tome $b \in G \setminus H$. Então $K = \{e, a, b, ab\}$ é um subgrupo de G . Tome $c \in G \setminus K$. Temos: $G = \{e, a, b, c, ab, ac, bc, abc\} = \{a^i b^j c^k \mid i, j, k \in \{0, 1\}\}$. A função $\varphi: \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow G$ que leva $(\bar{i}, \bar{j}, \bar{k}) \mapsto a^i b^j c^k$ é um isomorfismo. Portanto, $G \simeq \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.
 - **Caso 2.2: G possui um elemento de ordem 4:** Seja $a \in G$ um elemento de ordem 4 e seja $H = \langle a \rangle$. Tome agora $b \in G \setminus H$. Então, temos que $G = \langle a, b \rangle = \{e, a, a^2, a^3, b, ab, a^2b, a^3b\}$. Temos também que $b^2 \in H$, pois $b^2 \notin \{b, ab, a^2b, a^3b\}$. Também temos que $ba \notin \{e, a, a^2, a^3, b\}$. Provamos assim que:

$$|G| = 8; \quad G = \langle a, b \rangle; \quad a^4 = e; \quad b^2 = a^u; \quad ba = a^s b; \quad \text{para algum } u \in \{0, 1, 2, 3\} \text{ e } s \in \{1, 2, 3\}.$$

Vejam agora quais são as possibilidades para $u, s \in \{0, 1, 2, 3\}$.

Como $\mathcal{O}(bab^{-1}) = \mathcal{O}(a) = 4$, logo $s = 1$ ou 3 . Além disso, $b^2 \notin \{a, a^3\}$. Logo $u = 0$ ou 2 e $s = 1$ ou 3 . Portanto, temos as seguintes possibilidades:

$$|G| = 8; \quad G = \langle a, b \rangle; \quad a^4 = e; \quad b^2 = e; \quad ba = ab. \tag{4}$$

$$|G| = 8; G = \langle a, b \rangle; a^4 = e; b^2 = e; ba = a^3b. \quad (5)$$

$$|G| = 8; G = \langle a, b \rangle; a^4 = e; b^2 = a^2; ba = ab. \quad (6)$$

$$|G| = 8; G = \langle a, b \rangle; a^4 = e; b^2 = a^2; ba = a^3b. \quad (7)$$

Pela parte b) do Teorema 2.5, em cada um dos 4 casos temos no máximo um grupo, a menos de isomorfismos e esses grupos de fato existem. No caso (4), tome $G = \mathbb{Z}_4 \times \mathbb{Z}_2$, com $a = (\bar{1}, \bar{0})$ e $b = (\bar{0}, \bar{1})$. No caso (5), tome $G = D_4$, com $a =$ rotação $(\frac{\pi}{2})$ e $b =$ reflexão. No caso (6), tome $G = \mathbb{Z}_4 \times \mathbb{Z}_2$, com $a = (\bar{1}, \bar{0})$ e $b = (\bar{1}, \bar{1})$. No caso (7), tome $G = Q_8$, com $a = A$ e $b = B$, como definidos anteriormente.

Portanto, a menos de isomorfismos, existem exatamente cinco grupos de ordem 8. São eles: $\mathbb{Z}_8, \mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, D_4$ e Q_8 .

4 Conclusão

Através da aplicação de resultados advindos da Teoria de Grupos, concluímos que, a menos de isomorfismos, os grupos de ordem ≤ 8 são: $\{e\}, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_4, \mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{Z}_5, \mathbb{Z}_6, D_3, \mathbb{Z}_7, \mathbb{Z}_8, \mathbb{Z}_4 \times \mathbb{Z}_2, \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, D_4$ e Q_8 .

Agradecimentos

Agradeço ao PET Matemática da Universidade Federal de Uberlândia que incentivou a realização desse estudo e a Secretaria de Educação Superior do Ministério da Educação (SESu-MEC) que, em conjunto com o Fundo Nacional para o Desenvolvimento da Educação (FNDE), apoiou financeiramente essa pesquisa.

Referências

- [1] DOMINGUES, H. H.; IEZZI, G. **Álgebra Moderna**. 2ª Edição. São Paulo: Atual, 1982.
- [2] GONÇALVES, A. **Introdução à Álgebra**. Projeto Euclides. Rio de Janeiro: IMPA - SBM, 1979.
- [3] GARCIA, A.; LEQUAIN, Y. **Elementos de Álgebra**. Projeto Euclides. Rio de Janeiro: IMPA - SBM, 2002.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

24 A 26 DE MAIO DE 2021

Detecção de danos em vigas de alumínio via modelagem *fuzzy* Detection of damage in aluminum beams via fuzzy modeling

Fernando Augusto Freitas

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
fernandoaugusto.mat@gmail.com

Rosana Sueli da Motta Jafelice

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
rmotta@ufu.br

José Waldemar da Silva

UFU, Faculdade de Matemática, Uberlândia, Minas Gerais, Brasil
zewaldemar@ufu.br

Carlos Alberto Gallo

UFU, Faculdade de Engenharia Elétrica, Uberlândia, Minas Gerais, Brasil
gallo@ufu.br

Julio Endress Ramos

PETROBRAS, CENPES, Rio de Janeiro, Rio de Janeiro, Brasil
julio.ramos@petrobras.com.br

Resumo

Palavras-chave

Impedância Eletromecânica.
Conjuntos Fuzzy.
ANFIS.
SHM.
Índice de dano.

Um dos grandes desafios do processo de *Structural Health Monitoring* (SHM) é determinar estratégias para avaliar o nível de danos em estruturas e equipamentos de engenharia no decorrer do tempo. O objetivo deste trabalho é modelar o nível de danos de uma viga de alumínio, a partir de um experimento realizado em uma câmara climática de controle de temperatura e umidade. Neste estudo são determinados Sistemas Baseados em Regras Fuzzy (SBRF) através do *Adaptive-Network-Based Fuzzy Inference System* (ANFIS), com a variável de saída sendo o nível de dano. As variáveis de entrada são obtidas por dois índices de danos, estabelecidos pelas assinaturas da impedância eletromecânica. A média das porcentagens de acertos dos SBRFs é 95%. Os resultados podem indicar possíveis entradas para SBRFs com o objetivo de identificar os níveis de danos, quando não se tem conhecimento destes valores de saída do sistema.

Abstract

One of the great challenges of the Structural Health Monitoring (SHM) process is to determine strategies to assess the level of damage to structures and engineering equipment over time. The objective of this work is to model the damage level of an aluminum beam, based on an experiment carried out in a climatic chamber to temperature and humidity control. In this study, Fuzzy Rule-Based System (FRBS) are determined through the Adaptive-Network-Based Fuzzy Inference System (ANFIS), with the output variable being the level of damage. The input variables are obtained by two damage indices, established by the electromechanical impedance signatures. The average percentage of correct answers for the FRBS is 95%. The results can indicate possible inputs for FRBS in order to identify the damage levels, when these system output values are not known.

Keywords

Electromechanical Impedance.
Fuzzy Sets.
ANFIS.
SHM.
Damage index.

1 Introdução

O estudo da impedância eletromecânica (Ω) permite detectar deformações em estruturas metálicas a partir de sensores acoplados na estrutura, os Titanato Zirconato de Chumbo (PZT). Isso é possível uma vez que as cerâmicas de PZT são excitadas em altas faixas de frequência, permitindo uma alta sensibilidade a mudanças estruturais, detectadas através de alterações nas curvas de impedância obtidas. A presença ou ausência de dano, em geral, é verificada a partir de índices de dano os quais são calculados utilizando assinaturas de impedância obtidas na estrutura em um estado de referência, ou *baseline*, e no estado investigado [1].

Desta forma, uma dificuldade encontrada neste método é conseguir quantificar danos a partir destas métricas coletadas em suas respectivas temperaturas, pois não se tem informação se o comportamento do incremento de dano e valores de métrica estão ou não diretamente ligados. Uma alternativa para este estudo é a construção de Sistemas Baseados em Regras Fuzzy (SBRFs), devido à possibilidade de manipulação de informações incertas ou a capacidade de representar matematicamente a imprecisão. Assim, o objetivo deste trabalho é determinar SBRFs por meio do *Adaptive-Network-Based Fuzzy Inference System* (ANFIS) que tornem possível a identificação do nível de dano em vigas de alumínio, em um experimento com controle de temperatura e umidade.

2 Fundamentação teórica

2.1 Teoria dos conjuntos *Fuzzy*

Definição 2.1. Um conjunto *fuzzy* universo U é caracterizado por uma função de pertinência: $\mu_A : U \rightarrow [0, 1]$. Esta função é chamada de pertinência do conjunto *fuzzy*. O valor $\mu_A(x)$ da função é interpretado como o grau com que o elemento x está contido no conjunto *fuzzy* A , ou seja, $\mu_A(x) = 0$ indica a não pertinência do x à A , e o $\mu_A(x) = 1$ indica pertinência total ao conjunto [2].

Um conjunto *fuzzy* também pode ser representado por um conjunto clássico de pares ordenados: $A = \{(x, \mu_A(x)) | x \in U\}$.

As funções de pertinência podem assumir diferentes formas em que são escolhidas de acordo com o problema em questão, neste trabalho foram utilizadas as trapezoidais.

Um SBRF é composto por quatro componentes: fuzzificação, base de regras, método de inferência *fuzzy* e defuzzificação [3]. Estas componentes são conectadas conforme a Figura 1(a). Neste trabalho, o método de inferência utilizado é o de Takagi-Sugeno, que é representado por meio de duas regras (1) e (2):

$$\text{Se } x \text{ é } A_1 \text{ e } y \text{ é } B_1 \text{ então } z_1 = p_1x + q_1y + r_1; \quad (1)$$

$$\text{Se } x \text{ é } A_2 \text{ e } y \text{ é } B_2 \text{ então } z_2 = p_2x + q_2y + r_2. \quad (2)$$

em que p_1, q_1, r_1, p_2, q_2 e r_2 são números reais. Na Figura 1(b) está representado o método de inferência de Takagi-Sugeno.

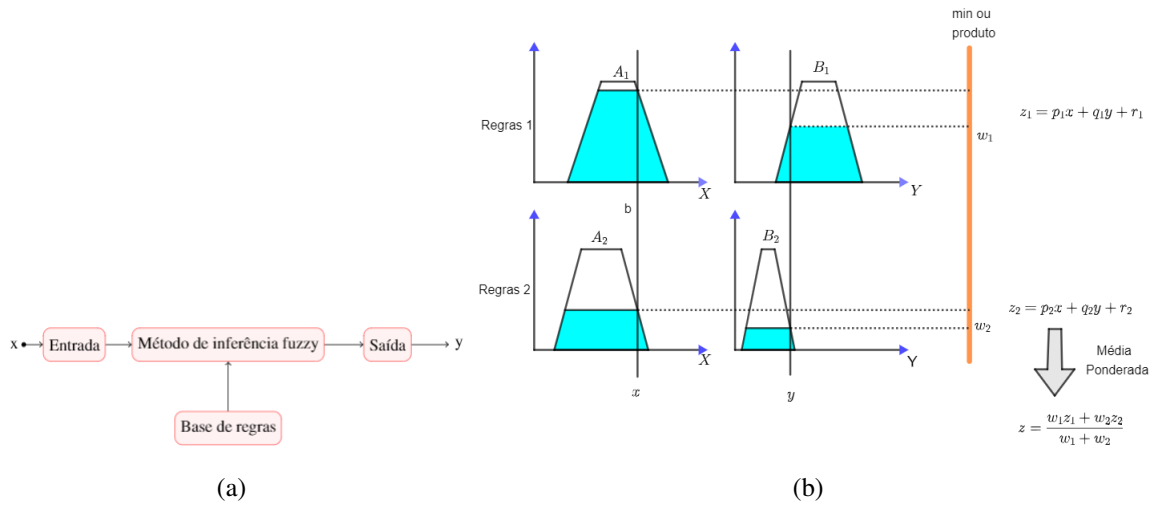


Figura 1: Arquitetura do SBRF (a) e método de inferência fuzzy de Takagi-Sugeno (b).

2.2 ANFIS

A ANFIS [4] é uma rede neural artificial que combina o método dos mínimos quadrados com o método do gradiente descendente para ajustar os parâmetros das funções de pertinência e os coeficientes das funções de cada regra, com o objetivo de minimizar o erro entre o observado e o esperado. Sua estrutura é conforme Figura 2(a) em que A_1, A_2, B_1 e B_2 são conjuntos fuzzy.

2.3 Índice de dano

Os índices de dano, ou métricas, que foram utilizados são o RMSD1 (desvio da raiz média quadrática) e o CCD (desvio do coeficiente de correlação)[1]. O índice RMSD1 pode ser calculado através da equação (3), enquanto que o CCD é computado por (4) que envolve também a diferença entre um e o coeficiente de correlação (CC), dado pela equação (5), conforme:

$$RMSD1 = \sqrt{\sum_{j=1}^n \frac{(Z_{B,j} - Z_{I,j})^2}{Z_{I,j}^2}}, \quad (3)$$

$$CCD = 1 - CC, \quad (4)$$

$$CC = \frac{1}{n} \sum_{j=1}^n \left\{ \frac{[Z_{B,j} - \bar{Z}_B][Z_{I,j} - \bar{Z}_I]}{S_{Z_B}S_{Z_I}} \right\}, \quad (5)$$

em que $Z_{B,j}$ é a parte real da assinatura *baseline* na frequência j e $Z_{I,j}$ é a parte real da assinatura investigada na frequência j , ambas considerando n pontos de frequência, \bar{Z}_B e \bar{Z}_I são as médias das assinaturas *baseline* e investigada, respectivamente, na faixa de frequência selecionada e S_{Z_B} é o

desvio padrão da assinatura *baseline* e S_{Z_I} é o desvio padrão da assinatura investigada, considerando todos os pontos de frequência de cada uma.

3 Experimento

Os dados experimentais utilizados para a aplicação da metodologia foram obtidos a partir de oito corpos de prova em um experimento em câmara climática, com controle de temperatura e umidade, instalada no Laboratório de Mecânica de Estruturas (LMEst) da Faculdade de Engenharia Mecânica da Universidade Federal de Uberlândia.

Os corpos de prova foram vigas de alumínio de 500 mm de comprimento, 38 mm de largura e 3,2 mm de espessura e em cada uma destas estruturas foi colada uma pastilha PZT de 1 mm de espessura e 20 mm de diâmetro a 100 mm da extremidade. Os processos de inserção de falhas adotado foi o de adição de massas coladas a 380 mm do centro da pastilha PZT, na extremidade oposta da viga. Estes corpos de prova podem ser vistos na Figura 2(b).

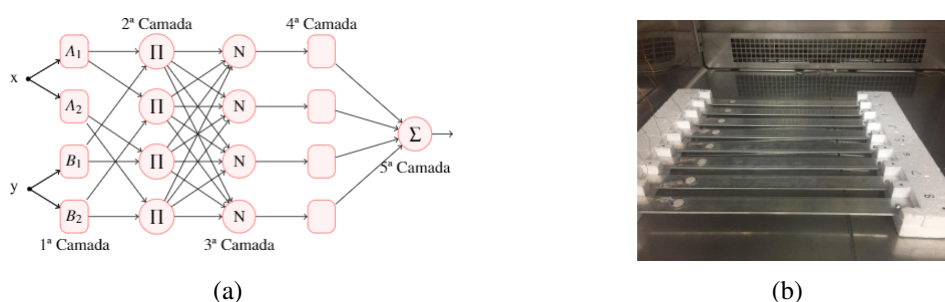


Figura 2: Arquitetura da ANFIS (a) e no experimento foram utilizadas 8 vigas com suas identificações, e localização das massas acrescentadas (b).

O experimento foi realizado nas temperaturas $-10, 0, 10$ e $20^{\circ}C$. Para os casos de danos por adição de massas foram considerados cinco níveis de falhas, sendo o primeiro sem falha, *baseline*, e outros quatro níveis progressivos. Para uma pequena variação controlada no experimento, foram realizadas as adições de massas nos corpos de prova de acordo com a Tabela 1.

	<i>Baseline</i>	Dano 1 (g)	Dano 2 (g)	Dano 3 (g)	Dano 4 (g)
Viga 1	sem dano	0,6	1,1	1,6	2,2
Viga 2	sem dano	0,6	1,1	1,6	2,2
Viga 3	sem dano	0,6	1,1	1,6	2,8
Viga 4	sem dano	0,6	1,1	1,6	2,7
Viga 5	sem dano	0,6	1,1	1,6	2,2
Viga 6	sem dano	0,6	1,1	1,6	2,2
Viga 7	sem dano	0,6	1,1	1,6	2,2
Viga 8	sem dano	0,6	1,1	1,6	2,2

Tabela 1: Massas referentes a cada adição de dano na estrutura.

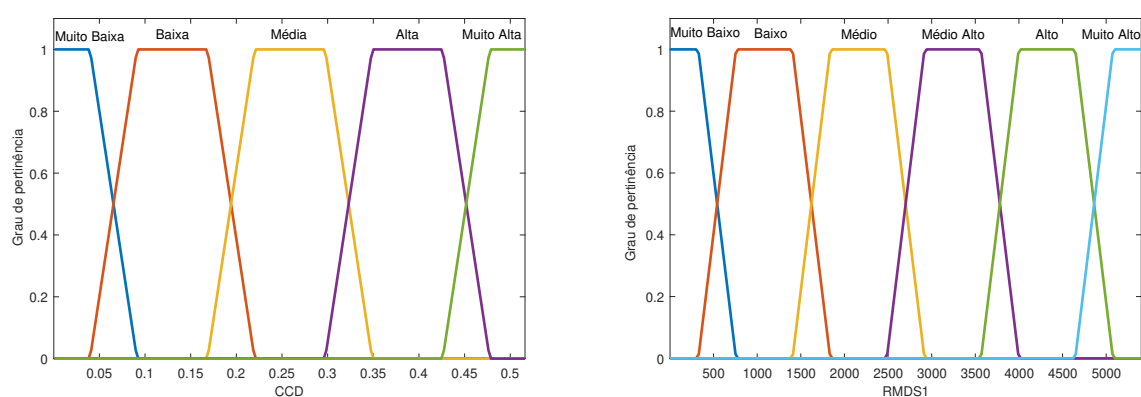
Em cada valor de temperatura foram coletadas 20 assinaturas de impedância afim de se obter uma quantidade razoável para verificar como a estrutura se comportava no dado momento. A partir de análises exploratórias prévias, verificou-se que os dados obtidos no PZT 2 não representavam assinaturas típicas de impedância, devido a problemas na aquisição e portanto, esses dados não foram utilizados no estudo.

4 Resultados

Considerando os dados do experimento para cada PZT, são computadas, para cada assinatura, as métricas CCD (4) e RMSD1 (3), que são as variáveis de entradas dos SBRFs. A variável de saída é o nível de dano obtido pela Tabela 1, observada no experimento por meio do incremento da massa na estrutura.

O treinamento através da ANFIS, no Matlab, é realizado tomando os índices das combinações de CCD e RMSD1 pares ou ímpares, tomadas de acordo com a ordem das coletas de assinaturas ou, em alguns treinamentos, em ordem crescente de RMSD1. A escolha do conjunto de dados utilizado no treinamento é o que melhor se aproximou dos resultados esperados. Assim, os dados não utilizados no treinamento foram considerados para validação. Desta forma, a quantidade de SBRFs construídos é 28, por ser 4 temperaturas e 7 PZTs considerados no estudo.

Para o treinamento de cada modelo são configuradas quantidades diferentes de funções de pertinência. Com o objetivo de ilustrar os gráficos das funções de pertinências, geradas pelo treinamento, no PZT1 para a temperatura de $20^{\circ}C$, tem-se a Figura 3. O sistema gerou 30 regras fuzzy e os parâmetros das funções de saída.



(a) Funções de pertinência para a entrada CCD.

(b) Funções de pertinência da entrada RMSD1.

Figura 3: Funções de pertinência das entradas.

A validação dos SBRFs é realizada, aproximando os valores de saídas dos SBRFs pelo inteiro mais próximo do nível de dano 0, 1, 2, 3 ou 4, e comparando com o dano observado no experimento.

Ademais, tem-se que a Tabela 2, apresenta as porcentagens de acerto dos SBRFs, considerando 50 valores de CCD e RMSD1 para cada temperatura e PZT.

PZT/Temperatura	-10°C	0°C	10°C	20°C
1	88%	96%	88%	100%
3	100%	90%	100%	98%
4	94%	100%	100%	100%
5	88%	94%	100%	100%
6	80%	96%	100%	100%
7	100%	92%	100%	94%
8	88%	88%	98%	88%

Tabela 2: Porcentagens de acerto dos SBRFs de acordo com cada temperatura.

Desta forma obtém-se uma porcentagem total de acertos médio de 95%, evidenciando a boa qualidade dos modelos.

5 Conclusão

Os resultados obtidos por meio dos SBRFs evidenciaram que a técnica pode ser uma ferramenta promissora para a detecção de dano, visto que obteve uma porcentagem média de acerto de 95%.

Agradecimentos

Todos os autores agradecem à Petrobras - Petróleo Brasileiro S.A. [5] pelo apoio neste projeto.

Referências

- [1] Palomino, L.V. Análise das métricas de dano associadas à técnica da impedância eletomecânica para o monitoramento de integridade estrutural. 2008. 133 f. In: Universidade Federal de Uberlândia, 2008. - UFU. **Dissertações e teses defendidas na UFU: 2008.** Uberlândia, 2008. Completo.
- [2] Zadeh, L.A. Fuzzy sets. **Information and control**, v. 8, n.3, p. 338-353, 1965.
- [3] Jafelice, R. S. M.; Barros, L. C.; Bassanezi, R. C., **Teoria dos Conjuntos Fuzzy com Aplicações.** São Carlos, SP: SBMAC, 2012, v. 17, n.119.
- [4] Jang, J.-S. R., **Adaptive-Network-Based Fuzzy Inference System**, IEEE Transactions on Systems, Man, and Cybernetics, 1993.
- [5] Petrobras - Petróleo Brasileiro S.A. – R&D Center (CENPES), CEP:21941-915, RJ. Brasil.



X MOSTRA DE INICIAÇÃO CIENTÍFICA DA FAMAT

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
24 A 26 DE MAIO DE 2021

Predição do modelo de crescimento econômico de Solow com parâmetro fuzzy Prediction of the Solow economic growth model with fuzzy parameter

Paulo Victor Santana

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil
santanapvicctor@gmail.com

Rosana Sueli da Motta Jafelice

UFU, Faculdade de Matemática, Uberlândia, MG, Brasil
rmotta@ufu.br

Julio Fernando Costa Santos

UFU, Departamento de Economia e Relações Internacionais, Uberlândia, MG, Brasil
julio.costa@ufu.br

Resumo

Palavras-chave

Conjuntos fuzzy .
Taxa de poupança.
ANFIS.
Modelo de Solow.
Crescimento econômico .

Para os economistas o estudo dos fatores que influenciam o crescimento econômico dos países tem sido de grande relevância. O objetivo deste trabalho é encontrar uma predição do modelo de Solow para três países, considerando a taxa de poupança como um parâmetro fuzzy, devido as incertezas que existem na dinâmica econômica. Este parâmetro é obtido através dos dados do consumo das famílias, do saldo de transações correntes e da taxa de juros real. A metodologia utilizada para determinar o sistema baseado em regras fuzzy é o *Adaptive-Network-Based Fuzzy Inference System* (ANFIS). Considera-se que os resultados são promissores, pois as aproximações numéricas da solução do modelo Fuzzy de Solow resulta gráficos qualitativamente similares aos do modelo clássico. Também, a quantidade de dados de países utilizados no treinamento do ANFIS é expressiva, sendo o *Root Mean Square Error* (RMSE) de aproximadamente 0.039, mostrando o potencial desta abordagem matemática.

Abstract

For economists the study of the influential factors for the economic growth of the countries has been of great relevance. The aim of this work is to forecast using Solow Model applied to three countries, considering the savings rate as a fuzzy parameter, due to the uncertainties that exist in the economic dynamics. This parameter is obtained through data on household consumption, the balance of current transactions and interest rate. The methodology used to determine the fuzzy rule-based system is the *Adaptive-Network-Based Fuzzy Inference System* (ANFIS). The results are considered to be promising, since the numerical approximations of the solution of the Solow Fuzzy model results in qualitatively similar graphs to the classic model. Also, the amount of country data used in ANFIS training is significant, with the *Root Mean Square Error* (RMSE) of approximately 0.039, showing the potential of this mathematical approach.

Keywords

fuzzy sets.
saving rate .
ANFIS.
Solow model.
Economic growth .

1 Introdução

Há países no mundo que possuem padrões de vida que eram inimagináveis há décadas atrás. Isso está diretamente ligado ao crescimento econômico que vivenciamos nos últimos tempos, em especial após a primeira revolução industrial. Assim, surge a importância de estudarmos modelos de crescimento econômico. Tais modelos podem nos fornecer uma estimativa de como a economia vai estar em certo período de tempo, e com isso podemos ter uma noção do padrão de vida de um determinado país em um momento futuro.

Este trabalho é baseado no modelo de Solow, que é um modelo de crescimento neoclássico proposto por Robert Solow com o intuito de responder uma simples pergunta: "por que uns países são mais ricos que outros"[1]. Tendo em vista a dificuldade em estimar o parâmetro (s) do modelo de Solow pelos métodos tradicionais econométricos, fizemos uso do ANFIS para determinar (s) como uma variável de saída de um Sistema Baseado em Regras Fuzzy (SBRF). As variáveis de entrada são: consumo das famílias utilizado no cálculo do Produto Interno Bruto (PIB), o saldo em transações correntes, também empregado na obtenção do PIB e a taxa de juros real. O PIB é a soma de todos os bens e serviços finais produzidos por um país, estado ou cidade, geralmente em um ano. O treinamento do ANFIS é realizado com a inserção de dados econômicos de 73 países, quase metade dos países existentes no mundo. Na próxima seção são apresentados conceitos importantes da teoria dos conjuntos fuzzy.

2 Teoria dos Conjuntos Fuzzy

A lógica fuzzy pode ser vista como extensão da lógica tradicional, em que é inserida uma noção de graus de pertinência de um elemento x a um conjunto U qualquer. Esta metodologia permite tratar matematicamente termos linguísticos, como ‘aproximadamente’, ‘em torno de’, dentre outros.

Definição 2.1. Um conjunto fuzzy A do universo U é caracterizado por uma função de pertinência $\mu_A : U \rightarrow [0, 1]$. Esta função é chamada de função de pertinência do conjunto Fuzzy A . A imagem de um elemento de U por μ_A é interpretado como o grau com que o elemento x está contido no conjunto fuzzy A , ou seja, $\mu_A(x) = 0$ indica a não pertinência do x à A , e $\mu_A(x) = 1$ indica pertinência plena ao conjunto A . Um conjunto fuzzy também pode ser representado por um conjunto clássico de pares ordenados da forma $A = \{(x, \mu_A(x)) | x \in U\}$.

Definição 2.2. Uma regra fuzzy é uma sentença do tipo "Se x é A então y é B ", em que A, B são conjuntos fuzzy em U e V , respectivamente, com $x \in U$ e $y \in V$.

Um sistema que é amplamente utilizado em problemas de modelagem, controle e classificação é o SBRF, que pode ser visto como um mapeamento entre a entrada e a saída na forma $y = f(x)$ onde $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$, veja Figura 1. As quatro componentes deste sistema são:

- Fuzzificação: transforma as componentes de entradas em conjuntos fuzzy em seus respectivos

domínios. Muitas vezes é necessário um especialista da área do fenômeno a ser modelado para auxiliar na construção das funções pertinência das entradas.

- Base de regras: composta por uma coleção de proposições do tipo “Se condição então ação”, em que a “condição” e a “ação” podem ser descritas linguisticamente com o conhecimento do especialista da área.
- Método de inferência fuzzy: nesta componente que cada regra fuzzy é “traduzida” matematicamente por meio da metodologia dos conjuntos fuzzy.
- Defuzzificação: é o processo de representar um conjunto fuzzy por um número real.

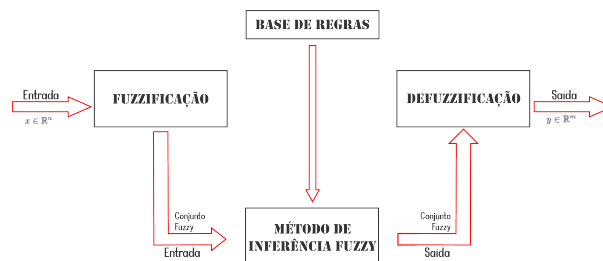


Figura 1: Sistema Baseado em Regras Fuzzy.

Na literatura existem vários métodos de inferência fuzzy, neste trabalho o método utilizado é o de Takagi-Sugeno [2]. A ferramenta matemática fundamental aplicada aos dados econômicos dos países em estudo é o ANFIS. Esta ferramenta é uma rotina de treinamento com sistema de inferência Fuzzy do tipo Takagi-Sugeno que utiliza um algoritmo de aprendizagem para identificar a partir de um conjunto de dados, as funções de pertinência de entrada, a base de regras e os parâmetros da saída do SBRF [3].

3 Modelo Clássico de Solow

O modelo de Solow relaciona as variáveis Trabalho (L), Eficiência do Trabalho (A) e Capital(K). A equação diferencial ordinária (EDO) que governa o modelo é dada por [4]:

$$\dot{k}(t) = s \frac{Y(t)}{A(t)L(t)} - (\delta + n + g)k(t), \tag{1}$$

em que t é o tempo, $k(t) = \frac{K(t)}{A(t)L(t)}$ é o capital por unidade de trabalho efetivo, $Y(t)$ é uma função produção, n, g são as taxas de crescimento de $L(t)$ e $A(t)$ respectivamente, δ é a depreciação do capital em relação ao tempo e s é a taxa da poupança, ou *saving rate*, em inglês. Neste trabalho, a função produção aplicada é a de Cobb-Douglas com $\alpha = 0.5$, dada por:

$$Y(t) = K(t)^\alpha (A(t)L(t))^{1-\alpha}. \tag{2}$$

Substituindo (2) na equação (1) obtém-se:

$$\dot{k}(t) = sk(t)^{0.5} - (\delta + n + g)k(t). \tag{3}$$

4 Modelo Fuzzy de Solow

Neste modelo, a taxa de poupança (s) que é um parâmetro da equação (3) é a variável de saída do SBRF obtido pelo ANFIS. O especialista na área econômica sugeriu que as variáveis de entrada sejam o consumo das famílias obtido no PIB (f), o saldo em transações correntes determinado no PIB (c) e a taxa de juros real (j). Assim, $s = s(f, c, j)$. Os termos linguísticos das variáveis de entrada são classificados da seguinte forma: de f são Muito Baixo, Baixo, Médio e Alto; de c , Baixo, Médio e Alto e da variável j , Baixa, Média e Alta. Os gráficos das funções de pertinências das variáveis de entrada são apresentados nas Figuras 2, 3 e 4.

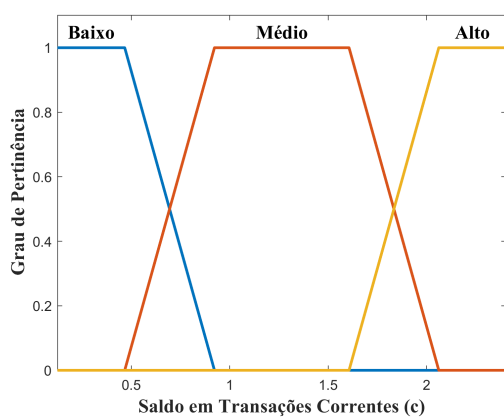


Figura 2: Funções de pertinência de c .

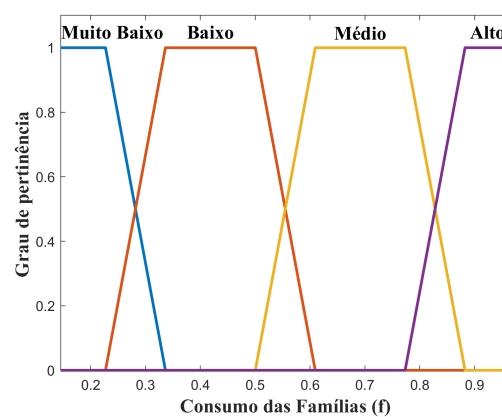


Figura 3: Funções de pertinência de f .

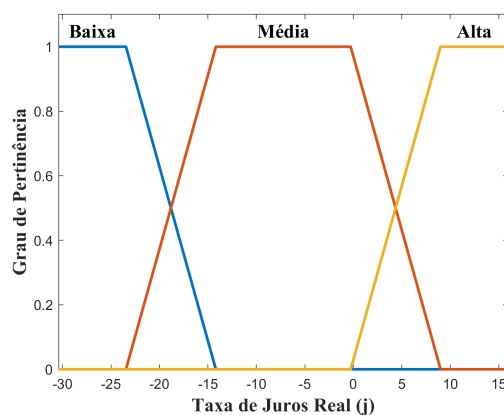


Figura 4: Funções de pertinência de j .

Este SBRF possui 36 regras fuzzy do método de Takagi-Sugeno, para exemplificar:

$$\text{Se } f \text{ é Alto e } c \text{ é Baixo e } j \text{ é Médio então } s = 0.5328f - 0.1104c - 0.01379j - 0.3935.$$

5 Resultados

Neste estudo foram coletados dados em diversas fontes, tais como o Fundo Monetário Internacional (FMI) [5] e outras [6], [7] e [8]. Para realizar o treinamento do ANFIS utilizou-se os seguintes dados: c, f, j e s de 73 países diferentes nos anos de 2016, 2017 e 2018. O RMSE dos valores reais em relação aos dados de treinamento foi de aproximadamente 0.039 no passo 100. Posteriormente, o SBRF foi validado com os dados do ano de 2019 de 54 países. O cálculo do máximo do erro absoluto entre os valores de s gerados pelo SBRF e pelos valores reais foi inferior a 0.17.

Escolhemos três países para analisar diferentes crescimentos econômicos futuros: Qatar, Brasil e Nigéria. Os parâmetros utilizados para a resolução numérica da equação (3) estão apresentados na Tabela 5 e a condição inicial é $k(0) = 1$. Tal resolução foi realizada utilizando o Método de Runge-Kutta de 4ª ordem [9].

País	n	g	δ	c	f	j	s SBRF	s Real
Qatar	0.01795	0.0271	0.072	0.5869	0.1555	4.92	0.4819	0.5250
Brasil	0.0075	0.0238	0.0478	0.2706	0.6499	0.77	0.2004	0.1714
Nigéria	0.0256	0.0129	0.0342	0.1850	0.8373	2.10	0.1203	0.1070

Tabela 1: Tabela de dados dos países avaliados.

Os gráficos das Figuras 5 e 6 apresentam as aproximações das soluções numéricas da equação (3) para os três países estudados, utilizando os valores de s obtidos pelo SBRF e dos dados, respectivamente.

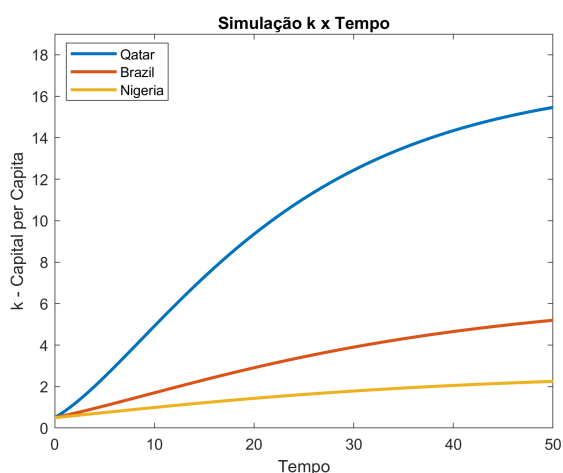


Figura 5: Solução da EDO com s gerado pelo SBRF.

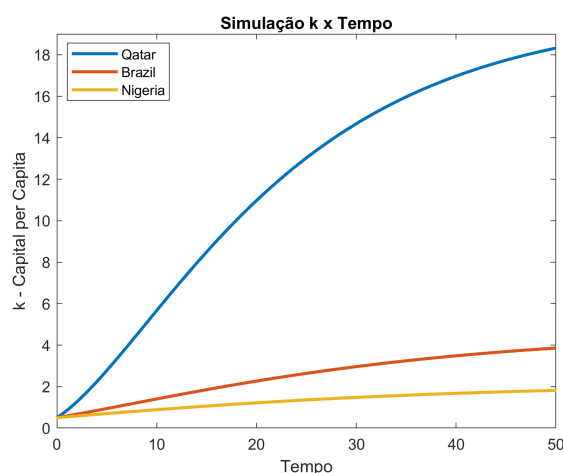


Figura 6: Solução da EDO com s real.

Avaliando os gráficos do crescimento econômico dos três países em função do tempo, obtidos pelos gerados pelo SBRF e dos dados reais, observa-se um comportamento similar e compatível com o esperado pela área econômica.

6 Considerações finais

Os resultados alcançados são condizentes com a realidade e incorporam a opinião do especialista na área econômica quanto a dependência da taxa de poupança em relação as variáveis de entrada inseridas no ANFIS. Estes resultados nos motivam na realização de trabalhos futuros, pois mostram que a teoria dos conjuntos fuzzy pode ser uma ferramenta de grande aplicabilidade nas ciências econômicas.

Agradecimentos

O primeiro autor agradece à CAPES pela bolsa de estudos concedida para realização do mestrado.

Referências

- [1] SOLOW R. M. **A Contribution to the Theory of Economic Growth**. The Quarterly Journal of Economics, Oxford University Press, vol. 70(1), pages 65-94.
- [2] JAFELICE, R. S. M.; BARROS, L. C.; BASSANEZI, R. C. **Teoria dos Conjuntos Fuzzy com Aplicações**. 2ª Edição. São Carlos, SP: SBMAC, 2012, 119 p.
- [3] JANG, J.-S.R. **ANFIS: Adaptive-Network-based Fuzzy Inference System**. IEEE Transactions on Systems, Man, and Cybernetics, 1993.
- [4] ROMER, D. **Advanced Macroeconomics**. 4ª Edição: McGraw-Hill, 2012.
- [5] IMF DATA. **Taxa de juros nominal**: Disponível em: <<https://data.imf.org/regular.aspx?key=61545867>>. Acesso em: 10 abr. 2021.
- [6] WORLD BANK DATA. **Inflação**: Disponível em: <<https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG>>. Acesso em: 10 abr. 2021.
- [7] WORLD BANK DATA. **Population growth**: Disponível em: <<https://data.worldbank.org/indicator/SP.POP.GROW>>. Acesso em: 10 abr. 2021.
- [8] PWT. **Penn World Table version 10.0**: Disponível em: <<https://www.rug.nl/ggdc/productivity/pwt/?lang=en>>. Acesso em: 10 abr. 2021.
- [9] BURDEN, R. L.; FAIRES, J. D. **Numerical Analysis**. Brooks/Cole, 2011.