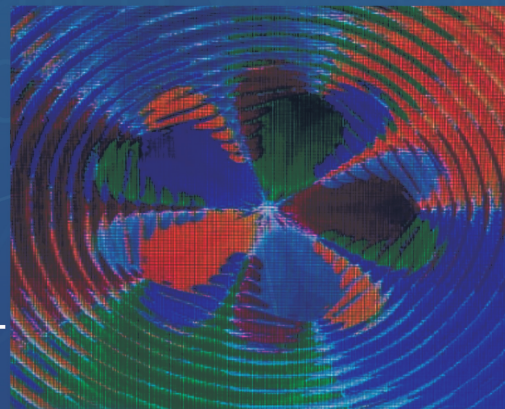


ERMAC 2023

ENCONTRO REGIONAL DE MATEMÁTICA
APLICADA E COMPUTACIONAL & SIMPÓSIO
1ª DÉCADA DO PPG-MCCT

30 de outubro a 01 de novembro de 2023

**ESCOLA DE ENGENHARIA INDUSTRIAL E METALÚRGICA DE VOLTA
REDONDA (EEIMVR) - UNIVERSIDADE FEDERAL FLUMINENSE**



Fonte: Souza, E. A. Métodos interativos para problemas não lineares, 2015. Dissertação (Mestrado em Modelagem Computacional em Ciência e Tecnologia) - Universidade Federal Fluminense, Volta Redonda, 2015.



Anais do ERMAC-RJ 2023

Programa de Pós-Graduação Modelagem Computacional
em Ciência e Tecnologia (PPG-MCCT/UFF)

Editores:

Panters Rodríguez-Bermúdez (UFF)

Tiago Araújo Neves (UFF)

Claudia Mazza Dias (UFRRJ)

ISBN: 978-65-272-0185-4

DOI: doi.org/10.29327/1343139

Volta Redonda, Dezembro 2023



Universidade
Federal
Fluminense

ERMAC 2023

& Simpósio 1ª década do PPG-MCCT

Universidade Federal Fluminense (UFF) - Volta Redonda, Rio de Janeiro
30 de outubro a 1 de novembro de 2023

Palestrantes Convidados



João Frederico
C. A. Meyer



Zochil González
Arenas



Reinaldo
R. Ramos



Maria Claudia
S. Boeres



Dayse Haime
Pastore



Sandra Mara
Malta

Minicursos

Curso Básico de Koopman Operator

Expositores: Amaury Alvarez Cruz (UFRJ),
Local: Sala/laboratório
Horários: Segunda 30/10: 17h30 às 19h30
e Terça 31/10: 15h30 às 17h30

Introdução aos Loops de Código

Expositores: Rosemary Miguel Pires (UFF),
Local: Sala/laboratório
Horários: Segunda 30/10: 17h30 às 19h30
e Terça 31/10: 15h30 às 17h30

Introdução ao Método de Elementos Finitos

Expositores: Issac Pinheiro dos Santos (UFES)
Local: Sala/laboratório
Horários: Segunda 30/10: 17h30 às 19h30
e Terça 31/10: 15h30 às 17h30

Modelagem Matemática: Uma Ferramenta no Ensino da Matemática

Expositores: Cláudia Mazza Dias (UFRRJ),
Marcela Lima Santos Pereira (UFRRJ)
Local: Sala/laboratório
Horários: Terça 31/10: 17h30 às 19h30
e Quarta 01/11: 8h às 10h

Teoria Espectral de Grafos e Aplicações

Expositores: Francisca Andrea Macedo
França (UFF),
André Ebling Brondani (UFF)
Local: Sala/laboratório
Horários: Terça 31/10: 17h30 às 19h30
e Quarta 01/11: 8h às 10h

Introdução aos Métodos Pseudo-Espectrais

Expositores: Daniel G. Alfaro Vigo (UFRJ)
Local: Sala/laboratório
Horários: Terça 31/10: 17h30 às 19h30
e Quarta 01/11: 8h às 10h



Mais informações:

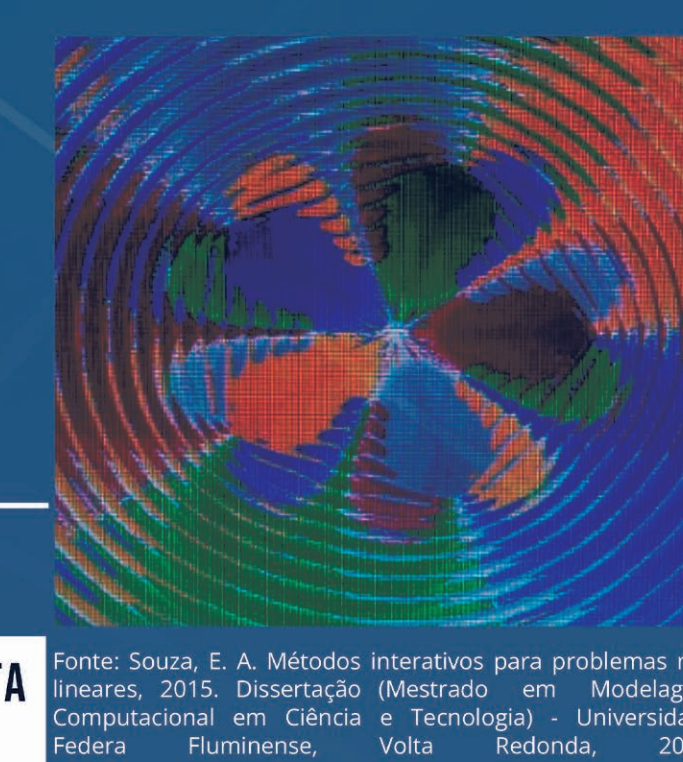
<https://www.even3.com.br/encontro-regional-de-matematica-aplicada-e-computacional-ermac-rj-simposio-1-decada-ppg-mcct-295052/>

ERMAC 2023

ENCONTRO REGIONAL DE MATEMÁTICA
APLICADA E COMPUTACIONAL & SIMPÓSIO
1ª DÉCADA DO PPG-MCCT

30 de outubro a 01 de novembro de 2023

ESCOLA DE ENGENHARIA INDUSTRIAL E METALÚRGICA DE VOLTA
REDONDA (EEIMVR) - UNIVERSIDADE FEDERAL FLUMINENSE



Universidade
Federal
Fluminense

Dados Internacionais de Catalogação na Publicação (CIP)

AN532 Anais do Encontro Regional de Matemática Aplicada e Computacional (ERMAC-RJ) & Simpósio Primeira Década PPG-MCCT. Anais...Volta Redonda(RJ) UFF, 2023

Disponível em www.even3.com.br/anais/encontro-regional-de-matematica-aplicada-e-computacional-ermac-rj-simposio-1-decada-ppg-mcct-295052

ISBN: 978-65-272-0185-4

DOI: doi.org/10.29327/1343139

1. Matemática 2. Ciência da computação, informação e obras gerais

UFF

CDD - 370

CORPO EDITORAL

COMISSÃO CIENTÍFICA

ALEXANDRE SANTOS FRANCISCO (UFF)
ANDRÉ EBLING BRONDANI (UFF)
AMAURY ALVAREZ CRUZ (UFRJ)
CLAUDIA MAZZA DIAS (UFRRJ)
DAYSE HAIME PASTORE (CEFET-RJ)
GUSTAVO BENITEZ ALVAREZ (UFF)
ISAAC PINHEIRO DOS SANTOS (UFES)
PANTERS RODRÍGUEZ BERMUDEZ (UFF)
REINALDO RODRIGUEZ RAMOS (UFF/ UNIVERSIDAD DE LA HABANA)
SANDRA MARA MALTA (LNCC)
TIAGO ARAÚJO NEVES (UFF)
WAGNER RAMBALDI TELLES (UFF)
ZOCHIL GONZALEZ ARENAS (UERJ)

COMISSÃO ORGANIZADORA LOCAL

AFONSO A. DE CARVALHO PERES
ANA PAULA MARTINAZZO
ANDRÉ EBLING BRONDANI
ANDRÉ GUSSO
EMERSON SOUZA FREIRE
GUSTAVO BENITEZ ALVAREZ
JORGE A. RODRIGUEZ DURÁN
NADJA V. VASCONCELOS DE AVILA
PANTERS RODRÍGUEZ BERMÚDEZ (PRESIDENTE)
REINALDO RODRIGUEZ RAMOS
TIAGO ARAÚJO NEVES (VICE-PRESIDENTE)
VANESSA DA SILVA GARCIA

WELLINGTON CARLOS DE JESUS

SECRETÁRIA DO EVENTO

CAMILA SOUZA SILVA

PROMOVENDO REFLEXÕES SOBRE O USO DA MODELAGEM MATEMÁTICA ATRAVÉS DA LEI DE NEWCOMB-BENFORD

Marcelo Ferreira Farias
Universidade Federal Rural do Rio de Janeiro
mfarias@ufrj.br

Lavinia Beatriz Duarte da Silva
Universidade Federal Rural do Rio de Janeiro
laviniabduarte@ufrj.br

Victor Augusto do Carmo Duarte
Universidade Federal Rural do Rio de Janeiro
victoracd2002@ufrj.br

Resumo: A lei de Newcomb-Benford é um fenômeno estatístico que descreve a distribuição das frequências dos dígitos iniciais em conjuntos de dados provenientes de uma ampla gama de fenômenos do mundo real. Essa lei tem sido aplicada com sucesso em diversas áreas, incluindo auditorias financeiras, análise de dados, detecção de fraudes e investigações forenses, mas alguns abusos são cometidos, o que nos leva à discussão do mau uso das ferramentas matemáticas na tomada de decisões em problemas reais. Neste trabalho, relatamos a experiência na realização de uma atividade com uma turma de ingressantes do curso de Matemática, da UFRRJ, no qual mostramos as distribuições de dados numéricos em diferentes contextos e suas similaridades com a previsão da lei, discutimos os limites do uso de argumentos matemáticos na modelagem matemática em problemas reais, em acordo com as preocupações da Matemática Crítica, e propomos um experimento feito pelos próprios alunos envolvendo a distribuição dos primeiros dígitos significativos dos números de seguidores nas redes sociais.

Palavras-chave: Ensino de Matemática. Educação Matemática Crítica. Probabilidade. Lei de Newcomb-Benford.

Introdução

A lei de Newcomb-Benford estabelece que, em muitos conjuntos de dados numéricos, as frequências relativas de ocorrência dos dígitos iniciais seguem um padrão específico, diferente de uma distribuição equiprovável, que seria o mais natural, a priori. De acordo com essa lei, o dígito “1” ocorre como o primeiro dígito significativo em cerca de 30,1% dos casos; o dígito “2” aparece menos, aproximadamente em 17,6% das ocorrências, e decresce sucessivamente até o dígito “9”, com cerca de 4,6% de frequência relativa. Assim, a probabilidade de ocorrência de um dígito inicial não é uniforme e segue uma distribuição logarítmica (BENFORD, 1938).

A observação dessa lei remonta ao trabalho de Simon Newcomb (1835-1909), um astrônomo e matemático canadense, que observou a tendência de certos dígitos ocorrerem com mais frequência em tabelas de logaritmos consultadas por diversos pesquisadores de diferentes áreas do conhecimento, pois as páginas que traziam números começados com o dígito “1” estavam mais desgastadas por serem as mais usadas, seguidas pelas páginas de números iniciados em “2”, e assim sucessivamente até o 9 (NEWCOMB, 1881). No entanto, foi o matemático Frank Benford (1883-1948) que, em 1938, popularizou essa observação empírica e formulou uma explicação teórica para o fenômeno. Desde então, a lei de Newcomb-Benford tem sido objeto de extenso estudo e aplicação em várias áreas do conhecimento.

Além da fundamentação teórica, tem sido confirmada empiricamente em uma ampla variedade de conjuntos de dados reais. Estudos têm mostrado que essa distribuição é observada em fenômenos tão diversos como populações de cidades, números contidos em jornais, valores de contas bancárias, comprimentos de rios, entre outros (BENFORD, 1938; GONSALVES, 2020).

Embora a Lei de Newcomb-Benford tenha se mostrado eficaz em várias áreas, aplicações equivocadas e controversas têm conquistado grande repercussão nas mídias. Um caso emblemático se refere a uma tentativa de imputar fraudes aos resultados das eleições presidenciais brasileiras de 2014, que foi utilizado para desacreditar o processo eleitoral de 2018 e, também, o de 2022. Um vídeo produzido em 2015, de poucos minutos, se propagou nas redes sociais de apoiadores de um candidato marcado pelas constantes declarações contrárias às urnas eletrônicas, em épocas próximas aos pleitos de 2018 e 2022. A tática de desacreditar as eleições tem sido usada recorrentemente por candidatos de extrema-direita, em diversos países do mundo, nos últimos anos. Países com Peru, Alemanha, Estados Unidos da América, Bolívia e Israel tiveram resultados de pleitos presidenciais questionados (GIOVANAZ, 2021).

A justificativa da probabilidade de fraude nas eleições de 2014 no Brasil ter sido grande, dada por um suposto especialista, que posteriormente se mostrou associado ideologicamente ao espectro político do próprio candidato que alegava fraude no processo eleitoral, envolvia argumentos estatísticos e o uso da Lei de Newcomb-Benford não explicitados. Diversas matérias veiculadas na imprensa brasileira apontaram a falsidade das alegações do vídeo (MOTA, 2018; UOL, 2018). A ampla divulgação do vídeo de caráter científico fraudulento levou a maior instância da Justiça Eleitoral brasileira, responsável por organizar o sistema eleitoral do Brasil, a se manifestar oficialmente, em seu sítio: “a simples análise por seção de votação ou por zonas eleitorais, que possuem poucas urnas, não é eficaz, pois o universo amostral é muito pequeno” (TRIBUNAL SUPERIOR ELEITORAL, 2022).

Sabemos como resultados matemáticos e estatísticos são parte da estrutura de argumentações de diversas naturezas. Ver a Matemática como um sistema infalível, livre de influências e interesses, onde se valoriza soluções matematizadas de modelos em detrimento a outras soluções obtidas por técnicas não matematizadas limita a capacidade de interpretação da realidade em sua enorme complexidade. Bassanezi (2010) nos alerta para este fato ao considerar a modelagem como processo de criação de modelos onde estão definidas as estratégias de ação do indivíduo sobre a realidade, carregada de interpretações e subjetividades do próprio modelador.

Diversos pensadores buscam refletir sobre a finalidade da educação matemática promovendo posturas democráticas, debatendo questões sociais, econômicas, culturais, cidadania e estruturas de poder que de certo modo são perpetuadas com o uso da Matemática. Essas atitudes educacionais têm em Skovsmose (1996, 2000, 2001, 2007) um dos seus mais importantes proponentes. Para o autor, no trabalho pedagógico, professor e alunos devem assumir o papel de investigadores da sua realidade social, criando possibilidades múltiplas para a construção do conhecimento. Ele propõe uma exploração matemática na qual o objetivo principal é criar uma contraposição ao que denomina paradigma do exercício, que pressupõe uma organização da aula em que os exercícios matemáticos só apontam para uma resposta correta única e definitiva, fazendo com que a matemática não permita diálogo.

Há na obra de Skovsmose uma preocupação com o desenvolvimento da *materacia*, capacidade de interpretar e manejar sinais e códigos e de propor e utilizar modelos para encontrar soluções para os problemas enfrentados diariamente e, também, elaborar abstrações sobre as representações propostas para os sistemas retirados da realidade (D’AMBROSIO, 2004). “*Materacia* não se refere apenas às habilidades matemáticas, mas também à competência de interpretar e agir numa situação social e política estruturada pela matemática.” (SKOVSMOSE, 2000, p. 68).

Borba e Skovsmose (2001) chamam de “ideologia da certeza a estrutura geral e fundamental de interpretação para [...] questões que transformam a matemática em uma linguagem de poder.” Para os autores, os alunos deveriam ser persuadidos contra ideias de que o argumento matemático é superior por sua própria natureza e convencidos que a matemática é somente uma maneira possível de olhar o fenômeno.

Com essas mesmas preocupações em torno dos objetivos da educação matemática, desenvolvemos uma atividade com a turma de 25 ingressantes do segundo semestre letivo de 2023 do curso de Matemática, da Universidade Federal Rural do Rio de Janeiro, Campus Nova Iguaçu. Ela teve como objetivo propor uma reflexão sobre a aplicabilidade da matemática na tentativa de se estabelecer uma visão definitiva da realidade. Usamos a Lei de Newcomb-Benford por enxergar um potencial para trazer tais

reflexões a respeito dos limites do uso de argumentos matemáticos na tomada de decisões frente a um problema real.

Faremos o relato da nossa experiência, onde introduzimos o funcionamento dessa lei, mediante o estudo de logaritmos, unindo conceitos matemáticos e tecnologia (programação na linguagem Python), explicitando situações nas quais a Lei de Newcomb-Benford é aplicada com sucesso, mas também mostraremos os limites da aplicabilidade com uma compilação dos dados obtidos na atividade.

A Lei de Newcomb-Benford

A Lei de Newcomb-Benford descreve uma distribuição de frequências dos dígitos iniciais diferente daquela esperada em uma distribuição uniforme (com probabilidade $\frac{1}{9}$ para cada dígito de 1 a 9). Segundo a lei, dado um dígito significativo a , sua frequência de ocorrência se dá pela expressão:

$$P(a) = \log \left(\frac{a+1}{a} \right) = \log \left(1 + \frac{1}{a} \right),$$

para $a \in \{1, 2, 3, \dots, 9\}$ (BENFORD, 1938).

Khosravani e Rasinariu (2018) destacam que também existe uma expressão logarítmica para o segundo, terceiro e para o n -ésimo dígito. A regra geral usando logaritmos é basicamente a mesma utilizada apenas para o primeiro dígito. Tome a_1 sendo o primeiro e a_2 sendo o segundo dígito. Dessa forma, a regra geral logarítmica se dá por:

$$P(a_1 a_2) = \log \left(1 + \frac{1}{a_1 a_2} \right),$$

para $a_1 \in \{1, 2, 3, \dots, 9\}$ e $a_2 \in \{0, 1, 2, \dots, 9\}$.

A Lei de Newcomb-Benford tem sido demonstrada e justificada através de diversas abordagens teóricas ao longo dos anos. Uma importante contribuição para a validação teórica dessa lei foi feita por Hill (1995), cuja prova parte do pressuposto de que uma distribuição de dígitos é invariante em relação à base se, e somente se, satisfaz à Lei de Newcomb-Benford, resultado provado por Pinkham (1961). Desse modo, Hill fornece uma base matemática sólida para a referida lei, mostrando que a distribuição de dígitos iniciais é uma consequência natural da invariância de base. Santos et al. (2003), por sua vez, destrincha a prova de Hill.

Metodologia da Atividade

A palestra “A Lei de Newcomb-Benford e Aplicações” foi planejada para ser realizada em um período de 50 minutos. Consistiu de uma apresentação de 30 minutos (com o auxílio de data show e slides) e de uma atividade didática de 20 minutos. Na apresentação, falamos resumidamente sobre a história dos logaritmos; as propriedades operacionais dos logaritmos decimais (característica e mantissa); a disposição das mantissas dos logaritmos decimais em tabelas de logaritmos; a revolução técnica e tecnológica propiciada pela invenção; a observação das páginas mais usadas de uma tabela de logaritmos por Simon Newcomb e sua posterior conjectura; o enunciado da lei de Newcomb-Benford, e diversas aplicações em conjuntos de dados mais variados, bem como casos nos quais a lei não se aplica (por exemplo, na escolha do nono dígito que compõe o Cadastro da Pessoa Física-CPF) e casos em que a lei pode levar a conclusões falsas (o caso da suposta fraude nas urnas eletrônicas brasileiras). Após a apresentação, abrimos um espaço para que os discentes mostrassem suas percepções e dúvidas. Demonstraram muita receptividade, interesse e espanto com as previsões de distribuições dos primeiros dígitos propostas da Lei de Newcomb-Benford em diversos conjuntos de dados (áreas de terrenos da União, populações das cidades brasileiras, números da sequência de Fibonacci, durações de notas musicais nas sinfonias de Beethoven, entre outros). Em seguida, passamos à explicação da atividade.

A atividade didática teve como referência uma informação vista no documentário A Era dos Dados (2020), no episódio intitulado Dígitos (Temporada 1, Episódio 4), exibido através da plataforma de *streaming* Netflix. Uma das entrevistadas, a professora Jennifer Golbeck, da Faculdade de Estudos da Informação na Universidade de Maryland, EUA, diz que utilizando como espaço amostral uma quantidade considerável de modelos, na rede social Instagram, ou em qualquer outra, os números de seguidores de um conjunto de pessoas (escolhidas de forma aleatória) seguirá a Lei de Newcomb-Benford.

Antes de começarmos a atividade, verificamos se a turma tinha acesso à internet, o que era essencial para a realização da atividade. Vimos que a grande maioria da turma tinha smartphones com acesso à internet. Aos discentes que não tinham acesso à internet, pedimos que se sentassem juntos aos colegas com acesso à rede. Distribuímos uma ficha contendo uma tabela na qual os estudantes marcariam nas linhas enumeradas de 1 a 9, representando os primeiros dígitos significativos dos dados que seriam coletados.

Solicitamos aos alunos e às alunas que acessassem suas contas do Instagram (todos tinham acesso à rede social, por isso a escolha). Passamos às instruções da atividade: cada discente deveria escolher aleatoriamente o perfil de um seguidor seu e anotar o primeiro dígito significativo do número de seguidores do escolhido; em seguida, pedimos que escolhessem aleatoriamente um seguidor do seu seguidor, e novamente anotassem o primeiro dígito do número de seguidores; procedendo assim, sucessivamente, até analisar 15 contas. Foram coletados ao todo 282 dados numéricos. Levamos pouco mais de 10 minutos para que todos os discentes anotassem os dados observados, mostrando como a atividade pode ser rapidamente realizada.

Em seguida, coletamos os dados registrados pela turma, obtivemos as frequências dos primeiros dígitos significativos e passamos à etapa final da atividade que era a análise dos dados coletados. Realizamos uma implementação na linguagem de programação Python. O código, que analisamos abaixo, já estava pré-programado, faltando somente o preenchimento das frequências observadas de cada dígito significativo dentro do conjunto de dados. Antes do preenchimento, fizemos uma explicação de cada linha do código para a turma.

Código 1: Código em Python para tratamento dos dados e comparação com o previsto para a Lei de Newcomb-Benford.

```

1 import matplotlib.pyplot as plt #biblioteca para plotagem de graficos
2 from math import * #biblioteca para operacoes matematicas
3 import numpy as np #biblioteca para operacoes com vetores
4
5 b = [81, 44, 35, 28, 25, 25, 11, 20, 13] #quantidade de cada digito
6
7 observed = np.array(b)/sum(b) #proporcao observada em cada digito
8 expected = np.array([log((i + 1)/i, 10) for i in range(1,10)]) #benford
9
10 digitos = ['1', '2', '3', '4', '5', '6', '7', '8', '9'] #lista de digitos
11 largura = 0.3 #parametro para largura das barras no grafico
12 posicoes = np.arange(9) #parametro para posicoes das barras no grafico
13
14 plt.bar(posicoes, expected, largura, label='Frequencia teorica')
15 plt.bar(posicoes+largura, observed, largura, label='Frequencia observada')
16 plt.xticks(posicoes+largura, digitos)
17 plt.xlabel('Dígitos') #legenda do eixo x
18 plt.ylabel('Frequencia') #legenda do eixo y
19 plt.legend()
20 plt.show()

```

Analisando o Código 1, tem-se a importação de bibliotecas pertinentes nas linhas 1-3. Elas são responsáveis, respectivamente, pela representação dos dados na forma de gráficos de barras, por operações matemáticas como o cálculo do logaritmo decimal de um número real e por operações envolvendo vetores. Na linha 5, é definido um vetor de 9 posições associadas aos 9 dígitos a serem analisados: cada posição receberá a quantidade de vezes que determinado dígito foi o dígito significativo no número de

seguidores de uma conta no Instagram.

Na linha 7, o vetor com a frequência de cada dígito na primeira posição é efetivamente calculado com base nos resultados obtidos da amostra de dados analisada. Como indicado, a operação é feita a partir da razão entre o número de vezes em que cada dígito ocupou a primeira posição e o número total de dígitos. Na sequência, a distribuição conforme a Lei de Newcomb-Benford é calculada para cada dígito de 1 a 9, e é armazenada em outro vetor de 9 posições, que posteriormente será comparado ao primeiro.

Da linha 10 à linha 20, a representação gráfica dos dados encontrados é construída e exibida a partir de alguns parâmetros necessários para o uso da biblioteca importada na linha 1. Entre eles estão a largura das barras, as legendas associadas a cada uma, e outros. Essa sequência de comandos permite plotar o gráfico que associa a cada dígito duas barras: uma representando a frequência observada desse dígito na amostra de dados (na primeira posição), e outra representando a frequência teórica prevista pela Lei de Newcomb-Benford.

Como se pode perceber na Tabela 1 e na Figura 1, as frequências relativas dos primeiros dígitos foram próximas às previsões da Lei de Newcomb-Benford, mas não tiveram o comportamento decrescente de frequência relativa prevista pelo modelo. Iniciamos uma discussão com os alunos sobre as incompatibilidades encontradas. O primeiro questionamento, feito pelos próprios alunos, foi sobre a obrigatoriedade do conjunto de todos os números de seguidores de uma rede social seguir a distribuição prevista para a lei, e se seguisse por qual o motivo não encontramos a distribuição. A amostra de dados coletados poderia ser considerada pequena e, se continuássemos a coleta, a tendência era a convergência entre as previsões teóricas e as frequências relativas observadas, o que demonstra um raciocínio probabilístico baseado na chamada Lei Fraca dos Grandes Números (ROSS, 2010). Segundo Kruger e Yavadalli (2017), não há consenso na literatura para que o tamanho de uma amostra seja adequada, mas que seja de pelo menos 50 a 100 para lei a ser observada (caso seja seguida), e que um tamanho de amostra de 500 ou mais pode ser preferido para uma análise adequada.

Tabela 1: Frequência relativa observada na atividade e a previsão teórica pela lei de Newcomb-Benford.

Dígito	Frequência Relativa	Previsão Teórica
1	28,72%	30,1%
2	15,60%	17,60%
3	12,41%	12,49%
4	9,92%	9,69%
5	8,86%	7,91%
6	8,86%	6,69%
7	3,9%	5,79%
8	7,09%	5,11%
9	4,60%	4,57%

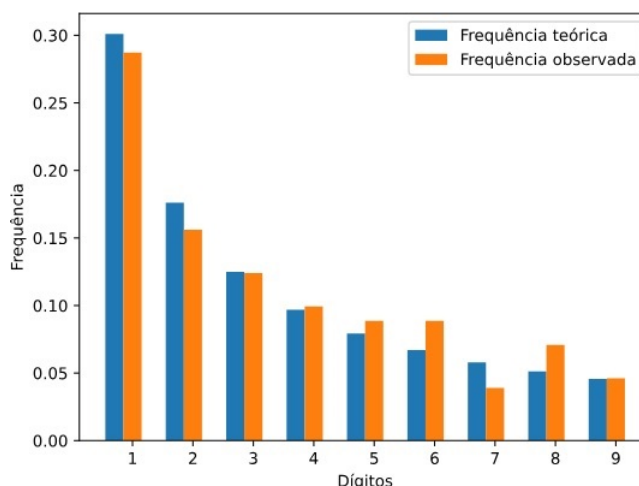


Figura 1: Gráfico obtido a partir do Algoritmo 1, contendo a frequência relativa dos primeiros dígitos significativos obtidas na atividade.

Outros apontamentos feitos pela turma merecem menção: algumas contas de seguidores utilizadas pelos discentes eram privadas, ou possuíam outra ferramenta que não se permitia ver os seguidores, o que acarretaria uma ideia irreal do número de seguidores; outros levantaram a hipótese de contas terem um número falso de genuínos seguidores, tendo em vista que a rede social analisada permite compra de pacotes de seguidores.

No fechamento da atividade, mostramos como o levantamento dessas questões torna evidente o fato de como uma tomada de decisão com o auxílio de ferramentas matemáticas poderosas têm suas

limitações de uso, podendo nos levar a diversas conclusões. Finalizamos a atividade sem fazer apontamentos para as verdadeiras causas da (in)compatibilidade, como aliás defendem Alrø e Skovsmose (2002), ao realçar o elemento básico na realização de uma investigação: as atividades devem ser abertas, de tal forma que os caminhos a seguir e as possíveis conclusões a que se pode chegar não sejam conhecidos de antemão.

Conclusões

Este trabalho busca combinar a análise empírica da Lei de Newcomb-Benford em conjuntos de dados diversos com a proposta de uma atividade didática que promova a compreensão e o interesse dos alunos pelo tema e que introduza o ensino de logaritmos, ou probabilidade, fornecendo uma visão da sua aplicação tanto em um contexto de pesquisa quanto de educação matemática.

Borba e Skovsmose (2001) dizem que “por meio de modelos matemáticos, também nos tornamos capazes de ‘projetar’ uma parte do que se torna realidade. Tomamos decisões baseados em modelos matemáticos e, dessa forma, a matemática molda a realidade”. Devemos como educadores desafiar a ideologia da certeza e esse poder formatador da matemática por meio de um currículo baseado na incerteza, com trabalhos com situações abertas, nas quais surgem ambiguidades, evidenciando os limites da Matemática na solução de um problema real, pelo questionamento dos interesses envolvidos na escolha de modelos matemáticos, não aceitando a neutralidade da matemática e suas soluções infalíveis. É assim que enxergamos a proposta de Borba e Skovsmose (2001) para propiciar uma formação crítica em Matemática. E o fato do nosso experimento ter resultado em uma distribuição da frequência relativa distinta da prevista, podendo parecer inconsistente à primeira vista, realça exatamente o que queríamos discutir: os limites da aplicabilidade da Matemática.

Consideramos que a atividade foi atrativa por ter despertado interesses diversos nos alunos da turma: se interessaram sobre a coleta de dados das aplicações trabalhadas, pediram detalhes da programação do Código 1, pediram a indicação do documentário, referências a livros, etc.

É interessante também notar que esta atividade pode ser desenvolvida e adaptada para turmas do Ensino Médio, por ser este tema motivador para se trabalhar História da Matemática, logaritmos e probabilidades. E sugerimos um cuidado especial, que nos gerou uma certa apreensão durante a realização da atividade, na certificação prévia do acesso à internet por parte dos alunos ou da instituição.

Referências

ALRØ, H.; SKOVSMOSE, O. **Dialogue and learning in Mathematics Education: intention, reflection, critique**. Dordrecht: Kluwer Academic Publishers, 2002. 284 p.

BASSANEZI, R. C. **Ensino-aprendizagem com modelagem matemática: uma nova estratégia**. São Paulo: Contexto. 2010.

BENFORD, F. The law of anomalous numbers. **Proceedings of the American Philosophical Society**, v.78, n.4, 551-572, 1938.

BORBA, M. C.; SKOVSMOSE, O. A Ideologia da Certeza em Educação Matemática. In: SKOVSMOSE, O. **Educação Matemática Crítica – A Questão da Democracia**. Campinas: Papirus, 2001.

D’AMBROSIO, U. A relevância do projeto indicador nacional de alfabetismo funcional – INAF como critério de avaliação da qualidade do ensino de matemática. In FONSECA, M. C. F. R. (Org.), **Letramento no Brasil: habilidades matemáticas**. São Paulo, SP: Editora Global, 2004. pp. 31-46.

DÍGITOS (temporada 1, ep. 4). **A Era dos Dados** [Série documental]. Direção: Nick Brigden. Produção: Christopher Collins, Lydia Tenaglia, Erik Osterholm, Latif Nasser, David Mettler. Estados

Unidos: Zero Point Zero Production Inc., 2020. (46 min.), son., color.

GIOVANAZ, D. Veja o que aconteceu em 7 países onde a direita alegou fraude eleitoral sem provas.

Brasil de Fato. 15 ago. 2021. Disponível em: <https://www.brasildefato.com.br/2021/08/15/veja-o-que-aconteceu-em-7-paises-onde-a-direita-alegou-fraude-eleitoral-sem-provas>. Acesso em: 29 set. 2023

GONSALVES, R. Benford's Law – a simple explanation. **Towards Data Science**, 2020. Disponível em: <https://towardsdatascience.com/benfords-law-a-simple-explanation-341e17abbe75>. Acesso em: 13 jul. 2023

HILL, T. Base-invariance implies Benford's Law. **Proceedings of the American Mathematical Society**, v.123, n.3, 1995.

KHOSRAVANI, A.; RASINARIU, C. Emergence of Benford's Law in Music. **Journal of Mathematical Sciences: Advances and Applications**, v.54, 11-24, 2018.

KRUGER, P.; YAVADALLI, V. The power of one: Benford's law. **South African Journal of Industrial Engineering** Vol 28(2), pp 1-13, 2017.

MOTA, M. et al. Mensagens com conteúdo #FAKE sobre fraude em urnas eletrônicas se espalham nas redes. **O Globo**, 06 out. 2018. Disponível em: <https://g1.globo.com/fato-ou-fake/noticia/2018/10/06/mensagens-com-conteudo-fake-sobre-fraude-em-urnas-eletronicas-se-espalham-nas-redes.ghtml>. Acesso em: 08 set. 2023.

NEWCOMB, S. Note on the frequency of use of the different digits in natural numbers. **American Journal of Mathematics**, v.4, n.1, 39-40, 1881.

PINKHAM, R. On the distribution of first significant digits. **The Annals of Mathematical Statistics**, v.32, n.4, 1223-1230, 1961.

ROSS, S. **Probabilidade : um curso moderno com aplicações**. Porto Alegre : Bookman, 2010.

SANTOS, J., TENÓRIO, J., SILVA, L. Uma aplicação da teoria das probabilidades na Contabilometria: a Lei de Newcomb-Benford como medida para análise de dados no campo da Auditoria Contábil. **Contabilidade, Gestão e Governança**, v.6, n.1, 35-54, 2003.

SKOVSMOSE, O. Critical mathematics education: some philosophical remarks. In: **International Congress on Mathematics Education. Selected lectures**. Sevilha: S. A. E. M. 1996, p. 413-425.

SKOVSMOSE, O. Cenários para investigação. **BOLEMA – Boletim de Educação Matemática**, Rio Claro, n. 14, p. 66-91, 2000.

SKOVSMOSE, O. **Educação Matemática Crítica – A Questão da Democracia**. Campinas: Papirus, 2001.

SKOVSMOSE, O. **Educação Crítica: Incerteza, Matemática, Responsabilidade**. Tradução de Maria Aparecida Viggiani Bicudo. São Paulo: Cortez Editora, 2007. 304 p.

TRIBUNAL SUPERIOR ELEITORAL. **Fato ou boato: é mentira que Lei de Benford provou fraude nas eleições brasileiras**. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2022/Maio/fato-ou-boato-e-mentira-que-lei-de-benford-provou-fraude-nas-eleicoes-brasileiras>. Acesso em: 8 set. 2023.

UOL. **Vídeo que põe suspeita sobre urnas cita lei matemática que não prova fraude**. São Paulo, 10/10/2018. Disponível em: <https://noticias.uol.com.br/comprova/ultimas-noticias/2018/10/10/video-com-suspeitas-sobre-urnas-cita-lei-matematica-que-nao-prova-fraude.amp.htm?>. Acesso em: 21/09/2023.

SHAMIR SECRET SHARING COMO ELEMENTO MOTIVADOR NO ENSINO DA MATEMÁTICA

Thays Rocha

Laboratório Nacional de Computação Científica (LNCC-MCTI)
tferreira@posgrad.lncc.br

Renato Borseti

Laboratório Nacional de Computação Científica (LNCC-MCTI)
renatojp@posgrad.lncc.br

Fábio Borges

Laboratório Nacional de Computação Científica (LNCC-MCTI)
borges@lncc.br

Resumo: O objetivo deste estudo é empregar a criptografia, mais especificamente o método de *Shamir Secret Sharing*, como uma ferramenta educacional para enriquecer o engajamento nas aulas de matemática. A intenção é utilizar essa abordagem para introduzir, revisar, consolidar e aprofundar conceitos matemáticos, concentrando-se especialmente no ensino do polinômio interpolador de Lagrange e Corpo Finito, base do método criptográfico em questão. Isso proporcionará uma compreensão mais profunda e significativa destes tópicos, ao mesmo tempo em que demonstra uma aplicação prática no cotidiano dos alunos. Além disso, essa abordagem estimula o desenvolvimento do raciocínio lógico, a resolução de problemas e a colaboração em trabalho em equipe.

Palavras-chave: Educação Matemática. Shamir Secret Sharing. Criptografia. Polinômio Interpolador de Lagrange. Corpo Finito.

Introdução

Um dos grandes desafios do ensino da matemática é a motivação dos alunos, principalmente quando se é utilizado as abordagens tradicionais de ensino, como defende D’ambrosio (2007). E como algumas temáticas dentro da matemática são, por si só, abstratas, tal dificuldade fica mais evidente. Pouco se é mostrado da aplicação do que é estudado em aula. E quando isso ocorre, é de maneira superficial.

Conforme a perspectiva de Borges (2008), o educando não irá aprender álgebra booleana ou cálculo, por exemplo, apenas com a justificativa que é a base da computação e a base de uma máquina que calcula seno, respectivamente. Ele precisa ver acontecer, ver uma aplicação daquilo que estuda em sala de aula.

Com isso, muito é debatido acerca do tratamento de temáticas transversais de forma integradora ao currículo, que contribuem para plena formação do educando, explorando a conexão entre os conceitos discutidos em sala de aula e como eles se aplicam na sociedade. Isso ocorre porque esses componentes já estão integrados em vários documentos curriculares da educação básica no Brasil, onde podemos citar a Base Nacional Comum Curricular (BNCC), os Parâmetros Curriculares Nacionais (PCN) e a Lei de Diretrizes e Bases da Educação (LDB), conforme Brasil (2018); Brasil (2008) e Brasil (1996), respectivamente.

Neste cenário, a criptografia entra como uma ferramenta para auxiliar a motivar os alunos no estudo da matemática. Com ela conseguimos abordar temáticas importantes de álgebra, teoria dos números e matemática discreta. Além de contribuir para o aprimoramento de competências essenciais, como raciocínio lógico, resolução de problemas e colaboração no trabalho em equipe.

Portanto, o propósito deste estudo é empregar a criptografia como meio para introduzir, rever, consolidar e até mesmo aprofundar tópicos matemáticos, com foco particular no ensino do polinômio interpolador de Lagrange e Corpo Finito, conferindo-lhe maior relevância e significado.

Sendo assim, as seções subsequentes estão organizadas da seguinte forma: Primeiramente é apresentado um resumo da criptografia e como ela pode ser utilizada como elemento motivador para o ensino. A

Seção seguinte apresenta a relação dos conteúdos matemáticos que se baseia o método criptográfico utilizado, a saber, o polinômio interpolador de Lagrange e Corpo Finito. Para na seção posterior apresentar o método criptográfico de Múltiplas Partes *Shamir Secret Sharing*. E por fim, apresentar as considerações finais e as perspectivas para possíveis trabalhos futuros.

Criptografia

A origem do termo “Criptografia” vem das palavras gregas “kryptós”, que significa “escondido”, e “gráphein”, que significa “escrita”, como visto em Yan (2013). Consiste na habilidade de reescrever mensagens, números e até mesmo efetuar operações matemáticas de maneira incompreensível, tornando-as legíveis somente para o destinatário autorizado a decifrá-las e compreendê-las. E segundo Singh (2014), diferente da esteganografia que tem objetivo de esconder uma informação, a criptografia tem como objetivo ocultar seu significado, ou seja, a mensagem que for transmitida estará toda misturada e sem sentido para quem não conhece o protocolo de decifração.

A criptografia foi primordialmente empregada para fins militares, sendo um exemplo o uso da Máquina Enigma durante o contexto da Segunda Guerra Mundial para a comunicação das forças alemãs, conforme afirma Tavares (2017). Entretanto, atualmente, está presente nas mais diversas atividades do cotidiano, como por exemplo transações bancárias, votações, comunicação, assinatura digital, entre outros.

A criptografia se torna um excelente elemento motivador para o ensino de matemática, porque, conforme afirma Borges (2008), ela consegue abordar desde temáticas mais simples, como o conceito elementar da contagem, como temáticas bem mais complexas, como curvas elípticas e computação quântica, tudo dependendo do método criptográfico utilizado.

Em Tavares (2017) e Olgin e Groenwald (2011), temos a utilização da criptografia direcionado a matemática abordada na educação básica, mais especificamente as temáticas operações matriciais e funções, respectivamente. Assim como em Carvalho (2020), com ênfase em cursos técnicos, temos a utilização da criptografia direcionado a análise combinatória, porcentagem e números binários. Já em Lunkes (2022) e Fusco (2005), os autores abordam a criptografia como motivador no estudo do homomorfismo e das estruturas algébricas, respectivamente, no contexto do ensino superior.

Assim, destacamos a atualidade e a relevância desta temática no dia a dia dos estudantes, abrindo caminho para atividades pedagógicas em matemática que abordem a codificação e decodificação de mensagens. Isso capacita os alunos a desenvolverem tanto a independência para comunicar de maneira cifrada, quanto a habilidade de criar estratégias para resolver desafios.

Relacionando o Conteúdo

Dentre as várias temáticas da matemática em que a criptografia se fundamenta, concentraremos nossa discussão naquela que serve de base para o método criptográfico empregado neste trabalho, a interpolação de Lagrange e Corpo Finito.

A interpolação é um método muito utilizado na matemática para se encontrar uma função $g(x)$ que se aproxima da função $f(x)$, quando a função $f(x)$ possui operações de integração ou diferenciação extremamente difícil, ou até mesmo impossível, ou quando são conhecidos apenas um conjunto de valores numéricos da função dada um conjunto de pontos específicos, como explica Ruggiero e Lopes (1997). Dentre tantos métodos de interpolações disponíveis na literatura, podemos citar o Polinômio Interpolador de Lagrange.

Conforme apresentado em Burden (2016), temos o seguinte teorema para tal método:

Teorema 1: Se $x_0, x_1, \dots, x_n$, são $n + 1$ pontos distintos, e f uma função cujo valores são dados para esses números, então existe um único polinômio $P(x)$ de grau n máximo com

$$f(x_k) = P(x_k) \quad \text{para } k = 0, 1, \dots, n.$$

Este polinômio é dado por

$$P_n(x) = f(x_0)L_{n,0}(x) + \cdots + f(x_n)L_{n,n}(x) = \sum_{k=0}^n f(x_k)L_{n,k}(x), \quad (1)$$

e $L_{n,k}(x)$ é definido por

$$L_{n,k}(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)} \quad (2)$$

$$L_{n,k}(x) = \prod_{i=0, i \neq k}^n \frac{(x-x_i)}{(x_k-x_i)}. \quad (3)$$

Entretanto, assim como aplicado por Burden (2016), iremos utilizar daqui por diante $L_{n,k}(x)$ como $L_k(x)$ com o objetivo de evitar possíveis confusões quanto ao seu grau. Bem como, ao longo deste texto, não faremos distinção entre polinômios e funções polinomiais.

Utilizando o conjunto de pontos $(1,2)$, $(3,4)$, $(5,1)$, $(7,3)$ e $(8,9)$, é possível realizar uma representação geométrica do método em questão, por meio da Figura 1 com base na Equação 1. Em outras palavras, ao calcular cada valor de n , obteremos um correspondente $P_n(x_n)$ e, posteriormente, o valor de $P(x) = \frac{5x^3}{24} - \frac{5x^2}{2} + \frac{199x}{24} - 4$, indicado no gráfico pela cor preta.

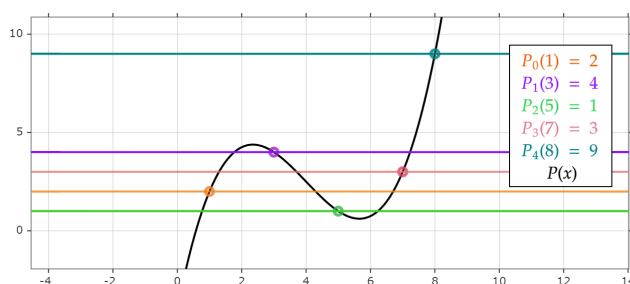


Figura 1: Representação Geométrica do Exemplo do Polinômio Interpolador de Lagrange.

Fonte: autores.

Dizemos que um determinado conjunto não-vazio é um corpo quando se verifica a satisfação das propriedades algébricas das operações de adição e multiplicação, as quais consistem na associatividade, comutatividade, existência de elemento neutro e inversos aditivos para a adição, bem como a associatividade, comutatividade, existência de elemento neutro e inversos multiplicativos, juntamente com a propriedade distributiva em relação à operação de adição, de acordo com Garcia e Lequain (2006). Sendo assim, Corpo Finito consiste em um corpo que possui um conjunto finito de elementos, diferente do conjunto dos números reais $\mathbb{R}$, por exemplo, que possui um número infinito de elementos. Dentre as definições e teoremas referentes a temática, baseada em Howie (2005), podemos destacar:

Teorema 2.1: Seja K um corpo finito, se $|K| = p^m$ para algum primo p e inteiro $m \geq 1$, então todo elemento de K é raiz do polinômio $X^{p^m} - X$ sobre subcorpo primo $\mathbb{Z}_p$.

Teorema 2.2: Seja p um primo e $m \geq 1$ um inteiro, então existe, a menos de isomorfismos, exatamente um corpo de ordem p^m .

Definição 1: Dado p primo e um inteiro $m \geq 1$, o corpo de ordem p^m é chamado de **Corpo de Galois** de ordem p^m , e é denotado por $GF(p^m)$.

Definição 2: Seja $\alpha \in K^*$, onde $K^* = K - \{0\}$ é um corpo finito, define-se a ordem do elemento α como sendo o menor inteiro m tal que $\alpha^m = 1$.

Teorema 3: Seja α um elemento não-nulo do corpo finito $GF(q)$, então $\alpha^{q-1} = 1$.

Com o objetivo de ilustrar as operações de adição e multiplicação definidas, bem como a ordem e elemento primitivo do corpo finito, consideraremos o $GF(5)$ com as operações definidas na Tabela 1 abaixo.

Tabela 1: Operações mod 5

(a) Adição mod 5						(b) Multiplicação mod 5					
+	0	1	2	3	4	.	0	1	2	3	4
0	0	1	2	3	4	0	0	0	0	0	0
1	1	2	3	4	0	1	0	1	2	3	4
2	2	3	4	0	1	2	0	2	4	1	3
3	3	4	0	1	2	3	0	3	1	4	2
4	4	0	1	2	3	4	0	4	3	2	1

Tomando o elemento 3, utilizando a Tabela 1b, temos as Equações (4) a (9), nos oferece a informação de que 4 é a ordem do elemento 3, como afirmam os teoremas e definições citadas anteriormente.

$$3^1 = 3 \quad (4)$$

$$3^2 = 3 \cdot 3 = 4 \mod 5 \quad (5)$$

$$3^3 = 3 \cdot 3^2 = 2 \mod 5 \quad (6)$$

$$3^4 = 3 \cdot 3^3 = 1 \mod 5 \quad (7)$$

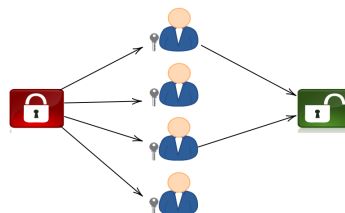
$$3^5 = 3 \cdot 3^4 = 3 \mod 5 \quad (8)$$

$$3^6 = 3 \cdot 3^5 = 4 \mod 5 \quad (9)$$

Shamir Secret Sharing

Na criptografia, temos a área de Computação Segura de Múltiplas Partes, que permite que várias partes colaborem na computação sem revelar suas entradas individuais, mantendo a privacidade das informações enquanto obtêm resultados precisos. Um ótimo exemplo são os métodos de *Secret Sharing*, que particionam um dado sigiloso em partes e distribuem para diferentes entidades. Só é possível reconstruir o dado sigiloso combinando um número suficiente destas partes.

Na Figura 2 temos uma ilustração de um esquema de *Secret Sharing* baseado em um limiar, onde podemos notar que o segredo é particionado em quatro chaves, um para cada participante, onde para recuperação do segredo é necessário a união de pelo menos dois destes. É válido reforçar que, por ser baseado em um limiar, qualquer subconjunto de dois participantes, ou mais, é suficiente para recuperar o segredo.


 Figura 2: Exemplo de *Secret Sharing*, baseado em um limiar.

Fonte: autores.

Dentre os métodos de *Secret Sharing*, podemos citar o *Shamir Secret Sharing*, formulado por Shamir (1979). Nele, a informação privada é chamada de segredo, e ao ser cifrado ele é reescrito em forma de polinômio, com a parte constante sendo o segredo. Ele é definido como (k, n) *Shamir Secret Sharing*, onde k é o quórum, valor mínimo de participantes necessários para recuperar o segredo; e n é o número de

participantes do grupo. Cada participante irá receberá suas respectivas chaves, $Sk_i = (x_i, f(x_i) \bmod p)$ para todo $i \in [1, n]$, ou seja, o valor do seu ponto, necessariamente diferente de zero, e o valor da função no ponto em questão, onde p é primo. Desta forma, asseguramos que qualquer união de $k - 1$, ou menos, não consegue obter qualquer informação referente ao segredo.

A reconstrução do segredo é baseada na interpolação de *Lagrange* sobre o Corpo Finito GF_p , que utilizando, pelo menos, k chaves recupera o polinômio de grau $k - 1$, e ao analisar o polinômio no ponto $x = 0$ iremos recuperar o segredo desejado.

Logo, seja n o número máximo de participantes e GF um corpo finito que contém o domínio dos possíveis segredos, onde a cardinalidade de GF é maior do que n . Sejam $x_1, x_2, \dots, x_n \in GF$ elementos não-nulos, distintos entre si e fixos. Temos:

Compartilhamentos: Dado um segredo $s \in GF$, quem inicia o esquema escolhe os coeficientes $a_1, a_2, \dots, a_{k-1} \in GF$ e define o polinômio $\mathcal{P}(X) = s + \sum_{i=1}^{k-1} a_i X^i$. O compartilhamento, ou *share*, do usuário P_i é $s_i := \mathcal{P}(x_i) \in GF$.

Reconstrução: Quando k usuários reúnem seus segredos $s_{i_1}, s_{i_2}, \dots, s_{i_k}$ eles reconstruem o segredo usando a interpolação polinomial calculando um polinômio único de grau $(k - 1)$ definido por $\mathcal{P}'$ onde o segredo é $\mathcal{P}'(0)$. A reconstrução funciona a partir do momento que $\mathcal{P}' = \mathcal{P}$ e $\mathcal{P}(0) = s$. Nenhum usuário consegue inferir nenhuma informação sobre o segredo a partir de sua *share*.

Exemplo Contextualizado e Proposta de Atividade

A seguir, apresentaremos uma situação hipotética que servirá como ponto de partida para introduzir a atividade proposta à turma, com o objetivo de explorar as temáticas mencionadas anteriormente de forma contextualizada. Além disso, com a atividade proposta, estimular o desenvolvimento do raciocínio lógico e promover o trabalho em equipe como parte fundamental para a resolução do desafio proposto.

Há uma empresa composta por cinco sócios formadores que, em comum acordo, decidiram que para assinar digitalmente qualquer documento referente à empresa seria necessário a aprovação de pelo menos três dos cinco sócios. A senha da empresa para sua assinatura digital é 1947. Aplicando o *Shamir Secret Sharing*, utilizando o primo $p = 20231$, cada sócio receberá uma das seguintes chaves:

$$Sk_1 = (20, 2767) \quad Sk_2 = (51, 7200) \quad Sk_3 = (83, 15808) \quad Sk_4 = (73, 12678) \quad Sk_5 = (31, 3900)$$

Escolhendo de forma aleatória, por exemplo, as chaves Sk_2 , Sk_3 e Sk_5 e aplicando no polinômio interpolador de Lagrange, calcularemos:

$$L_{Sk_2} = \frac{(x - x_3) \cdot (x - x_5)}{(x_2 - x_3) \cdot (x_2 - x_5)} = \frac{(x - 83) \cdot (x - 31)}{(51 - 83) \cdot (51 - 31)} = \frac{-x^2 + 114x - 2573}{640} \quad (10)$$

onde, L_{Sk_3} e L_{Sk_5} são calculadas de forma análoga. Aplicando na Equação (1), teremos então:

$$\begin{aligned} P(x_i) &= 7200 \cdot \left(\frac{-x^2 + 114x - 2573}{640} \right) + 15808 \cdot \left(\frac{x^2 - 82x + 1581}{1664} \right) \\ &\quad + 3900 \cdot \left(\frac{x^2 - 134x + 4233}{1040} \right) \end{aligned} \quad (11)$$

$$P(x) = 2x^2 + x + 1947 \quad (12)$$

Ao analisarmos o polinômio resultante no ponto $x = 0$, iremos obter o valor 1947, senha desejada. De forma análoga, se utilizássemos quatro ou cinco chaves, o procedimento seria o mesmo e chegaríamos no mesmo resultado. Entretanto, se utilizássemos apenas duas chaves, realizando o mesmo procedimento, o resultado que iríamos obter ser o polinômio $143x - 93$ e ao analisarmos no ponto $x = 0$, iremos obter o valor $(-93) \bmod 20231 = 20138$. Valor distinto da senha desejada. Vale ressaltar que cada sócio receberá somente seu conjunto de dados da chave, sem ter acesso às chaves dos outros sócios. Além

disso, mesmo que dois dos cinco sócios se unam para assinar em nome da empresa, a assinatura não será bem-sucedida, pois o método foi configurado para requerer a participação de pelo menos três deles.

Baseado no que foi exposto, e com o objetivo de abordar temáticas de forma colaborativa entre os alunos, podemos sugerir a atividade denominada “Cofre da Turma”. Nesse contexto, o conteúdo armazenado dentro do cofre somente pode ser compartilhado com aqueles que obtiverem acesso ao mesmo. Nesse cenário, o papel do professor assemelha-se ao de um “computador central” encarregado de receber as senhas necessárias para a abertura do cofre. Ele será responsável para analisar as estratégias e os cálculos feitos pelos alunos, que poderão ser auxiliados por softwares, dependendo da turma a ser aplicada. Também ficará a critério do professor, o nível de dificuldade da atividade, sendo um exemplo a não informação do quórum necessário para maior dificuldade na resolução. A turma é dividida em cinco grupos distintos, sendo que cada grupo recebe seu próprio conjunto Sk_i de chaves. A tarefa desses grupos consiste em colaborar entre si, a fim de encontrar a melhor estratégia para combinar suas chaves de modo a abrir o cofre com sucesso.

Os participantes têm a liberdade de explorar diferentes abordagens, incluindo a tentativa de inserção de senhas incorretas, bem como a possibilidade de tentar subjugar um membro de outro grupo, entre outras estratégias, sempre visando identificar a abordagem mais vantajosa. Cada uma dessas situações deve ser analisada a posteriori, uma vez que elas podem gerar discussões significativas em sala de aula. O uso de senhas incorretas, mesmo que o número mínimo necessário de k chaves seja atingido, não revela informação sobre a senha correta, por exemplo. Assim como a vulnerabilidade a diversos tipos de ataques, inclusive de suborno, que ficamos quando confiamos nossas senhas a terceiros. Com isso, pode-se pensar em atividades em que grupos distintos tentam desvendar a senha de outro grupo, por exemplo.

Considerações Finais

O objetivo central deste estudo é complementar a prática pedagógica de ensino da Matemática no processo de ensino-aprendizagem, buscando na interdisciplinaridade uma abordagem que enriqueça a compreensão dos conceitos matemáticos e promova uma aprendizagem mais significativa. Além disso, busca trazer ao debate a importância de trabalhar a matemática de forma mais articulada com a realidade. Desse modo, há a possibilidade de motivar o aluno à aprendizagem, mostrando a importância de aprender matemática. Isso resulta em um enriquecimento da educação matemática, promovendo uma abordagem mais abrangente e relevante para o desenvolvimento dos estudantes.

A criptografia entra como um recurso didático para, além de aguçar a curiosidade, aproximar o educando com problemáticas reais que lhe cercam, como defende Olgin (2011). Então, com métodos criptográficos podemos introduzir, revisar, reforçar e aprofundar alguns conteúdos matemáticos presente em sala de aula, seja na educação básica, técnica ou na educação superior.

O presente trabalho é uma proposta para professores de cursos de graduação que envolvam matemática e computação, tendo em vista que consegue atender muito bem ambos os cursos. A criptografia em si, tem demonstrado seu papel de facilitador na compreensão de temáticas, até então abstratas, em virtude de sua aplicação no cotidiano. No caso do Shamir Secret Sharing, apresentado no presente trabalho, abordamos temáticas da interpolação de Lagrange, bem como o conceito de Corpo Finito com a aritmética modular.

Para trabalhos futuros, temos outros métodos de *Secret Sharing* para aprofundar e utilizar como motivador em sala de aula, tanto para ensino básico como ensino superior. O *Blakley's Scheme*, por exemplo, proposto por Blakley (1979), pode ser utilizado como motivador para o ensino de matrizes para o ensino básico e de para o ensino de hiperplanos e resolução de sistemas lineares para o ensino superior.

Agradecimentos

A CAPES pelo apoio para o desenvolvimento do presente trabalho.

Referências

- BLAKLEY, G. R. Safeguarding cryptographic keys. In: IEEE. 1979 **International Workshop on Managing Requirements Knowledge (MARK)**. [S.l.], 1979. p. 313–318.
- BRASIL. **Lei de Diretrizes e Bases da Educação Nacional**. Lei número 9394, 20 de dezembro de 1996
- BRASIL. MINISTÉRIO DA EDUCAÇÃO. SECRETARIA DA EDUCAÇÃO BÁSICA. **Orientações curriculares para o Ensino Médio: Ciências da Natureza**. Ministério da Educação, vol. 2, 2008.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO. **Base Nacional Comum Curricular**. Brasília, 2018.
- BORGES, F. **Criptografia como Ferramenta para o Ensino de Matemática**. Anais da SBMAC. p. 822, 2008.
- BURDEN, R. L. et al. **Análise numérica**. Cengage Learning, 2016.
- CARVALHO, J. Z. D. S. **A aplicabilidade da criptografia no ensino de matemática no contexto da educação profissional**. 2020. 108 f. Dissertação (Mestrado Profissional em Matemática em Rede Nacional - PROFMAT) – Universidade Federal de Rondônia, Porto Velho, 2020.
- D’AMBROSIO, U. **Educação Matemática: da teoria à prática**. Papirus Editora, 2007.
- FUSCO, C. A. S. **Educação superior: estratégia de aula aliando pesquisa e recursos multimídias**. Associação Nacional de Pós-Graduação e Pesquisa em Educação, v. 28, 2005.
- GARCIA, A.; LEQUAIN, Y. **Elementos de álgebra**. Instituto de Matemática Pura e Aplicada, 2006.
- HOWIE, M. J., **Fields and Galois Theory**, *Springer Undergraduate Mathematics Series*, 2005.
- LUNKES, A. L. Z.; BORGES, F. Sobre a aplicação de homomorfismo na criptografia. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 9, n. 1, 2022.
- OLGIN, C. A., GROENWALD, C. L. O. Criptografia e conteúdos de Matemática do Ensino Médio. In: *IICNEM-Congresso Nacional de Educação Matemática*, 2011
- OLGIN, C. A. **Currículo no Ensino Médio: uma experiência com o tema criptografia**. 2011. 136f. Tese (Doutorado em Ensino de Ciências e Matemática). Universidade Luterana do Brasil, Canoas-RS. 2011.
- RUGGIERO, M. A. G.; LOPES, V. L. D. R. **Cálculo numérico: aspectos teóricos e computacionais**. Makron Books do Brasil, 1997.
- TAVARES, N. P. et al. Criptografia: Uma ferramenta de ensino das operações matriciais. In: *IV CONEDU*, 2017, Joao Pessoa. Anais. Campina Grande: CONEDU, 2017.
- SHAMIR, A. How to share a secret. **Communications of the ACM**, v. 22, n. 11, p. 612-613, 1979.
- SINGH, S., **O livro dos Códigos: A ciência do sigilo - do antigo Egito à Criptografia Quântica**, Record, 10a Edição, 2014.
- YAN, S. Y. **Computational number theory and modern cryptography**. John Wiley & Sons, 2013.

A NÃO EQUIVALÊNCIA ENTRE OS PREÇOS-SOMBRA E AS VARIÁVEIS DUAIS

Beatriz Akiria de Assis Quaresma
b203899@dac.unicamp.br

Antônio Carlos Moretti
amoretti@unicamp.br

Aurelio Ribeiro Leite de Oliveira
aurelio@ime.unicamp.br

Instituto de Matemática, Estatística e Computação Científica - IMECC
Universidade Estadual de Campinas - UNICAMP

Resumo: Ao se trabalhar com interpretação econômica em um problema de programação linear aprendemos que o preço-sombra da i -ésima restrição é o valor da taxa de variação que ocorre na função-objetivo quando se perturba o recurso b_i desta restrição em uma unidade. E, esta perturbação se reflete na i -ésima variável dual, w_i . Consequentemente, igualamos o preço-sombra da i -ésima restrição com a i -ésima variável dual, w_i . Essa afirmação é verdadeira sempre que temos um problema com solução ótima primal não degenerada, contudo, quando o problema tem solução degenerada isto pode não ser verdade. Dessa forma, o objetivo deste trabalho é apresentar de forma mais detalhada este problema e propor duas abordagens para calcular os preços-sombra corretos de um problema de programação linear com solução ótima primal degenerada.

Palavras-chave: Preços-sombra. Programação linear. Solução ótima degenerada.

Introdução

Considere um Problema de Programação Linear (PPL) primal (P) dado por:

$$(P) \quad \begin{cases} \min & z_p = c^T x \\ \text{s.a} & Ax \geq b \\ & x \geq 0 \end{cases} \quad (1)$$

associado a este PPL, (P), temos o PPL dual (D), que pode ser escrito como:

$$(D) \quad \begin{cases} \max & z_d = b^T w \\ \text{s.a.} & A^T w \leq c \\ & w \geq 0 \end{cases} \quad (2)$$

onde $A \in \mathbb{R}^{m \times n}$ é a matriz dos coeficiente, com $\text{posto}(A) = m$, $b \in \mathbb{R}^m$ é o vetor de recursos, $c \in \mathbb{R}^n$ é o vetor de custos, $x \in \mathbb{R}^n$ é o vetor das variáveis do primal e $w \in \mathbb{R}^m$ é o vetor das variáveis duais.

No estudo de programação linear definimos a variação da função-objetivo z_p como a derivada parcial de z_p em relação ao recurso b_i da i -ésima restrição como a i -ésima solução dual relativa a esta restrição (isto é, w_i), ou ainda:

$$\frac{\partial z_p}{\partial b_i} = w_i^* = p_i^+ = p_i^-,$$

onde p_i^+ é o preço-sombra positivo, isto é, a taxa de variação na função-objetivo para o aumento de uma unidade em b_i , e p_i^- é dito o preço-sombra negativo, ou seja, a taxa de variação na função-objetivo

para a redução de uma unidade em b_i . Entretanto, essa afirmação foi refutada por Strum (1969) quando escreveu uma nota sobre os “dois lados” do preço-sombra. O autor constatou, por meio de um exemplo, que o preço-sombra relativo ao “ganho em ter mais uma unidade de um recurso” era diferente do preço-sombra da “perda em ter uma unidade a menos”, em outras palavras, o preço-sombra positivo poderia ser diferente do preço-sombra negativo. Essa diferença aconteceu devido à solução do problema primal ser degenerada, e consequentemente possuir um problema dual com solução ótima alternativa, possibilitando assim diferentes preços-sombra.

Problema Primal Degenerado

Considere o problema apresentado por Strum (1969):

$$\begin{aligned}
 \max \quad & z = 2x_1 + 3x_2 \\
 \text{sujeito a} \quad & 3x_1 + x_2 \leq 48 \\
 & 3x_1 + 4x_2 \leq 120 \\
 & x_1 + 2x_2 \leq 56 \\
 & x_j \geq 0, \quad j = 1, 2
 \end{aligned} \tag{3}$$

A resolução gráfica do problema (3) pode ser vista na Figura 1. Com o gráfico é possível notar que a segunda restrição é fracamente redundante e ocasiona a degenerescência do problema.

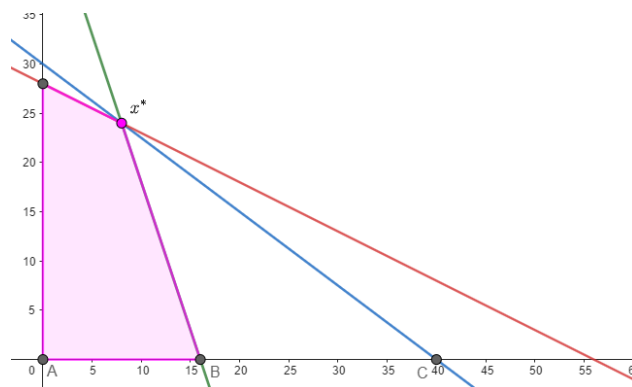


Figura 1: Representação gráfica do problema degenerado proposto por Strum (1969).

As Equações Básicas (EB) do problema em uma das bases ótimas são:

$$\begin{aligned}
 x_1 &= 8 - x_4 + 2x_5 \\
 x_2 &= 24 + \frac{1}{2}x_4 - \frac{3}{2}x_5 \\
 x_3 &= 0 + \frac{5}{2}x_4 - \frac{9}{2}x_5 \\
 (\max) z &= 88 - \frac{1}{2}x_4 - \frac{1}{2}x_5
 \end{aligned} \tag{4}$$

Com isso, o problema primal (3) tem solução ótima única $x^* = (8, 24)^t$, com o valor ótimo $z^* = 88$. Porém, como podemos observar pelas EB temos a variável básica $x_3 = 0$, ou seja, o problema primal tem solução ótima degenerada. Inicialmente, se quiséssemos calcular os preços-sombra do problema, assumiríamos o resultado das variáveis duais, isto é, $w^* = (0, \frac{1}{2}, \frac{1}{2})^t$. Inclusive, utilizando o solver do Excel temos o mesmo resultado, como é apresentado na Figura 2. Dessa forma, se planejássemos aumentar

uma unidade do recurso relativo a segunda restrição, por exemplo, concluiríamos que a variação no valor ótimo da função-objetivo seria de $1/2$, o que é equivocado, uma vez que não há variação no valor ótimo. Isso acontece pelo fato do solver emitir a solução dual, que corresponde a primeira base encontrada que satisfaz o critério de otimalidade, como preço-sombra e não busca por uma base que corresponde a um aumento no recurso e uma outra base que corresponde a redução do recurso.

Microsoft Excel 16.0 Relatório de Sensibilidade
Planilha: [Pasta1]Planilha1
Relatório Criado: 30/08/2023 15:32:47

Restrições

		Final	Sombra	Restrição	Permitido	Permitido
Célula	Nome	Valor	Preço	Lateral R.H.	Aumentar	Reduzir
\$E\$5	le	48	0	48	1E+30	0
\$E\$6	le	120	0,5	120	0	8
\$E\$7	le	56	0,5	56	4	0

Figura 2: Relatório de sensibilidade do problema do Strum (1969) emitido pelo solver do Excel.

Cálculo dos Preços-Sombra

Conhecendo o problema apresentado na seção anterior, apresentaremos duas abordagens para calcular corretamente os preços-sombra. A primeira, inspirada no trabalho do Akgül (1984), resolvendo PL bem menores que o inicial e a segunda utilizando a Análise Paramétrica.

Abordagem 1

Assumindo que o problema primal tenha solução ótima degenerada, é conhecido que o problema dual associado possui soluções ótimas alternativas. Visto que os preços-sombra estão associados a solução ótima do problema dual, que nesse caso é alternativa, é necessário caracterizar a região de todas as soluções ótimas do dual, e a partir disso calcular os verdadeiros preços-sombra.

Descrevemos a região de todas as soluções alternativas como $\hat{w} = w^* + D\lambda \geq 0$, onde D é a matriz cujas colunas são as direções simplex das variáveis não básicas, com os custos reduzidos iguais a zero, a partir do ponto ótimo w^* . Observe que os pontos $\hat{w}$, dados por λ , pertencem a toda a região formada pela combinação linear positiva das direções simplex, porém, é preciso manter a factibilidade do problema, por isso impomos que $\hat{w} \geq 0$. Contudo, estas informações são relacionadas a solução do problema dual e queremos formar a região absorvendo os dados diretamente do problema primal, sem criar o problema dual e resolvê-lo.

O vetor w^* é a solução do dual, com isso, sabemos que as variáveis estruturais do dual são os custos reduzidos das variáveis de folga do primal, já as variáveis de folga duais são dadas pelos custos reduzidos das variáveis estruturais do primal. Agora, para as direções extremas desse ponto, inicialmente suponha que a variável básica x_k do problema primal no ótimo seja degenerada, com isso $\bar{b}_k = 0$, este valor se transforma no custo reduzido da variável w_k no problema dual associado, o que vai indicar que a solução ótima é alternativa. Sabemos que a direção simplex desta variável dual w_k se encontra na coluna do custo reduzido igual a zero. Dessa forma, nas EB do problema primal, a direção simplex fica na linha da variável degenerada. Porém, devemos nos atentar com a ordem dos componentes do vetor, uma vez que seguem as posições das variáveis duais com relação aos custo reduzidos.

Sabendo disso, vamos caracterizar a região formada pelas soluções ótimas alternativas do problema proposto por Strum (1969). Queremos recuperar as informações de $\hat{w} = w^* + D\lambda$ pelas EB determinadas

em (4), para tanto vamos considerar as linhas degeneradas e a linha da função-objetivo:

$$\begin{array}{ccc|ccc} 0x_1 + 0x_2 & & +1x_3 - \frac{5}{2}x_4 + \frac{9}{2}x_5 & = & 0 \\ 0x_1 + 0x_2 & & +0x_3 + \frac{1}{2}x_4 + \frac{1}{2}x_5 & = & 88 \\ w_4 & w_5 & & w_1 & w_2 & w_3 \end{array}$$

Consequentemente,

$$w^* = \left(0, \frac{1}{2}, \frac{1}{2}, 0, 0\right)^t \text{ e } D = \left(1, -\frac{5}{2}, \frac{9}{2}, 0, 0\right)^t.$$

Assim, definimos a região formada por todas as soluções ótimas do dual por:

$$\hat{w} = w^* + D\lambda = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{5}{2} \\ \frac{9}{2} \\ 0 \\ 0 \end{bmatrix} \lambda = \begin{bmatrix} \lambda \\ \frac{1}{2} - \frac{5}{2}\lambda \\ \frac{1}{2} + \frac{9}{2}\lambda \\ 0 \\ 0 \end{bmatrix} \geq 0.$$

Ou seja, temos o seguinte sistema de inequações, já removendo as inequações redundantes:

$$(\Lambda) \quad \begin{cases} \lambda \leq \frac{1}{5} \\ \lambda \geq 0 \end{cases}.$$

Conhecendo a região das soluções ótimas do dual, que chamaremos de W^* , Akgül (1984) mostra como podemos calcular os preços-sombras corretos de um problema primal dado no formato canônico de minimização, (P) . Definindo $F(b) = \min \{c^t x : Ax \geq b, x \geq 0\}$ e a derivada direcional de F em b na direção de um vetor $u \in \mathbb{R}^m$ por $D_u F(b) = \max \{u^t w; w \in W^*\}$, o autor descreve o preço-sombra positivo da i -ésima restrição como a taxa de variação de $F(b)$ para um aumento de uma unidade em b_i por:

$$p_i^+ = D_{e_i} F(b) = \max \{e_i^t w; w \in W^*\}.$$

E o preço-sombra negativo da i -ésima restrição, isto é, a taxa de variação de $F(b)$ para uma redução de uma unidade em b_i fica:

$$p_i^- = D_{-e_i} F(b) = \max \{(-e_i)^t w; w \in W^*\} = \min \{e_i^t w; w \in W^*\}.$$

Com isso, para calcular os preços-sombra positivo e negativo de cada restrição precisamos apenas resolver esses dois pequenos PPL. Diante disso, vamos calcular os preços-sombras do problema do Strum (1969), entretanto, como o problema (3) não está no formato canônico de minimização os preços-sombra são dados por:

$$p_i^+ = D_{e_i} F(b) = \min \{e_i^t w; w \in W^*\} \text{ e } p_i^- = D_{-e_i} F(b) = \max \{(-e_i)^t w; w \in W^*\}.$$

Logo, os preços-sombra positivo e negativo corretos de cada restrição são:

• p_1^+ :

$$\begin{array}{ll} \min & z_\lambda = \lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = 0$. Assim, $p_1^+ = 0$.

- p_1^- :

$$\begin{array}{ll} \max & z_\lambda = \lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = \frac{1}{5}$. Assim, $p_1^- = \frac{1}{5}$.

- p_2^+ :

$$\begin{array}{ll} \min & z_\lambda = \frac{1}{2} - \frac{5}{2}\lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = \frac{1}{5}$. Dessa forma, $p_2^+ = 0$.

- p_2^- :

$$\begin{array}{ll} \max & z_\lambda = \frac{1}{2} - \frac{5}{2}\lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = 0$. Dessa forma, $p_2^- = \frac{1}{2}$.

- p_3^+ :

$$\begin{array}{ll} \min & z_\lambda = \frac{1}{2} + \frac{9}{2}\lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = 0$. Portanto, $p_3^+ = \frac{1}{2}$.

- p_3^- :

$$\begin{array}{ll} \max & z_\lambda = \frac{1}{2} + \frac{9}{2}\lambda \\ \text{sujeito a} & \lambda \in (\Lambda) \end{array}$$

cuja solução é $\lambda^* = \frac{1}{5}$. Portanto, $p_3^- = \frac{7}{5}$.

Em resumo, os preços-sombra corretos do problema proposto pelo Strum (1969) são:

$$\begin{array}{lll} p_1^+ & = 0 & p_2^+ & = 0 & p_3^+ & = \frac{1}{2} \\ p_1^- & = \frac{1}{5} & p_2^- & = \frac{1}{2} & p_3^- & = \frac{7}{5} \end{array}$$

Abordagem 2

Outra forma para encontrar os preços-sombra de acréscimo, p_i^+ , e decréscimo, p_i^- , é utilizando a abordagem da Análise Paramétrica (AP). Nesse caso, perturbamos isoladamente cada restrição i com $i = 1, 2, \dots, m$ ao longo das direções $d = +e_i$, para o preço-sombra positivo e $d = -e_i$, para o preço-sombra negativo, isto é, $b + \lambda d$ com $\lambda \geq 0$.

Para tal, definimos o conjunto $S = \{i : \bar{d} < 0\}$, onde $\bar{d} = B^{-1}d$, e verificamos que se $S = \emptyset$ então a base atual é ótima para todos os valores de $\lambda \geq 0$ paramos com o processo da AP e calculamos o preço-sombra, caso contrário, encontramos um intervalo crítico, $[0, \hat{\lambda}]$, no qual a base atual permanece ótima, de modo que $\hat{\lambda} = \min_{i \in S} \left\{ \frac{\bar{b}_i}{-\bar{d}_i} \right\}$, com $\bar{b} = B^{-1}b$. Ainda, como estamos interessados em descobrir apenas o p_i^+ e o p_i^- de cada restrição i , precisamos encontrar apenas o primeiro intervalo crítico, ou seja $\lambda = 1$. Porém,

devemos tomar cuidado quando $\hat{\lambda} = 0$, pois nesse caso não é possível fazer nenhum movimento com a base atual, portanto, temos que mudar a base pivoteando em cima da variável degenerada (utilizamos o Método Dual Simplex (MDS) para não romper a otimalidade do primal), consequentemente encontrando uma base ótima alternativa.

Assim, utilizando a AP com as devidas ressalvas, podemos encontrar os preços-sombra corretos de problemas de programação linear com solução ótima primal degenerada, como é o caso do problema do Strum (1969). Pelas EB (4) conseguimos representar a base ótima que a represente, logo:

$$B = [a^1, a^2, a^3] \quad \text{e} \quad B^{-1} = \begin{bmatrix} 0 & 1 & -2 \\ 0 & -\frac{1}{2} & \frac{3}{2} \\ 1 & -\frac{5}{2} & \frac{9}{3} \end{bmatrix}.$$

Agora, calculando os preços-sombra com a Análise Paramétrica temos:

- p_1^+ : Utilizando a direção $d = e_1 = (1, 0, 0)^t$, encontramos:

$$\bar{d} = B^{-1}d = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \implies S = \emptyset.$$

Como $S = \emptyset$, segue que a base atual é ótima para todos os valores de $\lambda \geq 0$. Calculando a variação no valor ótimo do problema vale que:

$$z(\lambda) = z^* + \lambda c_B^t \bar{d} = 8 + \lambda(2, 3, 0) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 8 + 0\lambda,$$

o que mostra que quando $\lambda = 1$ não ocorre variação na função-objetivo (nesse caso qualquer $\lambda \geq 0$ não afetará o valor ótimo inicial), ou seja, $p_1^+ = 0$.

- p_1^- : Para p_1^- usamos a direção $d = -e_1 = (-1, 0, 0)^t$, com isso:

$$\bar{d} = B^{-1}d = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \implies S = \{3\} \implies \hat{\lambda} = \min \left\{ \frac{0}{\frac{1}{2}} \right\} = 0.$$

Como $\hat{\lambda} = 0$ não é possível “andar” na direção d , por essa razão precisamos mudar a base ótima, isso significa que a terceira variável básica deve sair da base, ou seja, x_3 sairá da base. Pelo teste da razão do MDS temos $\min \left\{ \frac{\frac{1}{2}}{\frac{1}{5}} \right\} = \frac{1}{5}$, então a variável x_4 deve entrar na base. Assim, as novas EB no ótimo são dadas por:

$$\begin{aligned} x_1 &= 8 - \frac{2}{5}x_3 + \frac{1}{5}x_5 \\ x_2 &= 24 + \frac{1}{5}x_3 - \frac{3}{5}x_5 \\ x_4 &= 0 + \frac{2}{5}x_3 + \frac{9}{5}x_5 \\ (\max) z &= 88 - \frac{1}{5}x_3 - \frac{7}{5}x_5 \end{aligned} \tag{5}$$

Portanto, a nova base é dada por:

$$B = [a^1, a^2, a^4] \quad \text{e} \quad B^{-1} = \begin{bmatrix} \frac{2}{5} & 0 & -\frac{1}{5} \\ -\frac{1}{5} & 0 & \frac{3}{5} \\ -\frac{2}{5} & 1 & -\frac{9}{5} \end{bmatrix}.$$

Dessa forma, calculando $\bar{d}$ com a nova base, temos:

$$\bar{d} = B^{-1}d = \begin{bmatrix} -\frac{2}{5} \\ \frac{1}{5} \\ \frac{2}{5} \end{bmatrix} \Rightarrow S = \{1\} \Rightarrow \hat{\lambda} = \min \left\{ \frac{8}{\frac{2}{5}} \right\} = 20.$$

Como $\hat{\lambda} \neq 0$ e precisamos apenas de $\lambda = 1$, podemos calcular o preço-sombra negativo da primeira restrição, ou seja:

$$z(\lambda) = z^* + \lambda c_B^t \bar{d} = 8 + \lambda (2, 3, 0) \begin{bmatrix} -\frac{2}{5} \\ \frac{1}{5} \\ \frac{2}{5} \end{bmatrix} = 8 - \frac{1}{5}\lambda,$$

o que mostra que quando $\lambda = 1$ a variação na função objetivo é um decréscimo de $\frac{1}{5}$, isto é, $p_1^- = \frac{1}{5}$. Vale mencionar que nesse caso poderíamos verificar a variação na função-objetivo com λ até 20, uma vez que a base continua ótima dentro desse intervalo.

Por fim, continuando esse processo nas demais direções encontraríamos os outros preços-sombra.

Conclusões

Neste trabalho, mostramos que quando o problema primal é degenerado o preço-sombra de aumentar a disponibilidade de um recurso pode ser diferente do preço-sombra de reduzir a disponibilidade do mesmo recurso, e que o preço-sombra não é equivalente a solução ótima dual quando temos solução ótima primal degenerada. Vale destacar que apesar de exemplificar o erro gerado com o solver do Excel, esse erro também é encontrado por outros solvers mais robustos de PL (como CPLEX, GUROBI, etc). Por fim, ainda apresentamos duas formas para calcular os valores dos preços-sombra corretamente. A Abordagem 1, que emite o resultado do preço-sombra resolvendo um PPL bem menor, porém, precisamos caracterizar a região formada por todas as soluções ótimas alternativas, além de resolver até $2m$ problemas de programação linear (se quisermos saber todos os preços-sombra do PPL). Já a Abordagem 2, além de encontrarmos os preços-sombra corretos, indica o intervalo crítico no qual a variação é válida. Entretanto, em muitos casos precisamos realizar vários pivoteamentos, gerando um custo computacional maior.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Referências

- AKGÜL, M. A Note on Shadow Prices in Linear Programming. **Journal of the Operational Society**, v. 35, n. 5, p. 425–431, 1984.
- STRUM, J. E. Note on "Two-Sided Shadow Prices". **Journal of Accounting Research**, v. 7, n. 1, p. 130–162, 1969.

ANÁLISE DA PERFORMANCE DE ALGUNS MÉTODOS CLÁSSICOS DE CONSTRUÇÃO DE ROTAS PARA O PROBLEMA DO CAIXEIRO VIAJANTE

Luiz Fernando Nunes

Universidade Tecnológica Federal do Paraná

nunes@utfpr.edu.br

Resumo: Este trabalho apresenta uma comparação da performance de diversos algoritmos de construção de rotas para o clássico Problema do Caixeiro Viajante. A análise levou em conta a qualidade das soluções e o tempo computacional requerido. Utilizou-se 20 instâncias da TSP-Library, onde aplicou-se 9 técnicas aproximadas para a obtenção das soluções. Sabe-se que são inúmeras as técnicas de alta performance que têm surgido para resolver o referido problema. Apesar disso, técnicas significativamente mais simples, de fácil implementação e que exigem baixo esforço computacional, são muitas vezes utilizadas como soluções iniciais, onde outros métodos para melhoria de rotas podem ser aplicados, ou utilizadas na construção de populações de soluções iniciais para técnicas mais sofisticadas, como algoritmos genéticos. Comparações entre algoritmos são sempre complicadas, mas os resultados aqui obtidos podem dar uma boa ideia, para a escolha de um método não exato, capaz de gerar boas soluções aproximadas, em um tempo computacional aceitável.

Palavras-chave: Problema do Caixeiro Viajante. Pesquisa Operacional. Otimização

Introdução

O Problema do Caixeiro Viajante (PCV) consiste em estabelecer uma única rota que passe por cada nó de um grafo, uma e apenas uma vez, retornando ao nó inicial no final do percurso. Este roteiro Hamiltoniano deve ser feito de modo que a distância total percorrida seja mínima.

Não é conhecido exatamente quem foi o precursor do termo “Caixeiro Viajante” (SOUZA, 1997), mas o problema popularizou-se rapidamente, sendo inúmeras as aplicações para o mesmo (BENEVIDES, 2011), (COOK, 2014), (SOUZA, 1997).

Para a obtenção de soluções exatas para o problema, existem diversas formulações de programação matemática. Um modelo de programação linear inteira binária apresentado por (BODIN et. al., 1983) é descrito na sequência:

Seja o grafo $G(V, A)$ onde V representa o conjunto de vértices $|V| = n$, e A o conjunto de arcos. Se for admitido que os custos ou distâncias mínimas entre os nós da rede são dados pela matriz $C = [c_{ij}]$ simétrica, isto é, considera-se que $c_{ij} = c_{ji}$, assumindo ainda que $c_{ii} = +\infty, \forall i \in \mathbb{N}$ e considerando a matriz $X = [x_{ij}]$ das variáveis de decisão do problema, onde:

$$x_{ij} = \begin{cases} 1 & \text{se o arco } i - j \text{ está na rota} \\ 0 & \text{caso contrário} \end{cases}$$

tem-se a seguinte formulação de programação linear inteira binária para o problema

$$\min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (1)$$

sujeito a:

$$\sum_{i=1}^n x_{ij} = 1 \quad (j = 1, \dots, n) \quad (2)$$

$$\sum_{j=1}^n x_{ij} = 1 \quad (i = 1, \dots, n) \quad (3)$$

$$X = (x_{ij}) \in S \quad (4)$$

$$x_{ij} = 0 \text{ ou } 1 \quad (i, j = 1, \dots, n) \quad (5)$$

A função objetivo é representada pela equação (1). Os dois primeiros grupos de restrições (2) e (3), garantem que exatamente um arco (i, j) emana de cada nó da rota e exatamente um arco (i, j) é direcionado para cada nó j da rota. Nas restrições (4), o conjunto S pode ser qualquer restrição que evite a formação de subrotas. Estas restrições podem ser, entre outras:

$$S = \{(x_{ij}) : \sum_{i \in Q} \sum_{j \notin Q} x_{ij} \geq 1 \text{ para todo subconjunto não vazio } Q \text{ de } N\};$$

$$S = \{(x_{ij}) : \sum_{i \in R} \sum_{j \in R} x_{ij} \leq |R| - 1 \text{ para todo subconjunto não vazio } R \text{ de } \{2, 3, \dots, n\}\};$$

$$S = \{(x_{ij}) : y_i - y_j + nx_{ij} \leq n - 1 \text{ para } 2 \leq i \neq j \leq n \text{ para algum número real } y_i\};$$

As últimas restrições (5), indicam que as variáveis são binárias (0 ou 1).

Considerando a grande dificuldade na aplicação de procedimentos exatos, como o descrito anteriormente, baseados em programação matemática, muitas técnicas aproximadas, heurísticas e metaheurísticas, têm surgido nas últimas décadas. Entre estes procedimentos, podemos citar *Simulated Annealing* (BUDINICH, 1996), Redes Neurais (SIQUEIRA, 2005), Colônia de Formigas (DORIGO e GARAMBARDELLA, 1997), Busca Tabu (KNOX, 1994), Algoritmos Genéticos (JIN et al., 2003), Também comparações entre outros algoritmos heurísticos inspirados na natureza são encontrados na literatura (STUTZLE et al., 2000).

Neste trabalho, utilizou-se 9 técnicas de construção de rotas, descritas em (BODIN et. al. 1983). São procedimentos que variam dos mais simples, com características *greedy* e decisões estocásticas, até métodos mais complexos, que iniciam com uma envoltória convexa dos pontos representados pelos vértices do grafo associado ao problema, seguido de metodologias para a inserção no roteiro dos pontos que restaram no interior do polígono convexo formado.

Como justificativa para o uso destes procedimentos mais simples, que nem sempre fornecem soluções de qualidade, pode-se citar a existência de algoritmos para melhorias de rotas, que permitem entre outras coisas, desfazer cruzamentos de arcos, promovendo muitas vezes melhorias bastante significativas nos resultados. Entre estas técnicas de melhorias de rotas, pode-se citar os métodos conhecidos como 2-Opt e 3-Opt, que são frequentemente aplicados a partir de soluções iniciais (BIANCHI e BOWLER, 2005), (BODIN et. al. 1983), etc.

Estas soluções aproximadas, de rápida obtenção, podem também ser utilizadas para formação de populações iniciais para uso em algoritmos genéticos, ou outras técnicas.

Algoritmos utilizados

Na sequência, são apresentados os 9 algoritmos de construção de rotas utilizados neste trabalho. Em cada uma destas técnicas, são descritos os passos, o comportamento das técnicas no pior caso, bem como as complexidades computacionais dos mesmos (BODIN et. al. 1983).

Para facilitar a apresentação dos resultados nas tabelas, figuras e conclusões, os nomes destes procedimentos serão referidos pelas siglas em negrito, que seguem os nomes dos algoritmos apresentados abaixo.

1 – Vizinho mais próximo (**tspviz**):

P1. Escolha um nó inicial.

P2. Encontre o nó mais próximo do último nó adicionado na rota e adicione-o na rota.

P3. Repita o passo 2 até que todos os nós tenham sido incluídos na rota.

P4. Una o primeiro nó ao último nó da rota.

$$\text{Comportamento do algoritmo no pior caso: } \frac{\text{Comprimento do roteiro obtido}}{\text{Comprimento do roteiro ótimo}} \leq \frac{1}{2} \cdot [\log_2 n] + \frac{1}{2}$$

onde $[X]$ é o menor número inteiro $\geq X$ e n é o número de nós da rede.

Complexidade computacional: n^2

Observação: O algoritmo pode ser repetido n vezes, sendo que em cada vez, um novo vértice é escolhido para iniciar o processo. Neste caso, a complexidade computacional será n^3 .

2 – Inserção do mais próximo (**tspimp**):

- P1. Inicie com um subgrafo contendo apenas o nó i .
- P2. Encontre o nó k tal que c_{ik} seja mínimo e forme a subrota $i - k - i$.
- P3. Dada a subrota, encontre o nó k não pertencente a subrota, mais próximo de qualquer nó da subrota.
- P4. Encontre o arco (i, j) na subrota que minimiza $c_{ik} + c_{kj} - c_{ij}$ e insira k entre i e j .
- P5. Volte ao passo 3 até que todos os nós tenham sido incluídos na rota.

Comportamento do algoritmo no pior caso: $\frac{\text{Comprimento do roteiro obtido}}{\text{Comprimento do roteiro ótimo}} \leq 2$

Complexidade computacional: n^2

3 – Inserção do mais barato (**tspimb**):

- P1. Inicie com um subgrafo contendo apenas o nó i .
- P2. Encontre o nó k tal que c_{ik} seja mínimo e forme a subrota $i - k - i$.
- P3. Encontre o arco (i, j) na subrota e o nó k não pertencente a subrota, tais que: $c_{ik} + c_{kj} - c_{ij}$ seja mínimo. Insira k entre i e j .
- P4. Volte ao passo 3 até que todos os nós tenham sido incluídos na rota.

Comportamento do algoritmo no pior caso: é o mesmo do algoritmo **tspimp**.

Complexidade computacional: $n^2 \cdot \log_2 n$

4 – Inserção arbitrária (**tspiar**):

- P1. Inicie com um subgrafo contendo apenas o nó i .
- P2. Encontre o nó k tal que c_{ik} seja mínimo e forme a subrota $i - k - i$.
- P3. Selecione arbitrariamente um nó k não pertencente a subrota.
- P4. Insira o nó k entre i e j do arco tal que $c_{ik} + c_{kj} - c_{ij}$ seja mínimo.
- P5. Volte ao passo 3 até que todos os nós tenham sido incluídos na rota.

Comportamento do algoritmo no pior caso: $\frac{\text{Comprimento do roteiro obtido}}{\text{Comprimento do roteiro ótimo}} \leq \lceil \log_2 n \rceil + 1$

Complexidade computacional: n^2

5 – Inserção do mais distante (**tspimd**):

- P1. Inicie com um subgrafo contendo apenas o nó i .
- P2. Encontre o nó k tal que c_{ik} seja máximo e forme a rota $i - k - i$.
- P3. Dada a subrota, encontre o nó k não pertencente a subrota mais distante de qualquer nó da subrota.
- P4. Encontre o arco (i, j) na subrota que minimiza $c_{ik} + c_{kj} - c_{ij}$ e insira k entre i e j .
- P5. Volte ao passo 3 até que todos os nós tenham sido incluídos na rota.

Comportamento do algoritmo no pior caso: é o mesmo do algoritmo **tspiar**.

Complexidade computacional: não citado em (BODIN et. al., 1983).

6 – Inserção mais rápida (**tspimr**):

- P1. Tome um nó inicial para formar um circuito T com 1 nó e 0 arcos.
- P2. Dado o conjunto T_k , ache o nó z_k não pertencente a T_k mais próximo de um nó y_k em T_k .
- P3. Seja T_{k+1} a rota com $k + 1$ nós, obtida inserindo z_k imediatamente em frente a y_k .
- P4. Repita P2 e P3 até que todos os nós tenham sido incluídos na rota.

Comportamento do algoritmo no pior caso: é o mesmo do algoritmo **tspimp**.

Complexidade computacional: é a mesma do algoritmo **tspimp**.

7 – Inserção na envoltória Convexa (**tspchi**):

- P1. Forme uma envoltória convexa com os pontos do problema. Esta envoltória é uma subrota inicial.

P2. Para cada nó k que ainda não está na subrota, decidir entre que nós i e j da subrota, o nó k será inserido. Assim, para cada nó k , encontre (i, j) tal que $c_{ik} + c_{kj} - c_{ij}$ seja mínimo.

P3. Para todos os (i, k, j) encontrados no passo 2, determine (i^*, k^*, j^*) tal que: $(c_{i^*k^*} + c_{k^*j^*})/c_{i^*j^*}$ seja mínimo.

P4. Inserir k^* na subrota entre os nós i^* e j^* .

P5. Repetir os passos P2 até P4, até que um circuito Hamiltoniano seja obtido.

Comportamento do algoritmo no pior caso: desconhecido.

Complexidade computacional: maior que a complexidade computacional do algoritmo **tspimb**.

8 – Inserção do maior ângulo (**tspgai**):

P1. Forme uma envoltória convexa com os pontos do problema. Esta envoltória é uma subrota inicial.

P2. Encontre o nó k^* fora da subrota e o arco (i^*, j^*) na subrota, tal que o ângulo formado pelos arcos (i^*, k^*) e (k^*, j^*) é máximo.

P3. Inserir k^* na sub-rota entre os nós i^* e j^* .

P4. Repetir os passos P2 até P3, até que um circuito Hamiltoniano seja obtido.

Comportamento do algoritmo no pior caso: desconhecido.

Complexidade computacional: é a mesma do algoritmo **tspimb**.

9 – Diferença $\times$ raio de inserção (**tspdri**):

P1. Forme uma envoltória convexa com os pontos do problema. Esta envoltória é uma subrota inicial.

P2. Encontre o nó k^* fora da subrota e o arco (i^*, j^*) na subrota, tal que o produto: $(c_{i^*k^*} + c_{k^*j^*} - c_{i^*j^*}) \times (c_{i^*k^*} + c_{k^*j^*})/c_{i^*j^*}$ é máximo.

P3. Inserir k^* na subrota entre os nós i^* e j^* .

P4. Repetir os passos P2 até P3, até que um circuito Hamiltoniano seja obtido.

Comportamento do algoritmo no pior caso: desconhecido.

Complexidade computacional: é a mesma do algoritmo **tspimb**.

Os testes computacionais e resultados

Os algoritmos aproximados foram todos implementados utilizando a linguagem Python 3.8.3. Utilizou-se um computador notebook, com processador Intel(R) Core (TM) i7-8550U, 8th Gen. com 16GB de memória RAM e sistema operacional Windows 11.

Para testar os algoritmos, extraiu-se 20 instâncias do TSP-Library (REINELT, 1990) e (REINELT, 1991). Estas instâncias são referidas nas tabelas por: **att48**, **eil51**, **berlin52**, **st70**, **eil76**, **rat99**, **kroD100**, **kroE100**, **eil101**, **pr136**, **kroA150**, **kroB150**, **kroA200**, **kroB200**, **ts225**, **att532**, **rat575**, **pr1002**, **vm1084** e **rl1304**, com uma quantidade de nós variando de 48 até 1304 nós.

Na tabela 1 que segue, são apresentadas as 20 instâncias, com as quantidades de nós e seus respectivos valores ótimos. Nas colunas que seguem, são exibidos os resultados obtidos para as técnicas abreviadas por **tspviz**, **tspimp**, **tspimb**, **tspiar** e **tspimd**.

Tabela 1: Os resultados computacionais para as 20 instâncias escolhidas, utilizando os algoritmos de construção de rotas: **tspviz**, **tspimp**, **tspimb**, **tspiar** e **tspimd**.

Instâncias TSP-Library	Nós	Valor ótimo	tspviz	tspimp	tspimb	tspiar	tspimd
att48	48	33.523,709	42.626,505	37.984,53	36.246,871	36.181,941	36.687,026
eil51	51	429,118	539,995	484,978	496,252	447,321	489,803
berlin52	52	7.544,366	9.223,293	8.878,75	9.216,761	8.442,962	8.116,906
st70	70	677,871	794,547	773,065	784,528	698,966	750,058
eil76	76	545,388	636,317	618,084	623,391	596,061	609,830
rat99	99	1.219,244	1.416,735	1.465,88	1.484,801	1.385,282	1.316,581

kroD100	100	21.294,291	29.201,458	25.859,91	25.216,098	24.166,506	23.201,631
kroE100	100	22.068,759	24.907,022	27.380,66	25.491,687	22.971,698	24.799,114
eil101	101	640,970	762,034	736,369	727,965	689,415	724,091
pr136	136	96.770,925	122.777,119	107.844,18	110.320,800	105.562,592	107.482,415
kroA150	150	26.525,032	33.603,035	32.215,80	30.955,040	29.052,659	30.394,054
kroB150	150	26.127,358	34.835,250	31.619,76	30.806,591	28.624,327	29.301,653
kroA200	200	29.369,408	38.996,468	35.561,22	34.995,760	31.799,054	32.830,935
kroB200	200	29.440,413	37.466,144	36.276,30	35.858,749	32.484,722	33.811,724
ts225	225	126.645,934	152.894,347	149.925,13	173.628,128	137.568,966	159.404,527
att532	532	86.742,423	110.016,151	106.858,41	101.915,967	96.758,297	100.723,643
rat575	575	6.798,256	8.342,289	8.272,43	8.047,029	7.563,529	8.258,420
pr1002	1002	259.066,664	331.303,658	306.943,61	305.129,802	292.695,612	321.385,601
vm1084	1084	239.281,871	306.630,607	281.237,92	278.359,127	271.545,789	289.198,013
rl1304	1304	252.940,699	332.677,692	317.570,01	309.626,291	301.779,130	323.252,341

Na tabela 2 que segue, são apresentadas as 20 instâncias, com as quantidades de nós e seus respectivos valores ótimos. Nas colunas seguintes, são exibidos os resultados obtidos pelas técnicas abreviadas por **tspimr**, **tspchi**, **tspgai** e **tspdri**.

Tabela 2: Os resultados computacionais para as 20 instâncias, utilizando os algoritmos de construção de rotas: **tspimr**, **tspchi**, **tspgai** e **tspdri**.

Instâncias TSP-Library	Nós	Valor ótimo	tspimr	tspchi	tspgai	tspdri
att48	48	33.523,709	45.077,358	35.759,278	35.802,856	34.542,296
eil51	51	429,118	577,468	454,301	460,442	466,262
berlin52	52	7.544,366	10.241,823	8.147,391	7.725,475	8.051,149
st70	70	677,871	887,029	755,342	739,362	707,997
eil76	76	545,388	748,644	595,346	609,600	602,090
rat99	99	1.219,244	1.728,817	1.395,179	1.290,181	1.267,755
kroD100	100	21.294,291	29.330,050	22.990,690	23.637,089	21.672,288
kroE100	100	22.068,759	32.581,357	24.539,101	23.522,965	23.093,082
eil101	101	640,970	921,205	694,573	691,596	671,118
pr136	136	96.770,925	128.542,987	108.965,414	112.895,391	107.395,825
kroA150	150	26.525,032	38.598,426	29.483,554	28.837,293	28.522,831
kroB150	150	26.127,358	37.716,682	29.657,024	30.605,833	27.704,268
kroA200	200	29.369,408	41.342,049	32.742,784	33.259,726	32.155,679
kroB200	200	29.440,413	40.978,338	32.531,488	32.091,640	31.991,792
ts225	225	126.645,934	159.289,453	162.335,238	172.000,048	153.392,967
att532	532	86.742,423	123.413,193	103.529,745	97.964,608	93.348,630
rat575	575	6.798,256	9.925,758	7.929,055	7.755,913	7.776,060
pr1002	1002	259.066,664	355.534,081	319.122,315	314.999,957	299.875,022
vm1084	1084	239.281,871	352.386,306	582.365,072	579.982,365	547.823,128
rl1304	1304	252.940,699	389.858,417	337.303,319	322.065,602	292.002,878

Observa-se um desempenho muito ruim dos três procedimentos que fazem uso inicial de uma envoltória convexa (**tspchi**, **tspgai** e **tspdri**) para a instância **vm1084**. Se esta instância for retirada dos testes, os erros médios destes três métodos passam para 14%, 13% e 9%, respectivamente, ao invés dos

20%, 19% e 15%, respectivamente, caso a instância seja considerada. Possivelmente estes desempenhos ruins, para técnicas promissoras, tenham a ver com as características topológicas da instância **vm1084**.

Para ilustrar alguns dos testes computacionais, são apresentadas as seguintes figuras:

Na figura 1.a são apresentados os nós e a solução obtida pelo algoritmo **tspdri**, aplicado à instância **att48**. Na figura 1.b são apresentados os nós e o roteiro ótimo da mesma instância. Na figura 2, são apresentados os nós e a solução obtida pelo algoritmo **tspchi**, aplicado à instância **ts225**. Na figura 3, exibe-se os nós e o roteiro ótimo da mesma instância **ts225**.

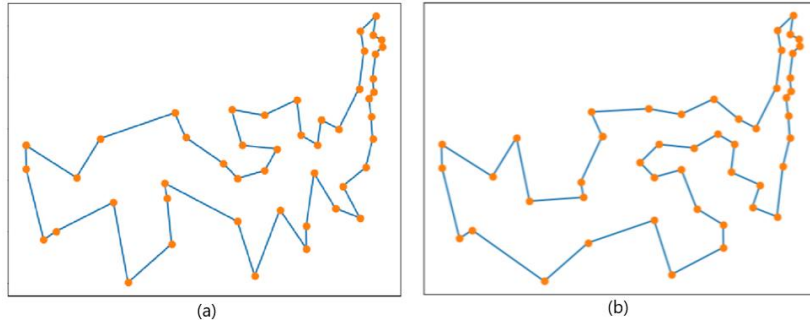


Figura 1: (a) os nós e a solução obtida pelo algoritmo **tspdri**, aplicado à instância **att48**. (b) os nós e o roteiro ótimo da instância **att48**.

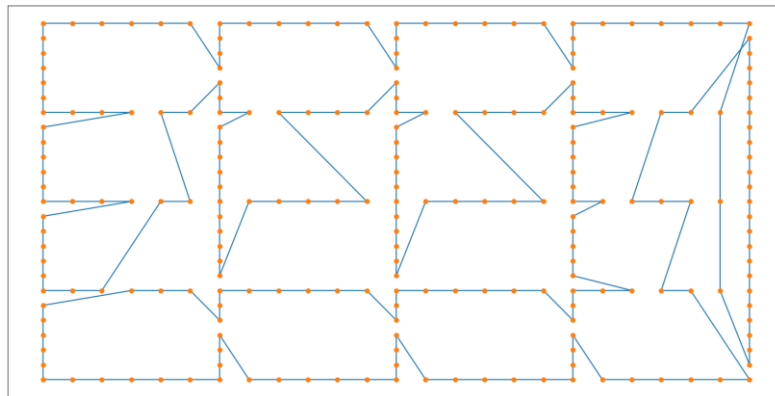


Figura 2: Os nós e a solução obtida pelo algoritmo **tspchi**, aplicado à instância **ts225**.

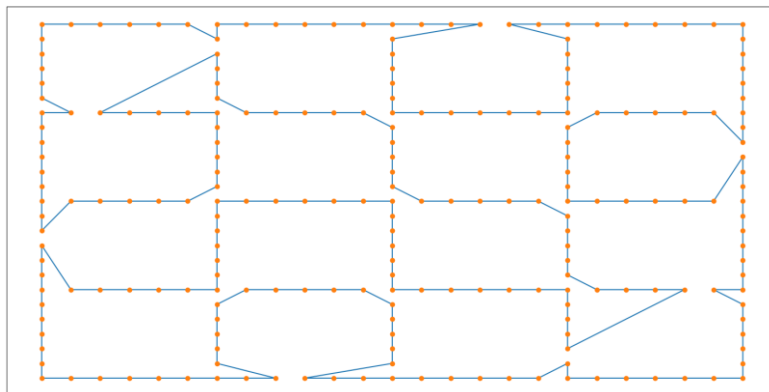


Figura 3: Os nós e a solução ótima da instância **ts225**.

Conclusões

Foram realizados testes computacionais para 9 técnicas clássicas de construção de rotas apresentadas na literatura (BODIN et al., 1983). Para testar os desempenhos destas técnicas, utilizou-se 20 instâncias do Problema do Caixeiro Viajante extraídas do TSP-Library.

Em todos os casos que puderam ser testados, estas técnicas atenderam às expectativas para o desempenho nos piores casos (*the worst case behavior*), previstos em (BODIN et al., 1983).

Sabe-se da dificuldade para se comparar o desempenho de algoritmos para resolver problemas muito sensíveis aos dados de entrada. Por exemplo, mesmo uma técnica bastante simples para resolver o referido problema, como o método **tspviz**, que exigiu baixos tempos computacionais para as instâncias utilizadas, poderia encontrar a solução ótima para grandes instâncias particulares, como uma instância com muitos nós posicionados em uma circunferência.

Este trabalho também resgata a importância da publicação histórica (BODIN et al., 1983), que até hoje tem sido citada por pesquisadores de Pesquisa Operacional, principalmente àqueles que trabalham com roteirização de veículos e *scheduling* de tripulações.

Referências

- BENEVIDES, P.F. **Aplicação de heurísticas e metaheurísticas para o Problema do Caixeiro Viajante em um problema real de roteirização de veículos**. Dissertação de Mestrado (Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática) – Universidade Federal do Paraná: UFPR, 2011.
- BIANCHI, L., KNOWLES, J. e BOWLER, J. **Local search for the probabilistic traveling salesman problem: Correction to the 2-p-opt and 1-shift algorithms**. *European Journal of Operational Research*, v. 162, n. 1, 2005. p. 206-219.
- BODIN, L. et al. **Routing and scheduling of vehicles and crews. The state of the art**. *Computers and Operations Research*, v. 10, n. 2, 1983. p. 63–211.
- BUDINICH, M. **A Self-Organizing Neural Network for the Traveling Salesman problem that is competitive with Simulated Annealing**. *Neural Computing*, v. 8, n. 2, 1996. p. 416-424.
- COOK, W. **Applications of the TSP. The Traveling Salesman Problem**, 2014. Disponível em: < <http://www.math.uwaterloo.ca/tsp/index.html> >.
- DORIGO, M. e GARAMBARDELLA, L.M. **Ant Colonies for the Traveling Salesman Problem**. *Biosystems*, v. 43, n. 2, 1997. p. 73-81.
- KNOX, J. **Tabu Search Performance on the Symmetric Traveling Salesman Problem**. *Computers Ops. Res.*, v. 8, 1994. p. 867–876.
- JIN, H.D., LEUNG, K.S., WONG, M.L. e XU, Z.B. **An Efficient Self-Organizing Map Designed by Genetic Algorithms for the Traveling Salesman Problem**. *IEEE Transactions On Systems, Man, And Cybernetics - Part B: Cybernetics*, v. 33, n. 6, 2003.
- REINELT, G. **TSPLIB**, 1990. Disponível em: < <http://comopt.ifl.uni-heidelberg.de/software/TSPLIB95/> >.
- REINELT, G. **TSPLIB. A traveling salesman problem library**. *ORSA Journal on Computing*, v. 3, n. 4, 1991. p. 376-384.
- SIQUEIRA, P.H. **Uma Nova Abordagem na Resolução do Problema do Caixeiro Viajante**. Curitiba. Tese de Doutorado (Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática) – Universidade Federal do Paraná: UFPR, 2005.
- SOUZA, L. V. de. **Técnicas de Roteirização de Veículos Aplicadas ao Transporte Escolar**. Dissertação de Mestrado (Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática) – Universidade Federal do Paraná: UFPR, 1997.
- STUTZLE, T., GRUN, A., LINKE, S. e RUTTGGER, M. **A comparison of nature inspired heuristics on the traveling salesman problem**. In: *International Conference on Parallel Problem Solving from Nature*, 6., 2000, Paris, France: *Proceedings of International Conference on Parallel Problem Solving from Nature - PPSN-VI*, Springer Verlag: *Lecture Notes in Computer Science*, v. 1917, n. 1, 2000. p. 661–670.

Aplicação de uma Formulação Matemática para Montagem de Planos de Estudos dos Alunos de um Curso de Engenharia

Marcelo Lira Brandão
Universidade Federal Fluminense
marcelolira@id.uff.br

Tiago Araújo Neves
Universidade Federal Fluminense
tneves@id.uff.br

Kelly Alonso Costa
Universidade Federal Fluminense
kellyalonso@id.uff.br

Thalya Salino Braga da Gama Furtado
Universidade Federal Fluminense
thalyafurtado@id.uff.br

Marco Antonio Moraes Pereira Júnior
Universidade Federal Fluminense
marcomoraes@id.uff.br

Resumo: A cada semestre, estudantes de graduação enfrentam a seleção de disciplinas, um processo que implica na análise de múltiplas turmas de disciplinas diversas, com vários conflitos de horário. Muitos alunos decidem com base em interesses pessoais e atividades não acadêmicas, aumentando assim, o risco de atrasar a formatura e de abandono devido à falta de um plano a longo prazo. Este trabalho apresenta os resultados da aplicação de um modelo matemático para auxiliar na criação de cronogramas de estudo, visando encurtar o tempo de graduação, usando dados reais. Os resultados indicam que o problema pode ser resolvido usando programação linear inteira, com tempos computacionais aceitáveis.

Palavras-chave: Otimização Combinatória. Plano de Estudos Personalizado. Formulação Matemática.

Introdução

A evasão escolar, caracterizada pela interrupção prematura dos estudos, é um desafio para nos sistemas educacionais de muitos países. Este processo resulta em indivíduos sem qualificação adequada, que tendem a enfrentar muitas dificuldades de entrada e permanência no mercado de trabalho, o que acaba se tornando um problema crítico para o próprio indivíduo e para a sociedade como um todo (OCDE, 2016).

Estudos indicam que o fenômeno da evasão possui diferentes causas (OCDE, 2016), e que o contexto socioeconômico tem forte influência na conclusão dos cursos, talvez até mais do que etnia e gênero. Estudantes vindos de contextos socioeconomicamente desfavorecidos possuem maior chance de abandonar os estudos devido às restrições financeiras ou problemas familiares.

Outra possível causa para esse fenômeno é o longo tempo necessário para concluir os programas de graduação, especialmente no ensino superior. Estudos mostram que menos de 40% dos alunos terminam o bacharelado dentro da duração teórica prevista (OCDE, 2019). O número se aproxima de 70% quando são considerados três anos adicionais, ou seja, um número considerável dos alunos leva cerca de 7 ou 8 anos para se formar. Outra preocupação são as taxas de conclusão, que estão diminuindo de acordo com os dados apresentados em 2016 e 2019 (OCDE, 2019), mesmo considerando o tempo adicional.

As causas para o longo tempo de conclusão podem ser várias, desde a formação inadequada no ensino médio até a necessidade de trabalhar e estudar simultaneamente. Contudo, uma das possíveis causas pode ser o mal planejamento da trajetória acadêmica. Algumas vezes, na tarefa de fazer seu plano de estudos, os alunos consideram apenas suas próprias preferências e atividades extra-classe, ignorando que algumas disciplinas são requisitos de muitas outras. Assim, quando tomam decisões sem um plano de longo prazo, os alunos estão atrasando sua progressão no curso e possivelmente adiando sua formatura. Em casos críticos, os alunos (e às vezes também suas famílias) esgotam seus recursos financeiros e não conseguem continuar os estudos.

Considerando o cenário descrito, este trabalho apresenta um modelo matemático para resolver o Problema de Construção de Horário Personalizado (PCHP). O PCHP consiste em encontrar o melhor conjunto de turmas para um aluno, considerando as disciplinas que ele já concluiu e as disciplinas e turmas disponíveis no semestre atual, priorizando as disciplinas importantes para a sua progressão no programa.

Uma ferramenta computacional, que usa o modelo matemático, foi construída e testada em casos reais do problema usando dados de um curso de Engenharia da Universidade Federal Fluminense (UFF). O objetivo da ferramenta é ajudar os alunos a fazer um plano de estudos otimizado, que priorize disciplinas críticas nas cadeias do curso e com isso, tentar reduzir o tempo médio que os alunos levam para concluir seus estudos. Os resultados obtidos pela ferramenta mostram que é possível resolver o PCHP de maneira exata em um tempo computacional viável. Como os longos tempos necessários para concluir a graduação estão relacionados ao fenômeno da evasão, uma ferramenta que tente melhorar os índices de alunos que se formam dentro da duração teórica pode ajudar na redução das taxas de evasão. Assim, este trabalho aborda, indiretamente, uma das possíveis causas da evasão escolar, que tem um impacto tão negativo na sociedade em termos econômicos e sociais.

Descrição do Problema e Literatura Relacionada

No Brasil, um estudante precisa concluir um conjunto de diferentes disciplinas para obter uma graduação e cada curso possui um conjunto diferente de disciplinas obrigatórias e optativas. Além disso, cada curso tem um conjunto de regras específicas (número máximo de disciplinas por semestre, número máximo de semestres para concluir o curso, etc). Estas regras podem variar dependendo da universidade e também entre os cursos de uma mesma universidade. Contudo, muitos cursos apresentam grande semelhança em suas grades curriculares (conjunto de disciplinas obrigatórias e opcionais). Cursos com muitas afinidades, como por exemplo, engenharias e cursos na área de Tecnologia da Informação possuem muitas disciplinas em comum. Cada disciplina possui um conjunto de requisitos (outras disciplinas) que devem ser completadas antes que os estudantes possam cursá-la. Um requisito de uma disciplina pode ter um conjunto próprio requisitos criando assim, uma cadeia de disciplinas interligadas. Por exemplo, no curso analisado neste trabalho, a disciplina de Física 3 tem como requisito a disciplina de Física 2, que tem como requisito Física 1; A disciplina de Cálculo 1 é requisito da disciplina de Cálculo 2 e de Física 1. Observe que um curso pode ter muitas cadeias de disciplinas e que, em alguns casos, várias cadeias podem ter interseções disciplinas que estão presentes em mais de uma cadeia).

Cada disciplina pode ter uma ou mais turmas. Cada turma possui um conjunto definido de professores e um conjunto de encontros semanais. Para otimizar os recursos, alunos de diferentes cursos podem compartilhar a mesma turma, e turmas diferentes de uma mesma disciplina podem apresentar diferenças na programação de seus encontros. Ressalta-se que, na universidade onde este estudo está sendo realizado (UFF), o aluno é responsável por fazer seu plano de estudos e, portanto, deve escolher em quais turmas das disciplinas elegíveis se inscrever. Uma disciplina é considerada elegível se o aluno já tiver concluído seus requisitos. Nesse sentido, os alunos competem por aquelas turmas que melhor se adequam aos seus interesses.

Ao fazer o plano de estudos, os estudantes muitas vezes consideram apenas suas preferências e atividades não acadêmicas, ignorando que algumas disciplinas são requisitos de muitas outras. Assim,

ao montar o plano de estudos sem uma visão de longo prazo, os alunos possivelmente estão atrasando o percurso das cadeias e talvez adiando a conclusão do curso. Em casos críticos, os alunos (e suas famílias) esgotam seus recursos financeiros levando-os a abandonar os estudos. Além disso, quanto maior o atraso, maior a chance de atingir o limite de tempo dos cursos, que é 150% do tempo teórico na UFF. Assim, em um curso previsto para quatro anos, o estudante tem até seis anos para sua conclusão. Como o fenômeno da evasão também está relacionado aos longos tempos de graduação, melhorar as taxas de alunos que se formam dentro do prazo teórico pode acarretar uma redução nas taxas de evasão. Nesse sentido, auxiliar os alunos a construir um plano de estudos que considere as cadeias de disciplinas dos cursos pode reduzir ou evitar atrasos e, assim, aumentar as chances dos alunos se formarem dentro do prazo teórico.

Este trabalho aborda o Problema de Construção de Horário Personalizado (PCHP), que consiste em encontrar o melhor conjunto de turmas para um aluno cursar em um semestre, considerando as disciplinas já cursadas e as turmas disponíveis, priorizando disciplinas obrigatórias e críticas nas cadeias do curso, observando as restrições de carga horária máxima e questões de choque de horários.

Os problemas de alocação de horários são NP-Completo em quase todos os casos universitários (SCHAERF, 1999), o que significa que, mesmo com o auxílio de computadores, eles são difíceis de resolver. Existem três grandes problemas na literatura neste contexto (SCHAERF, 1999): O Problema de alocação de horários em Escolas; Problema de Alocação de Cursos; e o Problema de Alocação de Exames. As abordagens mais comuns são meta-heurísticas ou baseadas em meta-heurísticas, que não podem garantir a obtenção de soluções ótimas. Abordagens exatas, como as formulações matemáticas, conseguem encontrar a solução ótima, mas, geralmente, demandam mais recursos computacionais, e são menos frequentes na literatura (TAN, 2021).

O problema da literatura mais próximo ao abordado neste artigo é o Problema de Particionamento de Estudantes (PPE) (SCHINDL, 2019). Neste problema o objetivo é alocar os alunos em turmas de forma que nenhum aluno seja alocado a turmas que se sobreponham no tempo (choque de horários). A preferência dos alunos por turmas e exigências de tempo livre também podem ser consideradas durante a otimização, bem como objetivos institucionais, como capacidades de turmas e distribuições de grupos de alunos. Embora o particionamento totalmente automatizado dos alunos obtenha bons resultados, essa abordagem tem sido criticada desde seus primeiros estudos (BUSAM, 1967). Entre as principais críticas estão o fato de os alunos perderem o poder de decisão, uma vez que eles são alocados pelo sistema, e o fato de que os objetivos institucionais são tratados com maior prioridade. É importante ressaltar que há poucos trabalhos que lidam com casos reais de aplicações do PPE na literatura (MÜLLER; MURRAY 2010). Alguns trabalhos são estudos detalhados e que merecem destaque são Müller e Murray (2010), Carter (2001) e Murray (2007). Estes trabalhos descrevem a estrutura universitária e como a responsabilidade pela resolução do problema é distribuída pelas unidades acadêmicas das universidades de Purdue e Waterloo. Essa é uma tarefa muito complexa e dar suporte aos responsáveis departamentais é importante, uma vez que eles têm um conhecimento muito maior dos detalhes envolvidos (necessidades das turmas, afinidade dos docentes por disciplinas específicas, etc) (MURRAY, 2007). Além disso, manter o senso de propriedade de cada departamento também é um fator importante na aceitação das soluções produzidas pelo processo de agendamento (automatizado ou manual) (MURRAY, 2007).

A grande maioria dos trabalhos na literatura resolvem os problemas pela perspectiva institucional. Ou seja, problemas de horários e/ou particionamento são geralmente resolvidos em função das necessidades institucionais. Poucos trabalhos abordam os problemas na perspectiva dos alunos (BRANDÃO, 2022).

O PCHP, foco deste trabalho, é proposto em Oliveira Neto (2017), onde autores apresentam um sistema que constrói planos de estudos personalizados visando reduzir os tempos de graduação utilizando uma meta-heurística baseada em Algoritmos Genéticos (TALBI, 2009). Em Brandão (2022), os autores propuseram um modelo matemático para o PCHP, e resultados preliminares indicavam que o problema poderia ser resolvido de maneira exata para casos reais. Este artigo é uma continuação do trabalho exposto em Brandão (2022) e apresenta os resultados da aplicação da modelagem matemática proposta para vários casos reais de alunos do curso de Engenharia de Produção da UFF. É importante enfatizar que o objetivo da ferramenta proposta é fazer sugestões de turmas, não atribuições. De acordo com a

Algoritmo 1 prioridade (Vértice v)

```

1: se ( $v.prioridade \leq 0$ ) então
2:    $aux = \alpha_v$ 
3:   para todo Vértice  $w \in N(v)$  faça
4:      $aux = aux + prioridade(w)$ 
5:   fim para
6:    $v.prioridade = aux$ 
7: fim se
8: retorne  $v.prioridade$ 
    
```

(a)

Algoritmo 2 prioridadeGrafo (Grafo G)

```

1: para todo Vértice  $v \in G.E$  faça
2:    $v.prioridade = 0$ 
3: fim para
4: para todo Vértice  $v \in G.E$  faça
5:   prioridade( $v$ )
6: fim para
    
```

(b)

Figura 1: Algoritmos para o cálculo da prioridade utilizados

política da universidade, a decisão final é do(a) estudante.

Modelagem Matemática

Na modelagem proposta a estrutura do curso foi modelada como um grafo $G(V, E)$, onde V é o conjunto de vértices que representam as disciplinas, e E é o conjunto de arcos que representam os requisitos. Assim, um arco partindo de a e chegando em b indica que a é requisito de b . Com isso, pode-se dizer que uma disciplina y é dependente de uma disciplina x se houver um caminho de x para y no grafo. Também pode-se definir a vizinhança de um vértice x como $N(x) = \{y : y \in V, \exists(x, y) \in E\}$. Cada vértice possui todas as informações sobre a disciplina que representa, como o número de turmas, seus horários, etc.

A Figura 1a mostra o algoritmo para o cálculo da prioridade para cada disciplina, que representa relevância desta nas cadeias do curso. Se a prioridade para a disciplina v já estiver calculada (maior que zero) o algoritmo encerra. Caso contrário, o valor é calculado recursivamente como a soma das prioridades de todos os vértices dependentes de v mais α_v (valor arbitrário que indica se a disciplina é obrigatória ou não). Neste trabalho, α_v assume 10 se a disciplina é obrigatória e 0 caso contrário, mas outros valores também podem ser usados. A Figura 1b apresenta o algoritmo para calcular a prioridade para todas as disciplinas de um curso. Primeiro, a prioridade de todas as disciplinas é definida como zero. Posteriormente, a função de prioridade é chamada para cada disciplina.

A formulação matemática também utiliza as seguintes notações:

Seja $S \subset V$ o conjunto de disciplinas já concluídos pelo(a) aluno(a). Seja R o conjunto de disciplinas elegíveis (disciplinas que possuem o conjunto de requisitos satisfeito) definido por $R = \{i : i \in V - S, \nexists(j, i) \in E, j \in V - S\}$.

Seja também q a carga horária semanal máxima (em horas), determinada pelo curso, que um estudante pode cursar. Seja C_i a carga horária semanal da disciplina i . Seja L_i o conjunto de turmas da disciplina i . Seja T o horizonte de tempo. Seja h_{ikt} uma constante que assume 1 se a turma k da disciplina i tem um encontro agendado para o intervalo t , e 0 caso contrário. Seja também p_i a prioridade da disciplina i , calculada usando os algoritmos da Figura 1. Por fim, seja x_{ik} a variável binária que assume 1 se a turma k da disciplina i é selecionada para compor a solução, e 0 caso contrário. Uma possível formulação matemática para o problema é descrita por (1)-(5).

A função objetivo (1) maximiza a prioridade total e a carga horária das disciplinas sugeridas. As restrições (2) impõem que no máximo uma turma de cada disciplina pode ser utilizada. As restrições (3) asseguram que a carga horária semanal seja respeitada. As restrições (4) evitam choque de horários nas turmas selecionadas e (5) são as restrições de integralidade e não-negatividade. A formulação prioriza disciplinas obrigatórias e disciplinas que são requisitos de muitas outras já que estes são fatores considerados no cálculo da prioridade de cada disciplina. A função objetivo também considera a carga horária das disciplinas pelo fato de que a carga horária cursada é um fator relevante para questões internas na universidade (como por exemplo, se dois alunos estão disputando uma vaga em uma turma, o estudante com mais horas cursadas tem preferência). Além disso, a formulação evita que mais de uma turmas de

uma disciplina sejam selecionadas, trata os conflitos de tempo, e assegura a carga horária máxima.

$$Max \quad \sum_{i \in R} \sum_{k \in L_i} p_i c_i x_{ik} \quad (1)$$

S.t.

$$\sum_{k \in L_i} x_{ik} \leq 1, \forall i \in R, \quad (2)$$

$$\sum_{i \in R} \sum_{k \in L_i} c_i x_{ik} \leq q, \quad (3)$$

$$\sum_{i \in R} \sum_{k \in L_i} h_{ikt} x_{ik} \leq 1, \forall t \in T \quad (4)$$

$$x_{ik} \in \{0, 1\} \quad \forall i \in R, \forall k \in L_i \quad (5)$$

Esta formulação foi usada em vários casos reais de estudantes de um curso da UFF e os resultados mostram que é possível resolver o PCHP usando a abordagem proposta. Estes resultados e uma breve discussão sobre eles estão expostos na próxima seção.

Resultados

Para os testes computacionais foram usados 16 casos reais de estudantes do curso presencial de Engenharia de Produção da Escola de Engenharia Industrial e Metalúrgica de Volta Redonda. Cada estudante corresponde a uma instância do problema e todos cederam seus dados voluntariamente. O curso conta com 54 disciplinas obrigatórias (cada uma com uma série de requisitos), divididas entre 10 períodos, além de disciplinas optativas, projetos finais e estágio. Ao todo, é preciso cumprir aproximadamente 3600 horas de atividades para se formar. Aqui cabe ressaltar que, apesar de usar um curso específico como estudo de caso, a formulação matemática criada é geral, e pode ser aplicada a qualquer curso que possua uma grade curricular estruturada. A escolha por este curso específico como estudo de caso se deu por conveniência, visto que os autores e voluntários estão ligados a este curso, o que facilita a obtenção de dados. Para a resolução das instâncias do problema, foi utilizado o software IBM Optimization Studio (IBM-OS) versão 12.10, que usa o CPLEX como motor de otimização. O computador utilizado tem processador i7 de terceira geração, 8 GB de memória principal e Sistema Operacional Linux.

Foram construídas duas versões da formulação proposta. Uma usando a interface do IBM-OS e outra usando diretamente as bibliotecas do CPLEX em Java, chamada de PHU neste trabalho. Um dos objetivos do estudo era verificar se existe algum *overhead* de tempo para o uso da interface do IBM-OS em relação ao uso direto das bibliotecas em Java. É importante observar que medir os tempos de execução do IBM-OS com exatidão não é fácil, mesmo com recursos disponibilizados pela ferramenta. Para obter uma comparação justa, foi estabelecido que seria medido o tempo de relógio da execução de duas tarefas executadas sequencialmente: 1) montagem do modelo matemático; e 2) resolução. Para realizar a medição, foi feito um programa que faz uma chamada do executável *oplrunJava*, que é o mecanismo de uso do IBM-OS em interface de comando, passando como parâmetros os arquivos de configuração (somente opções padrão foram usadas). O *oplrunJava* foi executado cinco vezes para cada instância e as médias calculadas. O PHU realiza as mesmas tarefas listadas anteriormente e também foi executado 5 vezes. A medição de tempo, mostra que os dois programas, estão estatisticamente empatados, e as médias diferem em milésimos de segundos, sendo a média de todas as instâncias de 91 ms para o IBM-OS e 95 ms para o PHU. Como são medições de tempo de relógio, as medidas podem ter interferências do sistema operacional, mas o fato dos dois programas executarem as mesmas tarefas, e terem tempos próximos indicam que a interferência não foi significativa. Os detalhes podem ser vistos na Tabela 1, que expõe os tempos computacionais, a quantidade de turmas que cada estudante pode escolher e a quantidade de turmas sugeridas pelas abordagens. Os resultados mostram que o uso do IBM-OS não

tem *overhead* significativo quando comparado ao uso direto das bibliotecas em JAVA. Este resultado é importante devido ao fato de que construir um modelo de otimização usando o IBM-OS é muito mais simples (aprox. 40 linhas) do que usando as bibliotecas diretamente (aprox. 450 linhas). Isto mostra que, para cenários cuja fonte de dados sejam arquivos e que não tenham necessidade de interações sofisticadas com o resolvidor durante o processo de otimização (caso deste projeto), o uso do IBM-OS é a melhor opção. Pelos resultados também percebe-se que a execução das ferramentas leva décimos de segundo, mostrando que a resolução de casos reais do PCHP via método exato é possível e viável. Neste ponto é importante mencionar que foi possível provar a otimalidade das soluções para todas as instâncias e que o computador utilizado é uma máquina comum, sem nenhum hardware especial. Também é importante mencionar que os horários sugeridos foram todos viáveis. Contudo, não foi feita uma verificação da aceitação das soluções fornecidas pelos estudantes. Isso se deve ao fato de que o objetivo desta fase do projeto era verificar a viabilidade técnica da modelagem, que foi atingido com sucesso. Uma verificação da aceitação das soluções propostas pelos estudantes será objeto de estudo futuro.

Tabela 1: Resultados Computacionais para Casos Reais

Instância	Tempo IBM-OS (ms)	Tempo PHU (ms)	Total de Turmas	Turmas Sugeridas
1	92,2	95,2	67	8
2	92,4	94,6	64	9
3	94	95,2	67	8
4	95,4	95	46	8
5	91,8	94,4	39	8
6	88	90	18	7
7	96,2	93,4	46	9
8	93,8	94,8	42	8
9	87,4	98,2	26	8
10	89,8	96,6	44	8
11	90,6	94,8	42	9
12	89,6	94,6	42	9
13	92,8	100,6	43	8
14	89,6	99,4	41	9
15	96,6	97,8	57	9
16	83	91	21	8
Médias	91,5	95,4	44,1	8,3

Também pode ser observado na Tabela 1 que a média de turmas disponíveis para escolha é de aproximadamente 44 turmas, enquanto a média de turmas sugeridas pela ferramenta é 8 turmas. Neste sentido, o número médio de possibilidades a serem analisadas por cada estudante pode ser descrito, de modo aproximado, por $44!/(8! \times (44 - 8)!)$, que totaliza mais de 5 mil possibilidades (considerando viáveis e inviáveis). Note que o período de inscrição em disciplinas na universidade é de apenas alguns dias, e que não existe, atualmente, uma ferramenta para auxiliar os alunos neste processo, o que torna esta uma tarefa árdua para ser executada. Mesmo considerando que não é possível se inscrever em múltiplas turmas de uma mesma disciplina, o número de possibilidades a serem analisadas ainda é muito grande, o que reforça a importância de uma ferramenta para auxiliar os estudantes neste processo.

Conclusões e Trabalhos Futuros

Este trabalho apresentou os resultados da aplicação de uma formulação matemática para o PCHP em vários casos reais de estudantes de uma universidade brasileira. Em todos os casos foi possível encontrar a solução ótima para o problema em menos de um segundo, o que mostra que o uso da ferramenta proposta é viável. Os resultados também mostram que o número de possibilidades a ser analisada por um estudante a cada semestre é muito elevado e portanto, o uso de uma ferramenta para este fim pode ajudar na realização desta tarefa. A ferramenta proposta faz recomendações baseadas na estrutura curricular do curso, priorizando disciplinas obrigatórias, com mais carga horária e que são relevantes para o progresso do aluno nas cadeias de disciplinas. Com isso, espera-se que, ao seguir as recomendações feitas pela

ferramenta, os alunos possam concluir seus cursos mais rapidamente, reduzindo assim, os tempos médios para a conclusão de curso. Como o tempo para a conclusão é um dos fatores que está ligado à evasão, espera-se que ao reduzir os tempos médios, as taxas de evasão também diminuam.

Como trabalho futuro, pode-se propor a aplicação da mesma metodologia a outros cursos. Há interesse especial de aplicação da metodologia em cursos onde não há tantas relações de requisitos entre as disciplinas. Espera-se que para estes cursos, a ferramenta consiga encontrar as soluções ótimas, mas com um tempo computacional mais elevado, visto que o número de possibilidades de turmas que podem se cursadas aumenta à medida que a quantidade de requisitos diminui. Também considera-se a inclusão de mais elementos na resolução do problema como por exemplo co-requisitos (disciplinas que precisam ser cursadas ao mesmo tempo), pedidos de tempo livre por parte dos estudantes, requisitos fracos (que disciplinas sugeridas, mas sem a relação de requisito formal) e requisitos de outras naturezas (há exemplos de disciplinas que possuem como requisito uma carga horária mínima já cursada). Outro trabalho futuro é a análise da aceitação das soluções apresentadas pelo corpo discente.

Referências

- BRANDÃO, M. L.; NEVES, T. A.; COSTA, K. A. Modelagem matemática para auxiliar estudantes na montagem do horário visando reduzir o tempo de graduação. In: Anais do LIV Simpósio Brasileiro de Pesquisa Operacional (SBPO), 2022, Rio de Janeiro.
- BUSAM, V. A. An algorithm for class scheduling with section preference. **Communications of the ACM**, v. 10, p. 567–569, 1967.
- CARTER, M. W. A Comprehensive Course Timetabling and Student Scheduling System at the University of Waterloo. In: Anais do Pract and Theory of Autom Timetabling III, 2001, Berlin. p. 64-82.
- MURRAY, K. and MÜLLER, T. and RUDOVÁ, H. Modeling and Solution of a Complex University Course Timetabling Problem. In: Anais do Pract and Theory of Autom Timetabling VI - LNCS, 2007, Berlin. p. 189-209.
- MÜLLER, T. e MURRAY, K. Comprehensive approach to student sectioning. **Annals of Operations Research**, v. 181, p. 249–269, 2010.
- OCDE. **Education at a Glance 2016: OECD Indicators**. Paris: OECD Publishing, 2016.
- OCDE. **Education at a Glance 2019: OECD Indicators**. Paris: OECD Publishing, 2019.
- OLIVEIRA NETO, R. F.; MARIANO, C. C. L.; FARIAS, M. S. R. Horário Universitário Personalizado por Metaheurística: Um Estudo de Caso. In: Anais do XLIX Simpósio Brasileiro de Pesquisa Operacional (SBPO), 2017, Rio de Janeiro.
- SCHAERF, A. A survey of automated timetabling. **Artif Intell Review**, v. 13, p. 87–127, 1999.
- SCHINDL, D. Optimal student sectioning on mandatory courses with various sections numbers. **Ann Oper Res**, v. 275, p. 209–221, 2019.
- TALBI, E.-G. **Metaheuristics: From Design to Implementation**. Hoboken: John Wiley & Sons, 2009.
- TAN, J. S. et al. A survey of the state-of-the-art of optimisation methodologies in school timetabling problems. **Expert Systems With Applications**, v. 165, p. 1–13, 2021.

Comparação entre Modelos Preditivos em Séries Temporais Hierárquicas para o Consumo de Energia Elétrica do Setor Industrial Brasileiro

Marlon Mesquita Lopes Cabreira

Universidade Federal Rural do Rio de Janeiro

marlonmlc23@gmail.com

Felipe Leite Coelho da Silva

Universidade Federal Rural do Rio de Janeiro

felipeleite@ufrj.br

Josiane da Silva Cordeiro Coelho

Universidade Federal Rural do Rio de Janeiro

josicordeiro@gmail.com

Resumo: Análise e previsão do consumo de energia elétrica do setor industrial brasileiro vem sendo realizado ao longo dos últimos anos devido a sua importância para economia brasileira, ao meio ambiente e a eficiência energética. Esse setor é o maior consumidor de energia elétrica da matriz energética no Brasil. Os dados de consumo de energia elétrica do setor industrial podem ser organizados em uma estrutura hierárquica composta por cada região (Norte, Nordeste, Centro-Oeste, Sudeste e Sul), no nível inferior, e o Brasil sendo o nível superior. Este trabalho tem o intuito de avaliar a capacidade preditiva dos modelos de suavização exponencial, Box-Jenkins, redes neurais autorregressivas (RNA) e de perceptron multicamadas (MLP). Estes modelos foram integrados as abordagens *bottom-up* e *top-down*. Os resultados mostram que o modelos de suavização exponencial forneceu as melhores previsões e, portanto, superando os outros modelos de previsão hierárquica.

Palavras-chave: Séries Temporais Hierárquicas. Energia Elétrica. Modelos de Previsão.

Introdução

O setor industrial é o maior consumidor de energia elétrica entre os setores que compõem a matriz energética brasileira. Esse setor tem importância para os indicadores do produto interno bruto (PIB) e para sociedade brasileira. Neste contexto, o planejamento energético deste setor é importante para o desenvolvimento econômico, ao meio ambiente e para eficiência energética. Dessa forma, análise e previsão de séries temporais hierárquicas do consumo de energia elétrica podem contribuir para o planejamento e desenvolvimento do setor industrial brasileiro.

As séries históricas do consumo de energia elétrica do setor industrial podem ser organizados em uma estrutura de hierárquica. Em que no topo (nível 0), se encontra os dados do Brasil, e na base (nível 1), os dados de cada uma das cinco regiões geográficas (Norte, Nordeste, Centro-Oeste, Sudeste e Sul).

Os modelos preditivos utilizados neste trabalho foram o de suavização exponencial, Box-Jenkins, redes neurais autorregressivas e de perceptron multicamadas (*MLP - Multilayer Perceptrons*). Cada uma desses modelos foram acoplados a duas abordagens em séries temporais hierárquicas, conhecidas como *bottom-up* e *top-down*. Na abordagem *bottom-up*, são geradas as previsões independentes para cada série no nível mais baixo da hierarquia e, em seguida, elas são agregadas para produzir a previsão dos níveis superiores da hierarquia. Por sua vez, a abordagem *top-down* gera as previsões da série temporal totalmente agregada e, em algumas situações, desagrega tais previsões para os níveis exigidos em proporções históricas.

O objetivo deste trabalho é comparar os modelos preditivos de suavização exponencial, Box-Jenkins, redes neurais autorregressivas e de redes neurais perceptron multicamadas para obtenção das projeções mensais do consumo de energia elétrica do setor industrial brasileiro acoplados as abordagens *bottom-up* e *top-down*. A capacidade preditiva das abordagens será avaliada pelas métricas do erro médio absoluto percentual (MAPE) e do erro quadrático médio (RMSE).

No setor industrial, Silva et al. (2018) introduziram uma abordagem *bottom-up* para obter as projeções de longo prazo do consumo de eletricidade para a indústria de papel e celulose no Brasil, considerando as medidas de eficiência energética. Silva et al. (2019) propõem um modelo que combina os modelos lineares hierárquicos com abordagem *bottom-up* para o consumo de energia elétrica do setor industrial no Brasil.

Com a evolução recente dos algoritmos de aprendizado de máquina, os mesmos têm sido usados para previsão hierárquica de eletricidade. Um estudo comparativo usando os modelos de redes neurais artificiais para previsão do consumo de eletricidade do setor industrial brasileiro é apresentado em Silva et al. (2022). Maiçara et al. (2015) realizaram a projeção do consumo de energia elétrica para o setor residencial no Brasil. Spiliotis et al. (2021) propõem uma abordagem de previsão hierárquica baseada em aprendizado de máquina. Cartinteirol et al. (2007) apresentou um modelo neural híbrido hierárquico para previsão de carga de longo prazo. Fliedner e Lawrence (1995) desenvolveram um estudo empírico com uma análise comparativa sobre o desempenho de previsão de séries temporais hierárquicas.

Metodologia

Neste trabalho, foram utilizados os dados do consumo mensal de energia elétrica do setor industrial do Brasil e por região geográfica para o período de janeiro de 1979 até dezembro de 2022. Os dados estão disponíveis no sítio eletrônico da Empresa de Pesquisa Energética (EPE). Para avaliar a capacidade preditiva dos modelos de suavização exponencial, Box-Jenkins, Redes Neurais Autorregressivas e de perceptron multicamadas acopladas as abordagens *bottom-up* e *top-down*, os dados foram divididos em dois grupos, o conjunto de treinamento e de teste. O primeiro conjunto de treinamento foi do período de janeiro de 1979 a dezembro de 2018, o segundo de 1979 até 2019 e assim sucessivamente, até considerar o último conjunto que finalizava em dezembro de 2022. Por sua vez, os conjuntos de teste sempre continham, somente, 12 observações a frente do conjunto de treinamento. Para cada período de treinamento e teste foram calculadas as métricas MAPE e RMSE.

Notação de Séries Temporais Hierárquicas

Os dados de consumo de energia elétrica da indústria brasileira podem ser desagregados em cinco grupos (regiões do Brasil), apresentando uma estrutura hierárquica, onde a base da estrutura é formada por cinco séries temporais (Norte, Nordeste, Centro-Oeste, Sudeste e Sul) e no topo da hierarquia se encontra a série totalmente agregada que representa os dados do Brasil.

Seja y_t o consumo de energia elétrica observado no mês t , para $t = 1, \dots, T$, com T representando a quantidade de meses da série. De acordo com Athanasopoulos et al. (2009) e Hyndman et al. (2011), uma série temporal hierárquica pode ser representada através de uma equação matricial na forma $Y_t = Sb_t$, onde S é uma matriz soma de ordem $(n \times m)$, em que n é a quantidade de séries em toda hierarquia e m é a quantidade de séries na base da hierarquia (quantidade de regiões, em particular), b_t é o vetor de observações na base da hierarquia de dimensão m no tempo t . Em particular, a estrutura neste trabalho possui $n = 6$ e $m = 5$. Também, podemos expressar a estrutura hierárquica de uma forma unificada usando a equação matricial $\tilde{Y}_h = SG\hat{Y}_h$, em que $\tilde{Y}_h$ representa um conjunto de previsões combinadas h passos à frente, $\hat{Y}_h$ são as previsões de base (ou seja, é o nível inferior da estrutura hierárquica), que são previsões independentes de todas as séries na hierarquia e G é uma matriz referente ao tipo de abordagem de previsão hierárquica que será utilizada.

Abordagem *Bottom-up*

Esta abordagem consiste em utilizar modelos preditivos em cada série no nível inferior e da hierarquia, em seguida, agrega-las para produzir previsões para toda estrutura hierárquica. Uma vantagem dessa abordagem é que prevendo nos níveis inferiores, nenhuma informação perdida devido a agregação. Pode-se representar a abordagem *bottom-up* utilizando a matriz G construída na forma:

$$G = [0_{m \times (n-m)} | I_m], \quad (1)$$

em que I_m é a matriz identidade de ordem m e $0_{m \times (n-m)}$ é a matriz nula de ordem $m \times (n-m)$.

Abordagem *Top-Down*

Nesta abordagem, primeiramente obtém-se a previsão para a série temporal do nível superior da hierarquia e, em seguida, desagrega-se para os níveis inferiores utilizando um conjunto de proporções $p_1, p_2, \dots, p_m$, em que m é a quantidade de séries na base da hierarquia. Estas proporções ditam como a previsão deve ser distribuída nas séries inferiores. Em geral, para um conjunto fixo de proporções, as abordagens *top-down* podem ser representadas como $\hat{Y}_{T+h} = SG\hat{y}_{T+h}$, onde a matriz G é construída na forma:

$$G = [p_m | 0_{m \times (n-1)}] \quad (2)$$

A primeira coluna inclui o conjunto de proporções que distribuem as previsões básicas do nível superior para o nível inferior. O restante das colunas zera as previsões básicas abaixo do nível mais alto de agregação. Como o interesse deste trabalho é em previsões para série totalmente agregada não será realizado a etapa de desagregação das previsões, somente a previsão da série univariada que ocupa o topo da hierarquia.

Modelo de Suavização Exponencial

O modelo de amortecimento exponencial foi introduzido na década de 1950 por Brown (1959), Holt (1957) e Winters (1960). Neste modelo as previsões são geradas a partir de uma média ponderada das observações passadas, sendo que o peso das observações decresce exponencialmente à medida que elas envelhecem. Existem vários modelos de suavização exponencial, eles variam pelas características da série temporal.

Hyndman e Athanasopoulos (2018) propuseram uma classificação que depende de três elementos: erro, tendência e sazonalidade da série temporal. O erro pode ser aditivo (A) ou multiplicativo (M), a tendência pode ser aditiva (A), aditiva com amortecimento (A_d) ou no caso de inexistência, nenhuma (N), por sua vez, a sazonalidade, pode ser aditiva (A), multiplicativa (M) ou nenhuma (N). Dessa forma, cada modelo de suavização exponencial utilizada neste trabalho pode ser classificada como ETS(Erro, Tendência, Sazonalidade). Além disso, foi utilizado o algoritmo *ets* proposto por Hyndman (2008) para ajustar os modelos de suavização exponencial.

Modelo de Box-Jenkins

O modelo de Box-Jenkins utiliza as funções de autocorrelação da série temporal para obter as componentes de tendência, sazonalidade e ruído. Estes modelos foram estudados por Box e Jenkins (1976) e busca integrar modelos autorregressivos (AR) e de médias móveis (MA) a um conjunto de observações.

O modelo utilizado foi o ARIMA sazonal (SARIMA), para incorporar a componente de sazonalidade que está presente nos dados. A estrutura do modelo ARIMA sazonal de ordem $(p, d, q) \times (P, D, Q)_s$ é dada por

$$\phi(B)\Phi(B^s)\Delta^d\Delta_s^D Z_t = \theta(B)\Theta(B^s)a_t \quad (3)$$

em que a_t é o ruído branco; $\phi(B)$ é o operador autorregressivo de ordem p ; $\theta(B)$ é o operador de médias móveis de ordem q ; $\Phi(B^s)$ é o operador autorregressivo sazonal de ordem P ; $\Theta(B^s)$ é o operador de médias móveis sazonal de ordem Q ; Δ^d é o operador da diferença simples e; Δ_s^D é o operador da diferença sazonal; s é o número de observações por ano (período). O modelo de Box e Jenkins foi obtido através do algoritmo proposto por Hyndman e Khandakar (2018).

Redes Neurais

As redes neurais artificiais foram apresentadas pela primeira vez por McCulloch e Pitts (1943). Desde então, observa-se aplicações de redes neurais nas diversas áreas da ciência inclusive em séries temporais. Toda a rede neural artificial tem uma camada de entrada e outra de saída. Também pode haver camadas

intermediárias contendo neurônios ocultos, a qual chama-se de camada oculta. Ambas camadas contêm uma quantidade de neurônios artificiais, em que são conectados entre si formando uma rede.

A camada de entrada é responsável por receber informações oriunda de um meio externo, no caso de uma séries temporal, em geral, são observações passadas, já a camada oculta, tem a tarefa de aplicar pesos as observações passadas e utilizam funções, que são chamadas de funções de ativação, para introduzir a não linearidade nas saídas das camadas ocultas. A camada de saída, em geral, entrega o valor da previsão no tempo futuro subsequente.

Neste trabalho foi utilizado dois tipos de redes neurais: a autorregressiva (RNA) e a perceptron multicamadas (MLP - *Multilayer Perceptrons*). Em todas as análises as redes possuem uma única camada oculta, além disso, no modelo RNA, as séries passaram por uma transformação Box-Cox para tornar a variância da série estacionária e no modelo MLP, foi utilizado diferenças simples e sazonais para transformar a série em estacionária.

Capacidade Preditiva

Para avaliar a capacidade preditiva de cada método preditivo utilizou-se as modelo do erro médio absoluto percentual (MAPE) e a raiz do erro quadrático médio (RMSE), considerando-se que N é quantidade de valores previstos, $\hat{y}_i$ é o valor previsto e y_i é o valor observado ambos para o instante i . Seguem as equações das métricas:

$$MAPE = 100 \times \frac{1}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right|, \quad RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}$$

Resultados

Nesta seção será apresentado as séries temporais do consumo de energia elétrica do setor industrial brasileiro e de cada região. Além disso, será exibido os resultados de capacidade preditiva.

Descrição das Séries Temporais

Na figura 1 é mostrado o consumo de energia elétrica do setor industrial brasileiro comparando as cinco regiões geográficas. O comportamento da série da região Sudeste é semelhante a série do Brasil e além disso, é a região que mais consome energia elétrica, com contrapartida, a região Centro-Oeste é a que menos consome. As regiões Sul e Norte possuem clara tendência de crescimento, aquela, ocupando a segunda região que mais consome energia elétrica por parte de suas industriais, esta, a penúltima. A região Nordeste teve uma tendência de queda depois de 2010. Em todas as séries é visível mudanças de nível e alguma sazonalidade.

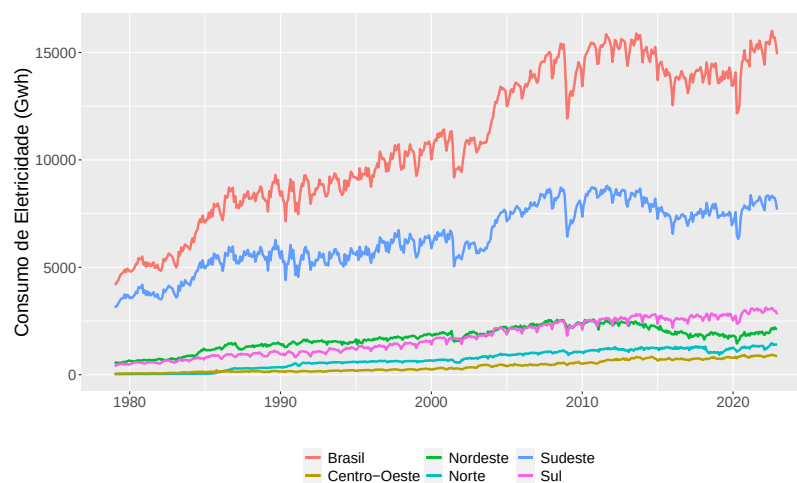


Figura 1: Consumo de energia elétrica no setor industrial do Brasil e região e dos estados brasileiros.

Comparação entre os Modelos Preditivos

Nesta seção serão apresentados os resultados das métricas MAPE e RMSE das previsões dos modelos de suavização exponencial, Box-Jenkins, Redes Neurais Autorregressivas e Redes Neurais MLP acoplados as abordagens *bottom-up* e *top-down*. Em todas as séries, foi aplicado uma transformação Box-Cox para controlar a variância dos dados, já que as séries de consumo de energia elétrica do setor industrial possui sazonalidade multiplicativa e muitas mudanças de nível.

Na tabela 1, pode-se verificar que os valores do MAPE médio são menores que 3,7% entre os modelos preditivos acoplados a abordagem *top-down* adotados neste trabalho. O modelos de suavização exponencial via abordagem *top-down* exibiu o menor valor para ambos os modelos médias, logo, obteve a melhor capacidade preditiva em comparação com as outras abordagens.

Tabela 1: Desempenho dos métodos preditivos nos dados de consumo de eletricidade do setor industrial brasileiro usando a abordagem *top-down*.

Teste	2018		2019		2020	
Suav. Exp.	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,20%	361,59	3,54%	547,72	6,10%	913,20
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,80%	356,99	1,53%	264,25	3,04%	488,75
Teste	2018		2019		2020	
Box & Jenk.	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,47%	271,66	2,46%	399,18	6,38%	945,95
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	3,31%	625,30	2,19%	375,53	3,16%	523,52
Teste	2018		2019		2020	
RNA	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,47%	266,39	1,57%	266,87	5,89%	910,29
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,67%	445,81	4,62%	755,97	3,24%	529,06
Teste	2018		2019		2020	
MLP	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,32%	273,89	3,01%	499,27	6,76%	991,07
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	4,03%	751,82	2,98%	539,15	3,62%	611,04

Na tabela 2, os valores de MAPE médios são menores que 4,6% entre os modelos preditivos acoplados a abordagem *bottom-up*. O modelo de suavização exponencial via abordagem *bottom-up* exibiu o menor valor médio em ambas as métricas, logo, obteve a melhor capacidade preditiva.

Tabela 2: Desempenho dos métodos preditivos nos dados de consumo de eletricidade do setor industrial brasileiro usando a abordagem *bottom-up*.

Teste	2018		2019		2020	
Suav. Exp.	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,44%	268,17	3,01%	408,05	5,65%	892,20
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,45%	449,84	1,57%	262,89	2,83%	470,63
Teste	2018		2019		2020	
Box & Jenk.	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,30%	249,01	2,58%	424,99	6,50%	977,58
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	3,84%	677,18	2,10%	362,86	3,27%	538,33
Teste	2018		2019		2020	
RNA	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	1,78%	308,97	1,91%	360,19	6,76%	1019,85
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	8,64%	1398,48	3,78%	704,15	4,57%	758,33
Teste	2018		2019		2020	
MLP	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	2,39%	410,96	3,56%	584,06	7,10%	1068,06
	2021		2022		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
	4,01%	841,99	3,51%	630,57	4,11%	707,13

Conclusões

Neste trabalho, considerou-se que o consumo de energia elétrica do setor industrial brasileiro pode ser organizado em uma estrutura hierárquica, onde a série histórica do Brasil se encontra no topo, e na base as séries históricas das cinco regiões geográficas. Neste contexto, foram utilizados os modelos preditivos de suavização exponencial, de Box-Jenkins, de redes neurais autorregressivas e de redes neurais MLP integrados as abordagens *bottom-up* e *top-down*, com o intuito de avaliar as previsões do consumo de energia elétrica do setor industrial no Brasil em cinco cenários de conjunto teste diferentes. Para avaliar a capacidade preditiva entre as abordagens foram utilizadas as médias das métricas MAPE e RMSE.

As abordagens consideradas neste trabalho apresentaram um MAPE médio inferior a 4,6%. O modelo de suavização exponencial acoplado a abordagem de *bottom-up* apresentou melhor capacidade entre todos os modelos analisados. Portanto, os resultados mostraram que o modelo de suavização exponencial pode contribuir para obtenção de projeções do consumo de energia elétrica no horizonte de curto prazo.

Para pesquisas futuras é interessante comparar os resultados deste trabalho usando modelos de redes neurais de memória longa.

Agradecimentos

Ao CNPq (407972/2021-8) e a CAPES pelo suporte financeiro.

Referências

- Athanasopoulos, G., Ahmed, R. A., e Hyndman, R. J. (2009). Hierarchical forecast for australian domestic tourism. **International Journal of Forecasting**, 25(1):146-166. ISSN 0169-2070. URL: <https://www.sciencedirect.com/science/article/pii/S0169207008000691>.
- Box, G. E. P., Jenkins, G. M., e Reinsel, G. C. (2015). **Time Series Analysis: Forecasting and Control**. Prentice-Hall.
- Carpinteiro, O. A., Leme, R. C., de Souza, A. C. Z., Pinheiro, C. A., e Moreira, E. M. (2007). Long-term load forecasting via a hierarchical neural model with time integrators. **Electric Power Systems Research**, 77(3):371–378. ISSN 0378-7796.
- Empresa de Pesquisa Energética (2021). **Séries estudos de energia, anuário de energia elétrica**. URL: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/anuario-estatistico-de-energia-eletrica>. Acesso em: 20 Jul. 2023
- Fliedner, E. B. e Lawrence, B. (1995). Forecasting system parent group formation: An empirical application of cluster analysis. **Journal of Operations Management**, 12(2):119–130. ISSN 0272-6963.
- Hyndman, R.J., & Athanasopoulos, G. (2018) **Forecasting: principles and practice**, 2nd edition, OTexts: Melbourne, Australia. URL: [OTexts.com/fpp2](https://otexts.com/fpp2). Acesso em: 20 Jul. 2023
- Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., Shang, H. L. (2011). Optimal Combination forecast for hierarchical time series. **Computational Statistics Data Analysis**, 55(9):2579-2589. ISSN 0167-9473. URL: <https://www.sciencedirect.com/science/article/pii/S0167947311000971>.
- Maçaira, P. M., Souza, R. C., e Oliveira, F. L. C. (2015). Modelling and forecasting the residential electricity consumption in brazil with pegelsexponential smoothing techniques. **Procedia Computer Science**, 55:328–335.
- Mcculloch, W. e Pitts, W. **Bulletin of Mathematical Biophysics**, p. 127–147.
- Sadownik, R., Barbosa, E. P. (1999). Short-term forecasting of industrial electricity consumption in Brazil. **Journal of Forecasting**, 18(3), 215-224.
- Silva, F. L., Cyrino Oliveira, F. L., e Souza, R. C. (2019). A bottom-up bayesian extension for long term electricity consumption forecasting. **Energy**, 167:198–210. ISSN 0360-5442. URL <https://www.sciencedirect.com/science/article/pii/S0360544218321984>.
- Silva, F. L., Souza, R. C., Cyrino Oliveira, F. L., Lourenço, P. M., e Calili, R. F. (2018). A bottom-up methodology for long term eletricity consumption forecast of an industrial sector - application to pulp and paper sector in Brazil. **Energy**, 144:1107-1118. ISSN 0360-5442. URL: <https://www.sciencedirect.com/science/article/pii/S0360544217321217>.
- Silva, F., da Costa, K., Rodrigues, P., Salas, R., e Linkolk, J. (2022). Statistical and artificial neural networks models for electricity consumption forecasting in the brazilian industrial sector. **Energies**, 15:588. <https://doi.org/10.3390/en15020588>
- Spiliotis, E., Abolghasemi, M., Hyndman, R. J., Petropoulos, F., e Assimakopoulos, V. (2021). Hierarchical forecast reconciliation with machine learning. **Applied Soft Computing**, 112:107756. ISSN 1568-4946. URL <https://www.sciencedirect.com/science/article/pii/S1568494621006773>.

MÉTODO PRIMAL SIMPLEX ESPECIALIZADO EM PROBLEMA DA PRODUÇÃO DE RAÇÃO

Marcio Lopes Gevegy 1
Universidade Federal Fluminense 1
marciogevegy@id.uff.br

Ricardo Silveira Sousa 2
Universidade Federal Fluminens 2
rsousa@id.uff.br

Wagner Rambaldi Telles 3
Universidade Federal Fluminens 3
wtelles@id.uff.br

Resumo:

O presente artigo apresenta a implementação e a aplicação do método Primal Simplex Canalizado (PSC) especializado na solução do problema da ração. Os problemas de produção de ração são do Instituto Federal do Espírito Santo (IFES). Para tal propósito, foram formulados 24 problemas (instâncias) que representam diferentes cenários da produção de ração, através de estudos preliminares e de uma visita técnica. No geral o PSC teve um desempenho bem satisfatório. O método fornece soluções robustas e com uma ótima precisão numérica, executando os cálculos em um curto tempo computacional.

Palavras-chave: Programação Linear. Método Primal Simplex Canalizado. Problema da Produção de Ração de Aves e Suínos.

Introdução

Em 1947, George Dantzig formulou o Problema de Programação Linear (PPL) e desenvolveu o método Simplex. O método Simplex é amplamente aplicado em problemas reais, inclusive na produção de ração animal, na literatura é conhecido como Problema da Mistura.

Tem-se que o Brasil ocupa a terceira posição entre os dez principais produtores de ração. O país é um dos principais produtores de proteína animal do mundo. Em 2021 foram produzidas 80,094 milhões de toneladas métricas (FEED, 2022). Cabe destacar que a alimentação representa o maior percentual dos custos de produção animal com cerca de 70% dos custos, especialmente no segmento de aves e suínos. Neste cenário, é crucial utilizar ferramentas para auxiliar a tomada de decisão fornecendo informações para se obter o custo mínimo de ração (NEME, 2000). Como o Método Simplex Especializado em problemas da produção de ração (GEVEGY, 2022).

O problema da ração consiste em encontrar a quantidade ideal de ingredientes a serem misturados de forma a obter um novo produto de custo mínimo. Na indústria alimentícia animal, formular uma ração balanceada é fundamental para atender as exigências nutricionais dos animais (TONIN, 2019). No contexto da produção de rações para aves e suínos, é essencial combinar alimentos de forma adequada, levando em conta a disponibilidade de insumos, exigências nutricionais, custos e entre outros aspectos. Embora seja possível formular uma ração de qualidade por meio de métodos tradicionais, é necessário o uso de ferramentas de gestão mais refinadas, por exemplo a Pesquisa Operacional e a Programação Linear (CASTRO; BORGERT e DE SOUZA, 2015).

No presente trabalho, foram utilizados dados da fábrica de ração do Instituto Federal de Educação Ciência e Tecnologia do Espírito Santo, em Alegre - ES (IFES - Campus Alegre), para abordar o problema da produção de ração e estudado e implementado o Método Primal Simplex para problemas com restrições e variáveis canalizadas.

Método Primal Simplex Canalizado

Considere o modelo geral com restrições e variáveis canalizadas de otimização linear:

$$\begin{aligned} \text{Min} \quad & f(x) = c^T x \\ \text{S.a.} \quad & d \leq Ax \leq e \\ & \alpha \leq x \leq \beta \end{aligned} \quad (1)$$

Em que c representa os custos do problema e o vetor x as variáveis de decisão. Os vetores d e e representam os limites inferiores e superiores das restrições. Já α e β são os limitantes inferiores e superiores das variáveis utilizadas no problema. Para mais detalhes veja Geveg (2022).

Sem perda de generalidade o método Primal Simplex Canalizado foi desenvolvido para o seguinte modelo:

$$\begin{aligned} \text{Min} \quad & f(x) = c^T x \\ \text{S.a.} \quad & Ax = b \\ & L \leq x \leq U \end{aligned} \quad (2)$$

Pressuponha que a matriz A cujo posto é igual a m , o vetor de custos c e x pertence ao conjunto $\mathbb{R}^n$, os vetores L , U pertencem ao conjunto $\mathbb{R}^n$ e b pertence ao conjunto $\mathbb{R}^m$. Do modelo 1 para o 2 ocorre uma mudança de variável, transformando o problema em um equivalente. Os detalhes desta manipulação encontra-se em Geveg (2022).

Dado estas considerações, é feito um particionamento nas colunas de A da seguinte forma:

$$A = [B, N1, N2] \quad (3)$$

A matriz B define a partição associada às variáveis básicas, as matrizes não básicas são representadas por $N1$ e $N2$, do qual a primeira representa a partição associada as variáveis no limitante inferior ($x_j = L_j$), e já a segunda o seu limite superior ($x_j = U_j$). De forma análoga, o vetor x pode ser particionado em:

$$x = [x_B, x_{N1}, x_{N2}] \quad (4)$$

Na equação (4) x_{N1} representa as variáveis no limitante inferior, já x_{N2} representa as variáveis no limitante superior.

O mesmo ocorre com o vetor c que é dado por:

$$c = [c_B, c_{N1}, c_{N2}] \quad (5)$$

Feito estas considerações, o modelo pode ser formulado da seguinte forma:

$$\begin{aligned}
 \text{Min} \quad & f(x) = c_B x_B + c_{N1} x_{N1} + c_{N2} x_{N2} \\
 \text{S.a.} \quad & Bx_B + N1x_{N1} + N2x_{N2} = b \\
 & L_B \leq x_B \leq U_B \\
 & x_{N1} = L_{N1} \\
 & x_{N2} = U_{N2}
 \end{aligned} \tag{6}$$

Para x_B , tem-se:

$$x_B = B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2} \tag{7}$$

Após realizar estas atribuições, se $L_B \leq x_B \leq U_B$, tem-se uma solução básica factível, caso contrário, tem-se uma solução infactível. As variáveis básicas do problema são representadas pelos valores de x_B . Já as variáveis não básicas x_{N1} e x_{N2} sendo a primeira em seu limitante inferior e a segunda no seu limitante superior respectivamente.

Ao substituir x_B no sistema, o modelo fica da seguinte forma:

$$\begin{aligned}
 \text{Min} \quad & f(x) = c_B x_B + c_{N1} x_{N1} + c_{N2} x_{N2} \\
 \text{S.a.} \quad & x_B = B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2} \\
 & L_B \leq x_B \leq U_B \\
 & x_{N1} = L_{N1} \\
 & x_{N2} = U_{N2}
 \end{aligned} \tag{8}$$

Agora considere as seguintes definições conforme visto em Gevegy (2022):

- I_B é o conjunto de índice das variáveis básicas;
- I_{N1} é o conjunto de índices das variáveis não básicas em seu limite inferior com $n1$ elementos;
- I_{N2} é o conjunto de índices das variáveis não básicas em seu limite superior com $n2$ elementos;
- $N1$ é a partição (matriz) não básica associada às variáveis não básicas que estão em seu limite inferior, isto é, $N1 = \{a_{.i} | i \in I_{N1}\}$;
- $N2$ é a partição (matriz) não básica associada às variáveis não básicas que estão em seu limite superior, isto é, $N2 = \{a_{.i} | i \in I_{N2}\}$.

Sendo assim, de acordo com o método Simplex é preciso realizar os cálculos envolvendo as matrizes e os vetores conforme descrito anteriormente da seguinte forma:

1. Cálculo da solução básica:
 $x_B = B^{-1}b - B^{-1}N1x_{N1} - B^{-1}N2x_{N2}$.
2. Cálculo do vetor multiplicador Simplex:
 $\pi = c_B^T B^{-1}$.
3. Cálculo dos custos reduzidos:
 $\hat{c}_j = c_j - \pi^T a_j, \forall j \in I_{N1} \cup I_{N2}$ (critério para escolha do índice que entra na base, q).
4. Direção Simplex:
 $d_B = B^{-1}a_q, q \in I_{N1}$ ou I_{N2} (índice não básico que entra na base).
5. Escolha do δ (tamanho do passo) de tal forma que a nova solução seja factível, ou seja:
 $L_{B_i} \leq x_{B_i} \pm \delta d_{B_i} \leq U_{B_i}, i = 1, \dots, m$.
6. Atualização da inversa da base:
 $(\bar{B}^{-1}) = EB^{-1}$, em que B^{-1} representa a inversa da base, $(\bar{B}^{-1})$ é a nova inversa da matriz base e E a matriz de transformação linear, veja Bazaraa (2010).

Na inicialização do método Simplex as variáveis de decisão são iniciadas no seu limitante inferior ou superior, isto é, $x_i = L_i$ sendo i variando de $1, \dots, n1$ ou $x_j = U_j$ sendo j variando de $1, \dots, n2$. Caso a solução básica inicial não seja factível, recorre-se à Fase I do método Simplex. Esta etapa objetiva buscar uma solução básica factível inicial. A Fase I consiste de mais uma implementação do próprio método. Logo, a fase II inicia-se então com esta solução básica factível e determina (encontra) a solução ótima para o problema original.

Resultados e Discussões

Neste trabalho, verifica-se o desempenho do PSC na otimização dos problemas de produção de ração do Instituto Federal do Espírito Santo (IFES). Para tal propósito, foram formulados 24 problemas (instâncias) que representam diferentes cenários da produção de ração. Por meio de estudos preliminares e de uma visita técnica, foi feito um estudo de caso no IFES em Alegre, ES. A formulação utilizada nas instâncias a seguir é uma representação dos problemas de produção de rações expressos em *kg*, mas a formulação mais adequada é tratada em termos de % conforme GEVEGY (2022). A solução encontrada pelo método PSC foi comparada com a do *software* CPLEX. As instâncias de 1 até 15, são equivalentes as rações de aves. Já as instâncias de 16 a 24 equivalem as rações para suínos.

No geral, os problemas formulados tem como características uma média de 21,88 linhas por 10,75 colunas. Além disso, apresenta um percentual de esparsidade de 70,56%. As instâncias relacionadas as aves possuem em média 22,67 linhas por 9,47 colunas e uma esparsidade de 69,36%. Já as instâncias relacionadas a suínos possuem uma média de 20,67 linhas por 12,89 colunas apresentando um percentual de esparsidade de 72,57%.

Cabe destacar que durante a execução das instâncias, o tempo registrado foi de 0,000000000001 segundos. Sendo assim, o tempo computacional é praticamente nulo.

A Tabela 1 apresenta o número de iterações do método PSC em comparação com o CPLEX para os problemas que envolvem aves.

Tabela 1: Número de iterações - PSC e CPLEX - Aves.

Instância	PSC		CPLEX
	fase I	fase II	
<i>instAves1</i>	43	4	5
<i>instAves2</i>	40	4	4
<i>instAves3</i>	40	4	5
<i>instAves4</i>	42	4	6
<i>instAves5</i>	44	5	4
<i>instAves6</i>	46	5	8
<i>instAves7</i>	34	4	1
<i>instAves8</i>	44	4	7
<i>instAves9</i>	44	3	6
<i>instAves10</i>	43	5	7
<i>instAves11</i>	18	4	9
<i>instAves12</i>	38	4	1
<i>instAves13</i>	46	5	4
<i>instAves14</i>	46	4	3
<i>instAves15</i>	65	5	6
MÉDIA	42,20	4,27	5,07

Fonte: O Autor (2022).

Nota-se que para algumas instâncias de aves, como *instAves15*, exigem um número maior de iteração como visto na fase I do método PSC. Esta instância aparenta ser mais complexa, o que pode necessitar de estratégias adicionais com o intuito de fazer o método convergir mais rápido.

Embora o número de iterações na fase I seja muito maior que na fase II mas na média foi inferior a 2m, em que m é o número de restrições (linhas) técnicas do problema. Este desempenho está de acordo com muitos trabalhos na literatura.

Nota-se que somando o esforço requerido pelo PSC nas fases I e II é muito maior que o CPLEX, mas considerando-se apenas a fase II o desempenho são similares. Vale ressaltar que o software foi desenvolvido há muito tempo e sofre diversas atualizações para melhorar a sua performance, em especial neste caso para se obter uma solução básica inicial factível.

Agora, a Tabela 2 apresenta o desempenho do PSC frente ao software CPLEX para problemas relacionados as rações suínas.

Tabela 2: Número de iterações - PSC e CPLEX - Suínos.

Instância	PSC		CPLEX
	fase I	fase II	
<i>inst1Suínos</i>	34	4	3
<i>inst2Suínos</i>	37	4	3
<i>inst3Suínos</i>	37	4	4
<i>inst4Suínos</i>	38	8	5
<i>inst5Suínos</i>	24	5	3
<i>inst6Suínos</i>	36	5	3
<i>inst7Suínos</i>	40	4	6
<i>inst8Suínos</i>	37	5	4
<i>inst9Suínos</i>	37	7	4
MÉDIA	35,56	5,11	3,89

Fonte: O Autor (2022).

No geral, nota-se um desempenho muito parecido para PSC fase II e CPLEX com relação aos problemas das aves (Tabela 1) e uma pequena redução na fase I para o PSC. Para verificar a precisão numérica do PSC, apresenta-se a solução ótima encontrada para todos os problemas na Tabela 3 a seguir, para PSC e CPLEX.

Durante a execução do método PSC, observou-se que nos casos onde há a fixação de alguns alimentos, reduzindo o espaço de busca, o PSC passa a ter um desempenho melhor. Note que o desempenho do CPLEX e PSC foram muito próximos ao resultados obtidos com os problemas da ração para aves e para suínos.

Portanto, novamente comparando a fase II do PSC com o CPLEX observa-se valores iguais ou superiores ao software comercial. Percebe-se que a iteração na fase II é equivalente à encontrada no CPLEX.

Cabe destacar que o PSC é uma opção válida ao CPLEX para solucionar problemas de produção de ração. Já em casos que demandem uma convergência mais rápida, o CPLEX pode ser recomendado. É necessário explorar estratégias de inicialização e refinamento para otimizar ainda mais o desempenho do PSC.

Tabela 3: Qualidade da solução PSC e CPLEX - Aves e Suínos/IFES.

Instâncias	PSC FO	CPLEX FO	Diferença PSC - CPLEX
<i>instAves1</i>	317,938263624112	317,938263610	0,00000001411
<i>instAves2</i>	312,737313637549	312,737313510	0,00000012755
<i>instAves3</i>	238,200271280072	238,200271370	-0,00000008993
<i>instAves4</i>	254,587177621759	254,587177600	0,00000002176
<i>instAves5</i>	300,102859384186	300,102859460	-0,00000007581
<i>instAves6</i>	226,920918645802	226,920916580	0,00000206580
<i>instAves7</i>	255,155999999823	255,156000000	-0,00000000018
<i>instAves8</i>	254,613130338601	254,613130270	0,00000006860
<i>instAves9</i>	250,549454619190	250,549454600	0,00000001919
<i>instAves10</i>	268,578678347934	268,578678360	-0,00000001207
<i>instAves11</i>	252,897099737040	252,897100990	-0,00000125296
<i>instAves12</i>	216,202000003176	216,202000000	0,00000000318
<i>instAves13</i>	217,630286507885	217,630286490	0,00000001788
<i>instAves14</i>	259,373527128053	259,373527030	0,00000009805
<i>instAves15</i>	266,925810178974	266,925810130	0,00000004897
<i>instSuínos1</i>	254,604283816199	254,60428378	0,000000036199
<i>instSuínos2</i>	244,478869885988	244,47886981	0,000000075988
<i>instSuínos3</i>	222,937456386946	222,93745644	-0,000000053054
<i>instSuínos4</i>	277,972291084490	277,97229105	0,000000034490
<i>instSuínos5</i>	282,104672523379	282,10467251	0,000000013379
<i>instSuínos6</i>	295,548511621892	295,54851155	0,000000071892
<i>instSuínos7</i>	244,931766947135	244,93176693	0,000000017135
<i>instSuínos8</i>	236,974388647334	236,97438853	0,000000117334
<i>instSuínos9</i>	223,612240931667	223,61224102	-0,000000088333

Fonte: O Autor (2022).

No geral nota-se que a qualidade da solução ótima do PSC se igualaram aos resultados do *software* CPLEX até a sétima casa decimal.

Destaca-se que nas instâncias *instAves3*, *instAves5*, *instAves7*, *instAves10*, *instAves11*, *instSuínos3* e *instSuínos9* o método PSC encontrou um valor ótimo com uma precisão melhor que a do CPLEX. Isto pode ter ocorrido devido ao valor de tolerância usado na implementação do PSC, sendo que, para estes problemas foram obtidos valores com uma ótima precisão numérica.

Assim, verifica-se que a implementação do PSC tem uma ótima estabilidade numérica para os problemas tratados.

Conclusão

O PSC foi estudado e implementado para solucionar problemas que retratem a produção de ração de aves e suínos da fábrica de ração do IFES/Alegre. O método foi usado em 24 problemas, sendo 15 deles relacionados à ração de aves e os outros 9 para suínos. A solução foi comparada com o *software* CPLEX e demonstrou uma ótima precisão da solução. O método PSC apresentou um baixo número de iterações (fase I e Fase II). Frente ao CPLEX foi inferior, mas os *software* é uma implementação comercial desenvolvida e aprimorada há muito tempo.

O método PSC auxilia na tomada de decisão na produção da fábrica. Além disso, informa o quanto cada ingrediente contribui tanto na quantidade em quilo de ração quanto no custo.

Assim, o método se mostrou eficaz na solução dos problemas reais da produção de ração de forma precisa e rápida. O presente trabalho evidencia que o PSC é uma ferramenta que pode contribuir muito no setor de produção de ração.

Como trabalhos futuros pretende-se adicionar estratégias para reduzir a quantidade de iteração na fase I. Implementar uma interface gráfica para facilitar a gestão dos ingredientes e utilizar o método para fazer a formulação da ração de aves e suínos considerando a formulação em termos de porcentagem como tratado em diversos trabalhos na literatura.

Referências

- ALL ABOUT FEED. World feed production increases to over 1.2 billion tons. 2022. Disponível em : <https://www.allaboutfeed.net/market/market-trends/world-feed-production-increases-to-over-1-2-billion-tons/>. Acesso em 27/03/2022.
- BAZARAA, M. S.; JARVIS, J.; SHERALI H. D. **Linear programming and network flows**. New Jersey: John Wiley & Sons, 2010.
- CASTRO, L. Y.; BORGERT, A; DE SOUZA, F, R. Definição do mix de produção em uma indústria de lácteos com uso da programação linear: um estudo de caso. **Anais do Congresso Brasileiro de Custos-ABC** . Foz do Iguaçu, PR, 2015.
- GEVEGY, M. L. **Método Simplex Canalizado para Produção de Ração: Um Estudo de Caso**. 2022. 138 f. Dissertação (Mestrado em Modelagem Computacional), – Universidade Federal Fluminense - RJ - Volta Redonda, 2022.
- NEME, R. **Digestibilidade verdadeira e biodisponibilidade da lisina sulfato e da lisina hcl determinadas em aves**. Universidade Federal de Viçosa (2000).
- TONIN, H. N. **Modelo de programação linear aplicado a redução de custo na formulação de ração para nutrição de bovinos**. 2019. 49 f. Dissertação (Mestrado em Engenharia Química), – Universidade de Uberaba - MG - Uniube, Uberaba, 2019.

PREVISÃO DO CONSUMO DE ENERGIA ELÉTRICA ATRAVÉS DE MODELOS DE SÉRIES TEMPORAIS UNIVARIADOS E MULTIVARIADOS

Rafael Santos de Oliveira

Universidade Federal Rural do Rio de Janeiro

fael.rso1@gmail.com

Felipe Leite Coelho da Silva

Universidade Federal Rural do Rio de Janeiro

felipeleite@gmail.com

Josiane da Silva Cordeiro Coelho

Universidade Federal Rural do Rio de Janeiro

josicordeiro@gmail.com

Resumo: O setor industrial é o maior consumidor de energia elétrica no Brasil e notavelmente é um dos mais importantes pelo seu grande impacto econômico e social. O objetivo deste trabalho é comparar os modelos de previsão autoregressivos vetoriais e o regressão dinâmica aplicados aos dados de energia elétrica do setor industrial. Para o modelo de regressão dinâmica será utilizado como variável exógena o PIB industrial. A capacidade preditiva dos modelos será avaliada pelas métricas MAPE e RMSE. Os resultados indicam que o modelo de regressão dinâmica apresentou a melhor capacidade preditiva.

Palavras-chave: Séries Temporais. Energia Elétrica. Modelos de Previsão. Produto Interno Bruto.

Introdução

Nos últimos anos, o consumo de energia elétrica e o produto interno bruto (PIB) do setor industrial brasileiro vem sendo investigados devido a sua importância no desenvolvimento nacional em termos sociais e econômicos. Dessa forma, a análise e previsão desses conjuntos de dados são importantes para o planejamento do sistema elétrico e, também, para o desenvolvimento de setores da indústria.

Neste contexto, a análise de séries temporais é importante para investigar e prever o consumo de energia elétrica (Sadownik; Barbosa, 1999). Esta análise permite identificar padrões e tendências nos dados históricos, o que é fundamental para a previsão de eventos futuros. Também é possível identificar valores atípicos e anomalias nos dados, que podem indicar problemas ou eventos excepcionais que requerem atenção especial. Neste trabalho, os conjuntos de dados tratados consideram períodos trimestrais para poder adequar-se a alguns pressupostos dos modelos de séries temporais. Primeiramente, será realizada uma análise exploratória dos dados referentes a cada um dos seus conjuntos, depois a verificação dos pressupostos dos modelos e por fim, a predição dos mesmos.

A previsão do consumo de energia elétrica no setor industrial brasileiro tem sido realizada por diferentes modelos de séries temporais, tais como os modelos de redes neurais artificiais (Silva et al., 2022), modelos de regressão dinâmica integrados a uma abordagem bottom-up (Silva, et al. 2018; Silva et al., 2019.) e os modelos de amortecimento exponencial e de Box-Jenkins (Mendes et al., 2023).

De acordo com Amaro et al. (2017), a previsão da demanda de energia de curto prazo é difícil de ser realizada porque envolve uma grande influência de vários fatores sociais e econômicos. Mas, sugerem que os modelos univariados podem ser utilizados para prever a demanda de energia elétrica de forma eficiente. Por outro lado, Castro et al. (2016) mostraram que a inclusão de variáveis que se inter-relacionam com a variável de interesse, aumenta a capacidade de previsão, sugerindo a utilização de modelos multivariados.

O objetivo deste trabalho é comparar os modelos autoregressivos vetoriais (VAR) e o de regressão dinâmica para a previsão do consumo de energia elétrica do setor industrial. No modelo de regressão dinâmica, será utilizado o PIB como variável regressora. A capacidade preditiva dos modelos foi avaliada pelas métricas do erro médio absoluto percentual (MAPE) e a raiz do erro quadrático médio (RMSE).

Metodologia

Neste trabalho, foram utilizados os dados do PIB (em reais) e o consumo de energia elétrica (em MWh) do setor industrial brasileiro, considerando o período de janeiro de 2002 até dezembro de 2021. Esses dados estão disponíveis no sítio eletrônico do Instituto de Pesquisa Econômica e Aplicada (IPEA, 2022). Para testar o poder de ajuste e as generalizações das previsões, os dados foram divididos em dois conjuntos: o conjunto de treino, com intuito de exploração dos dados, ajustes e validação dos modelos e o conjunto de teste para avaliar a capacidade preditiva.

Na primeira parte do estudo foi apresentada uma análise exploratória das séries temporais para compreender seu comportamento e suas características de nível, tendência, ciclo, sazonalidade e ruído. Estas informações são importantes para identificar o tipo de modelo a ser ajustado.

O conjunto com as séries temporais do consumo de energia elétrica industrial e PIB foi dividido em duas partes: as primeiras 68, 72 e 76 observações trimestrais de cada série foram utilizadas como conjunto de treinamento, correspondendo aos anos de 2002 até 2018, 2002 até 2019 e 2002 até 2020, respectivamente, para estimar os parâmetros e as últimas 4 observações como o conjunto de teste, ou seja, para obter como previsão 2019, 2020 e 2021.

Dessa forma, foram ajustados os modelos VAR e o de regressão dinâmica. Em seguida, foram obtidas as previsões para cada uma das séries temporais considerando 4 passos à frente. Para selecionar o modelo mais adequado foram utilizadas as métricas de acuracidade do erro médio absoluto percentual (MAPE) e a da raiz do erro quadrático médio (RMSE).

Modelo Autorregressivos Vetoriais - VAR

Entre os modelos multivariados de séries temporais, os modelos VAR são os mais utilizados na prática, devido a sua semelhança com os modelos de regressão e a relativa facilidade de ajustar esses modelos às séries temporais reais (Box et al., 2015). Por exemplo, os parâmetros podem ser estimados usando métodos de mínimos quadrados que geram expressões de forma fechada para as estimativas. Esses modelos são utilizados em econometria para descrever o comportamento dinâmico de séries temporais econômicas (ou financeiras) e para produzir previsões.

O modelo autorregressivo vetorial de ordem p , VAR(p), para n variáveis é dado por

$$X_t = \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + a_t, \quad (1)$$

em que X_t é um vetor de ordem $(n \times 1)$, composto pelo t -ésimo valor das variáveis analisadas, ϕ_0 é um vetor de ordem $(n \times 1)$, composto por constantes, ϕ_k é uma matriz de ordem $(n \times n)$ composta por elementos ψ_{ijk} , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, $k = 1, 2, \dots, p$ e a_t é um vetor de ordem $(n \times 1)$, composto pelos t -ésimos erros aleatórios.

Modelo de Regressão Dinâmica

A regressão dinâmica consiste em estender os modelos ARIMA para permitir que outras informações sejam incluídas no modelo, com o intuito de melhorar a análise e previsão de séries temporais. Segundo Hyndman e Athanasopoulos (2018), o modelo de regressão dinâmica é definido por

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t \quad (2)$$

onde y_t é uma função linear de k variáveis preditoras $(x_{1,t}, \dots, x_{k,t})$, e ε_t geralmente é considerado um termo de erro não correlacionado (ou seja, é ruído branco).

Neste modelo permite-se que os erros de uma regressão contendam autocorrelação. Para enfatizar essa mudança de perspectiva, substitui-se ε_t por η_t na equação. A série de erros η_t assume-se que segue um modelo ARIMA.

$$y_t = \beta_0 + \beta_1 x_{1,t} + \cdots + \beta_k x_{k,t} + \eta_t \quad (3)$$

$$(1 - \phi_1 B - \cdots - \phi_p B^p)(1 - B)^d \eta_t = (1 + \theta_1 B + \cdots + \theta_q B^q) \varepsilon_t \quad (4)$$

em que ε_t é o erro do modelo ARIMA, ou seja, ruído branco e η_t é o erro do modelo de regressão.

Métricas para os modelos de previsão

Para selecionar o modelo de previsão mais adequado foram utilizadas as métricas de raiz do erro quadrático médio (RMSE) e o erro médio absoluto percentual (MAPE). Seguem as definições abaixo, considerando N a quantidade de valores previstos, $\hat{y}_i$ é o valor previsto e y_i é o valor observado ambos para o instante i :

1. A raiz do erro quadrático médio é definido por:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5)$$

2. O erro médio absoluto percentual (MAPE) é definido por:

$$MAPE = 100 \times \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (6)$$

Resultados e Discussões

Nesta seção, primeiramente será apresentada uma análise exploratória para as séries temporais do consumo de energia elétrica e PIB do setor industrial brasileiro. Em seguida, serão apresentados os resultados dos modelos VAR e o de regressão dinâmica.

Análise Exploratória

A Figura 1 mostra as séries temporais do consumo de energia elétrica e o PIB do setor industrial brasileiro. Pode-se observar nesta figura que ambas as séries temporais possuem tendência e sazonalidade.

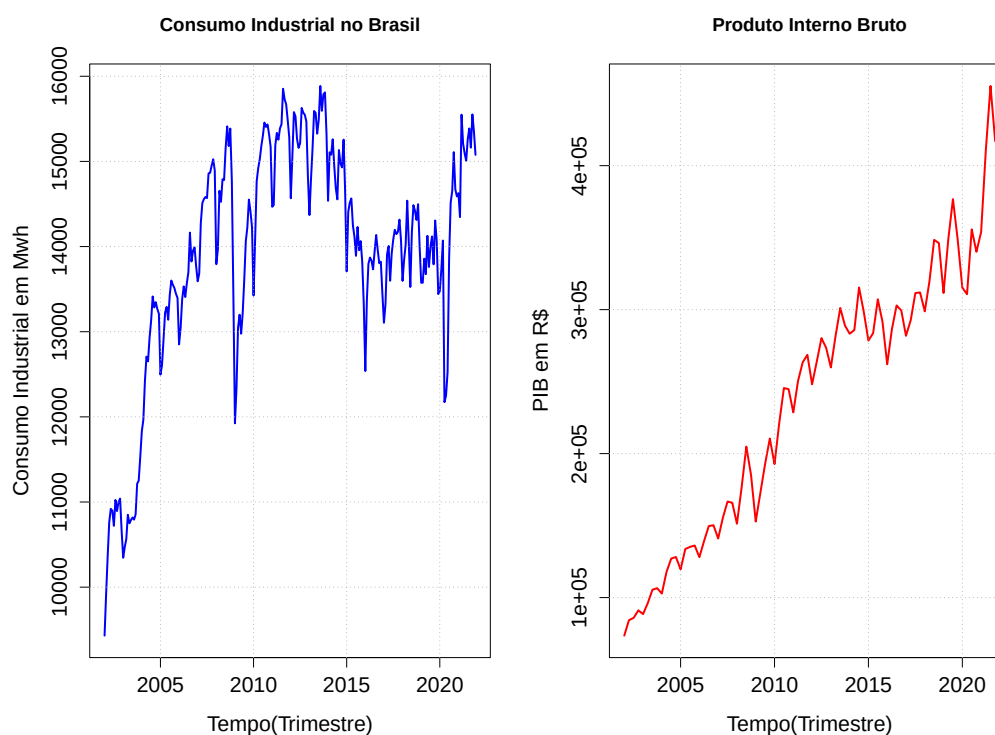


Figura 1: Séries Temporais do Consumo de energia elétrica industrial do Brasil e o PIB industrial brasileiro.

A série temporal do consumo de energia elétrica industrial possui bastante mudança de nível, com vales mais significativos em 2008, 2016 e 2020. Em 2008 teve a crise econômica mundial e em 2020 e 2021 a pandemia da COVID-19 e por estes motivos refletiram diretamente no comportamento da série temporal. No PIB também houve uma queda, porém menos acentuada. Nota-se também após a pandemia um crescimento significativo em ambas as séries temporais.

Uma maneira de confirmar a intuição anterior é utilizando a decomposição STL (Seasonal and Trend Decomposition using Loess). Essa decomposição divide os dados da série temporal nas componentes de tendência, sazonalidade e resíduo. Através disso, confirma-se a existência de tendência e a sazonalidade para o período trimestral.

Na Tabela 1, pode-se observar as medidas estatísticas para os dados de consumo de energia elétrica da indústria brasileira e para o PIB industrial brasileiro para o período de 2011 a 2021.

Tabela 1: Tabela de medidas estatísticas anuais do gasto energético da indústria e do PIB dos últimos dez anos.

	Média	Variância	Desvio padrão	Mínimo	Máximo	Co. Variação
Consumo	43494,45	5303194,35	2302,87	36943	47020	5,29
PIB	308146,31	2131700042,47	46170,34	228636,05	455358,93	14,98

A correlação entre os dados de consumo de energia elétrica da indústria do Brasil e do PIB industrial é 0,63. Com isto, nos dá evidências de que uma série temporal possa interferir e melhorar o modelo no comportamento da outra e podemos aplicar vários modelos multivariados para tentar melhorar a predição da série temporal endógena.

Modelo VAR

Antes de ajustarmos o modelo VAR(4), foi verificado pelos critérios de parcimônia AICc, BIC e AIC qual era o mais baixo. Para a adequabilidade do modelo proposto foi aplicado o teste de Portmanteau para a correlação serial. Aplicando este teste de correlação serial, apresentou valor-p aproximadamente de 0,05 indicando a ausência de correlação serial a 95% de confiança. Pelo teste da raiz unitária, calculando os autovalores associados a matriz de defasagens, todos os valores deram menores que 1. Logo, o modelo VAR(4) estimado é estável e portanto o modelo é adequado.

A Tabela 2, apresenta os valores das métricas para cada período dos conjuntos de teste. O modelo VAR(4) apresentou uma capacidade preditiva tendo um MAPE e RMSE médio de 4,3701 e 2177,3135.

Tabela 2: Tabela com as métricas MAPE e RMSE para o conjunto de teste do consumo de energia elétrica industrial para o modelo VAR(4).

Teste	2019		2020		2021		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
VAR(4)	1,33	690,60	5,79	3076,67	5,99	2764,68	4,37	2177,31

Modelo de Regressão Dinâmica

Na tabela 3 abaixo apresenta as métricas obtidas para o modelo de regressão dinâmica, considerando os diferentes conjuntos de teste. O modelo de regressão utilizou o PIB como variável regressora. Esse modelo apresentou um MAPE e RMSE médio de 4 e 2015,51, respectivamente.

Tabela 3: Tabela com as métricas MAPE e RMSE para o conjunto de teste do consumo de energia elétrica industrial considerando o modelo de Regressão Dinâmica.

Teste	2019		2020		2021		Média	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
DLM	2,68	1166,31	6,14	2881,89	3,18	1998,36	4	2015,52

Comparação entre os modelos VAR e de Regressão Dinâmica

A Tabela 4 apresenta as métricas de comparação para os conjuntos de treinamento e de teste para os modelos considerados neste estudo. O modelo de regressão dinâmica apresentou menores valores de MAPE e RMSE tanto para o conjunto de teste e treinamento (Tabela 4). Portanto, o modelo de regressão dinâmica apresentou a melhor capacidade preditiva.

Tabela 4: Tabela com as métricas MAPE e RMSE para os modelos VAR e de regressão dinâmica.

Conjuntos	Treinamento		Teste	
Métricas	MAPE	RMSE	MAPE	RMSE
DLM	2,12	1173,57	3,18	1998,36
VAR(4)	2,3	1368,8	5,99	2764,68

Observa-se na Figura 2(a) que ambos os modelos capturaram bem as variações ao longo da série temporal nos anos de 2000 a 2021, porém, nota-se que no ano de 2020 os modelos não se ajustaram bem no vale abrupto existente. A série predita pelo modelo de regressão dinâmica está mais próxima aos dados observados para a previsão trimestral de 2021 (Figura 2(b)), corroborando com as métricas apresentadas na tabela 4.

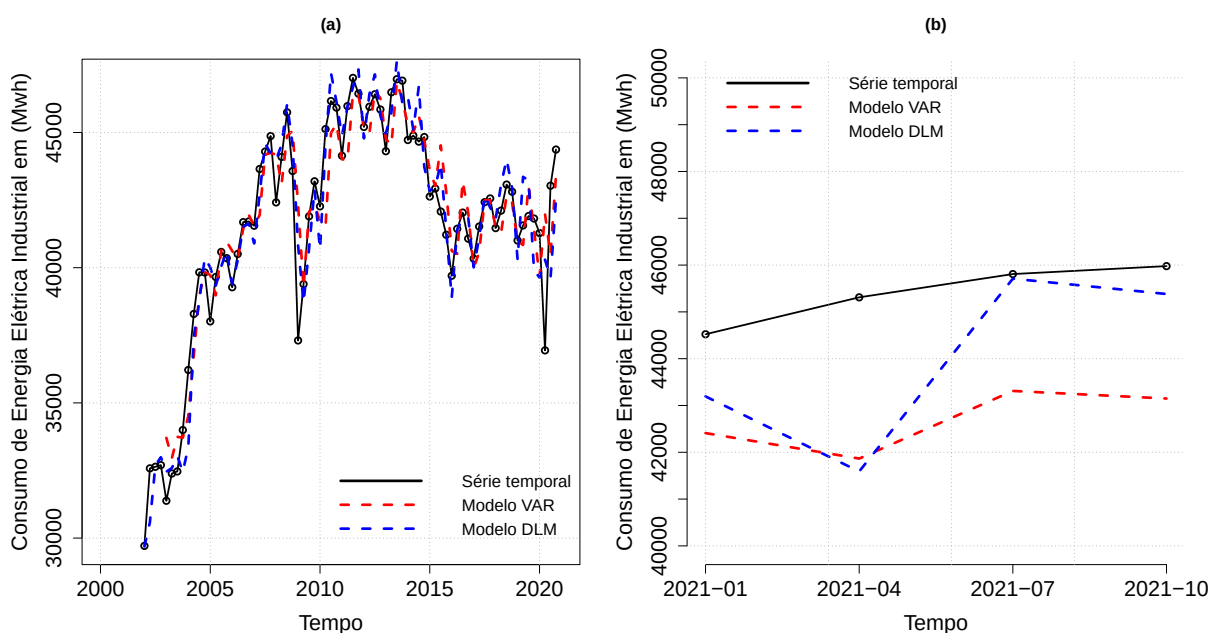


Figura 2: Ajuste (para o conjunto de treinamento) (a) e previsão com intervalo de confiança de 95% (para o conjunto de teste) (b) do consumo de energia elétrica do Brasil utilizando os modelos de regressão dinâmica e VAR(4).

Conclusões

Neste trabalho, aplicaram-se os modelos de regressão dinâmica e o VAR, a fim de prever o consumo de energia elétrica no Brasil. Os resultados mostraram que estes modelos podem contribuir para obtenção de projeções do consumo de eletricidade no horizonte de curto prazo. Os dados trimestrais analisados foram de janeiro de 2002 até dezembro de 2021, obtidos no sítio eletrônico do IPEA. As métricas de precisão MAPE e RMSE apontaram que o modelo de regressão dinâmica foi aquele com a melhor capacidade de predição e de ajuste dentre os modelos analisados. Como uma extensão deste trabalho, pretende-se para pesquisas futuras, investigar e comparar os resultados deste trabalho usando os modelos VAR-VEC e de redes neurais para realizar as previsões das séries temporais, investigar comportamentos das séries e verificar a relação contemporânea entre elas.

Agradecimentos

A CAPES pelo suporte financeiro.

Referências

- Amaro, R. S., Ceretta, P. S., Coronel, D. A., Bender, R., Filho, Oliveira, M. G. (2017). **Comparação da capacidade preditiva de modelos ARIMA para o consumo de energia elétrica no Brasil**, Revista Unemat de Contabilidade, 6(11), 173-192. DOI: <https://doi.org/10.30681/ruc.v6i11.1848>
- Box, G. E. P., Jenkins, G. M., e Reinsel, G. C. (2015). **Time Series Analysis: Forecasting and Control**. Prentice-Hall.
- Box, G. E.; Jenkins, G. M.; Reinsel, G. C. (1994). **Time series analysis: Forecasting and control**. 3a ed. New Jersey: Printice Hall.

- Box, G. E. P., Jenkins, G. M. (1970). **Time series analysis: forecasting and control**.
- Castro N. R., Gilio, I., Silva, A. F., Ozaki, V. G. (2016). **Modelos univariados e multivariados aplicados à previsão de valores de exportação: uma análise comparativa para o complexo soja**, Revista Espacios. 37(3). Recuperado em 10 abril, 2020, de: <https://www.revistaespacios.com/a16v37n03/16370315.html>
- Gujarati, D. N., Porter, D. C. (2011). **Econometria Básica**. 5a ed. São Paulo, SP: AMGH Editora.
- Hamilton, J. D. (1994). **Time Series Analysis**. Princeton University Press.
- Ipea (2022). Ipea data. URL <http://www.ipeadata.gov.br/Default.aspx>.
- Leite, S. Q. (2006). **Projeções para a demanda por energia elétrica no Brasil, 2006-2015**. Dissertação de Mestrado Profissionalizante, Faculdades Ibemec, Rio de Janeiro, RJ, Brasil. Recuperado em junho, 2018, de: http://s3.amazonaws.com/public-cdn.ibmec.br/portallibmeccontent/public/arquivos/df/dis_2006_12_-_sidimar_quezada_leite.pdf
- Mendes, R. F. da S., da Costa, K., da Silva, F. L. C., Coelho, J. da S. C., Vera-Tudela, C. A. R., & Pinto, R. V. (2023). **Forecasting models for the electricity consumption of the cement industry in Brazil**. OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA, 21(7), 6016–6031. <https://doi.org/10.55905/oelv21n7-011>
- Morettin, P. A., Toloi, C. M. C. (2020). **Análise de Séries Temporais: Modelos multivariados e não lineares**. Editora: Edgard Blucher Ltda, Volume 2.
- Morettin, P. A., Toloi, C. M. C. (2018). **Análise de Séries Temporais: Modelos lineares univariados** 3ª edição. Editora: Edgard Blucher Ltda, Volume 1.
- Silva, Felipe & da Costa, Kleyton & Rodrigues, Paulo & Salas, Rodrigo & López-Gonzales, Javier Linkolk. (2022). **Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector**. Energies. 15. 588. 10.3390/en15020588.
- Silva, Felipe & Souza, Reinaldo & Oliveira, Fernando Luiz & Lourenço, Plutarcho & Calili, Rodrigo. (2018). **A bottom-up methodology for long term electricity consumption forecasting of an industrial sector - Application to pulp and paper sector in Brazil**. Energy. 144. 10.1016/j.energy.2017.12.078.
- Silva, Felipe & Oliveira, Fernando Luiz & Souza, Reinaldo. (2018). **A bottom-up bayesian extension for long term electricity consumption forecasting**. Energy. 167. 10.1016/j.energy.2018.10.201.
- Hyndman, R.J., & Athanasopoulos, G. (2021) Forecasting: principles and practice, 3rd edição, OTexts: Melbourne, Australia. OTexts.com/fpp3. Acessado em 10/08.

Proposta metodológica para a revisão sistemática da literatura: estudo de caso associado à Teoria das Filas e Simulação

Kelly Cristine Franco de Azevedo
Universidade Federal Fluminense
kellyfranco@id.uff.br

Cecília Toledo Hernández
Universidade Federal Fluminense
ctoledo@id.uff.br

Eliane da Silva Christo
Universidade Federal Fluminense
elianechristo@id.uff.br

Resumo: O presente artigo oferece uma revisão sistemática sobre os temas de teoria das filas e simulação, utilizando as bases de dados da CAPES: *Scopus* e *Web of Science*. Ao abordar essas duas áreas que estão presentes em diversas disciplinas e mais especificamente na logística, a pesquisa visa examinar a evolução e as tendências das publicações ao longo do tempo e ver a pertinência das ferramentas para avaliar indicadores de eficiência dos processos logísticos. Mediante a análise e metodologia sistemática, o estudo explora a quantidade de artigos por ano, destacando anos de maior e menor atividade. A fusão dos dados da CAPES, *Scopus* e *Web of Science* forneceu *insights* sobre os padrões de pesquisa nessas áreas. Os resultados destacam anos de pico de publicações, mostrando flutuações naturais na produção acadêmica e contribui para o entendimento da dinâmica da pesquisa em teoria das filas e simulação, oferecendo uma visão das tendências e áreas de interesse ao longo do período estudado. Ao restringir a busca a área de logística, mesmo que a quantidade de artigos diminui, ainda mostra como ambas as ferramentas são usadas em pesquisas sobre este tema. Este resultado pode ser significativo para demonstrar a relevância atual deste tópico, o qual pode ser explorado em futuras pesquisas e projetos.

Palavras-chave: Teoria das Filas. Simulação. Logística. Revisão sistemática.

1. Introdução

Num ambiente globalizado a logística surge como uma vantagem competitiva utilizada como estratégia de criação de valor. Nesse sentido a avaliação de desempenho tem o intuito de promover informações sobre a performance das atividades logísticas (BALLOU, 2012). Quando se fala de desempenho logístico, os indicadores abrangem todas as atividades logísticas de transporte, aquisição de materiais e peças, gestão de estoque, gestão de armazenagem, gestão da manutenção, engenharia de infraestrutura, e distribuição física, sendo o tempo e a qualidade aspectos desafiadores da produção nos mercados transnacionais, por este motivo o indicador “Qualidade e competência logística” está entre os definidos em 2018 pelo Banco Mundial (MACHADO; DOS SANTOS, 2021).

Para Tavares (2018) na avaliação dos cenários sobre desempenho logístico é necessário a adoção de ferramentas de apoio à decisão que contribuam para a redução de custos e ganhos operacionais que por sua vez aumentam a capacidade de competição das empresas. Entre estas ferramentas se destacam a simulação computacional e a teoria das filas (DA SILVA et al., 2021).

A teoria das filas é um ramo da Pesquisa Operacional que tem como foco principal utilizar conceitos básicos de processos estocásticos e de matemática aplicada com o intuito de estudar o processo de formação de filas e quais são as características da mesma permitindo o dimensionamento

adequado de instalações, equipamentos e infraestrutura adequada para a quantidade de demanda (DE BRUNS; SONCIM; DE SINAY, 2001).

Outra ferramenta utilizada para estudar e avaliar os processos de fila em sistemas de produção e serviços é a simulação computacional. Da Silva et al. (2021) realizaram um estudo sobre a otimização dos processos logísticos do porto de Itajaí utilizando simulação e indicadores da teoria das filas, mostrando as particularidades de uso de cada uma das ferramentas. Neste contexto aparecem diferentes questões de pesquisa a serem respondidas: Qual dessas duas ferramentas, teoria de filas ou simulação, dá uma melhor resposta na otimização de processos logísticos relacionados com filas de espera em processos produtivos ou de serviços? Quais vantagens ou desvantagens estão associadas a cada ferramenta?

Neste sentido este artigo apresenta uma proposta metodológica para conduzir a revisão bibliográfica tendo como temática a utilização da teoria das filas e da simulação computacional de eventos discretos para avaliar indicadores de eficiência logística. Como resultado espera-se que a proposta metodológica auxilie no levantamento do estado da arte e no entendimento dos *gaps* ou lacunas que permitam desenvolver uma pesquisa aprofundada na temática com aplicação prática em estudo de caso nas operações portuárias.

2. Referencial teórico

2.1.Revisão Sistemática

Revisão sistemática é o nome dado a uma abordagem de pesquisa que tem como objetivo coletar, avaliar e sintetizar as evidências relevantes sobre um determinado assunto. É uma ferramenta essencial para consolidar o conhecimento existente, oferecer insights confiáveis e embasar decisões para realizar uma pesquisa. Conforme destacado por Biolchini (2005) as principais etapas que compõem o desenvolvimento de uma revisão sistemática são: planejamento, revisão, execução, revisão e análise dos resultados.

O planejamento é o ponto de partida, onde é definida a pergunta de pesquisa e os critérios de inclusão e exclusão dos estudos. Isso estabelece os parâmetros para a busca e seleção dos artigos relevantes. A etapa de revisão envolve a busca em bases de dados acadêmicas e outras fontes relevantes para identificar todos os estudos pertinentes ao tema.

Na fase de execução, os estudos são avaliados, seguindo critérios pré-definidos. A análise dos resultados implica na síntese das informações extraídas dos estudos selecionados, identificando padrões, tendências e relações entre os resultados. A revisão adicional pode ocorrer após a análise, permitindo uma validação ou ajuste das etapas anteriores.

Diante disso, na revisão sistemática o pesquisador pode se fundamentar nas contribuições dos autores de análises aprofundadas apresentadas nas obras, nutrindo sua própria exploração por meio da compreensão e ressonância com o pensamento preexistente (SEVERINO, 2014).

2.2.Teoria das Filas

A teoria das filas busca estudar a avaliação da espera em filas, empregando medidas representativas de desempenho do sistema (TAHA, 2008). As filas se formam quando a procura por um serviço é superior a capacidade de atendimento desse sistema (CAMELO et al., 2010). Portanto o foco dessa teoria é analisar quanto tempo um usuário espera em fila até o atendimento usando como instrumento modelos diversos de filas (HILLIER LIEBERMAN, 2013).

O sistema de filas pode ser descrito através de um usuário chegando e esperando ser atendido, passando pelo atendimento e logo após ser atendido saindo do sistema (BRONSON, 1985). O principal objetivo de um sistema de filas é relacionar as variáveis de chegada, atendimento e capacidade de serviço, mostrando como deve ser o funcionamento desse sistema (HILLIER; LIEBERMAN, 2013). As análises sobre os sistemas de filas buscam então descobrir um ponto de

equilíbrio entre a capacidade de serviço disponível aos usuários e os custos operacionais que são gerados até o atendimento (ALVES; MENEZES, 2010).

Segundo Fraga (2012), a notação de Kendall, é uma maneira de descrever os modelos para simplificação de sua análise, utilizando cinco características: $(a/b/c):(d/e/f)$ onde: “a” significa o processo de chegada, onde essa chegada é representada pelo intervalo entre eles; “b” o processo de atendimento que é representado pelo tempo em serviço; “c” o número de servidores, ou seja, a quantidade da fonte de serviço; “d” é a regra da fila representando como é o processo de chegada e atendimento do usuário no sistema; “e” é o número máximo de clientes no sistema; “f” o tamanho máximo da população que pode ser atendida.

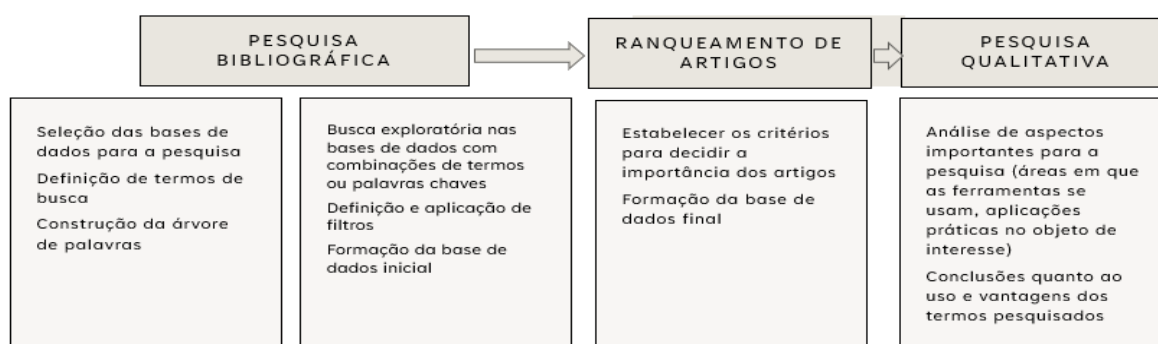
2.3. Simulação

A simulação pode ser definida como sendo uma técnica utilizada em diversas áreas com o objetivo de modelar e estudar sistemas não triviais, isso porque esta tem a habilidade de fazer a iteração entre problemas do meio corporativo e a academia (FELIX; PEIXOTO; EDWIGES, 2023), permitindo que se estude o comportamento, desempenho e interações em um ambiente ideal e virtual através da criação de modelos computacionais que buscam a semelhança do que acontece na vida real, proporcionando então detalhes de como esses sistemas podem evoluir ao longo do tempo. A simulação é cabível a tudo aquilo que é real, ou seja, significa reproduzir a performance de um sistema (CAMELO et al., 2010). É um método que emprega um modelo matemático para viabilizar a exploração e avaliação do comportamento e desempenho de um sistema, verificando a antecipação de comportamentos futuros envolvendo a criação de um ambiente virtual que espelha o sistema em análise. Isso proporciona a capacidade de explorar o comportamento e o desempenho do sistema sob diversas condições, sem a necessidade de realizar modificações no sistema real. Nesse contexto, a solução de integração se manifesta como um sistema complexo, cujo modelo é caracterizado por elementos estocásticos, dinâmicos e discretos (WIESNER, 2016).

3. Proposta metodológica

A proposta metodológica para a revisão sistemática da literatura sobre os temas de interesse pode ser apreciada na Figura 1.

Figura 1: Etapas para a revisão sistemática



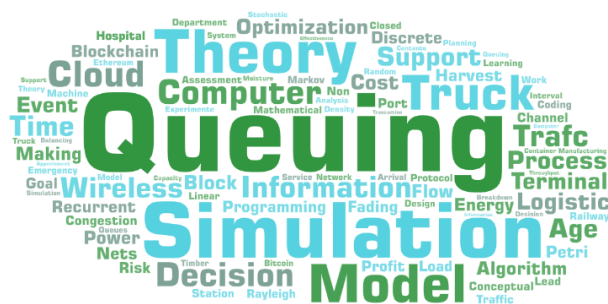
Fonte: Elaborado pelos autores

Para realizar uma revisão sistemática se parte de uma pergunta de pesquisa e de critérios de inclusão e exclusão de documentos e bibliografias para poder realizar o estudo. Nesta pesquisa pretende-se estudar aplicações de teoria de filas e simulação, para estabelecer qual das duas ferramentas dá uma melhor resposta na otimização de processos logísticos relacionados com filas de espera em processos produtivos ou de serviços.

Para a pesquisa bibliográfica, no contexto deste estudo, a *Web of Science* e a *Scopus* foram escolhidas para serem as duas plataformas de informações para a identificação e consolidação de estudos relacionados ao tema estipulado. Essas bases de dados merecem destaque devido à sua amplitude temática, abarcando variados campos do saber. Além disso, as plataformas em questão apresentam ferramentas de pesquisa avançadas, que concedem a capacidade de refinar os resultados de acordo com critérios específicos o que simplifica a seleção de estudos pertinentes e assegura um maior grau de validade e solidez à revisão bibliográfica.

Para iniciar a pesquisa foram formulados critérios de busca e empregou-se a árvore de palavras (Figura 2) como primeiro recurso para o desenvolvimento da pesquisa.

Figura 2: Árvore de palavras



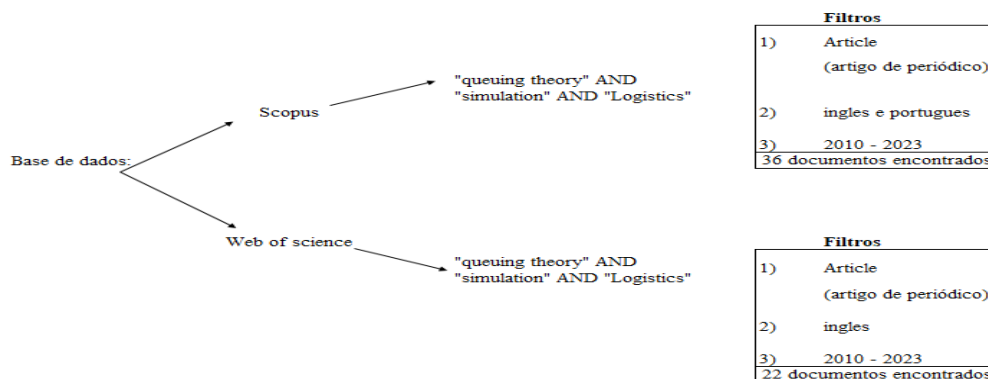
Fonte: Elaborado pelos autores

Essa abordagem permitiu uma representação das palavras mais frequentes relacionada a temática do trabalho, auxiliando na identificação de tendências e conteúdos predominantes.

A busca exploratória nas bases de dados se deu a partir da estruturação dos termos utilizando lógica booleana, uma abordagem que possibilita a organização precisa das palavras-chave com base nas particularidades das bases de dados. Esse método permitiu criar consultas de pesquisa refinadas, combinando os termos de maneira estratégica para obter resultados mais pertinentes e alinhados com os objetivos da investigação. A aplicação da lógica booleana proporciona uma abordagem sistemática para moldar a busca, maximizando a relevância das informações recuperadas e otimizando o processo de descoberta acadêmica.

Após a seleção das bases de dados e a definição dos termos de pesquisa, foram implementados filtros com o intuito de assegurar a qualidade dos dados adquiridos. Essa abordagem permitiu discernir e excluir informações que pudessem ser desatualizadas, imprecisas, antiquadas ou desprovidas de relevância para a investigação em questão. Um exemplo dos filtros utilizados é representado na Figura 3 para um conjunto de palavras-chaves.

Figura 3: Filtros base de dados



Fonte: Elaborado pelos autores

Na Tabela 1, é possível observar os resultados das pesquisas booleanas em cada uma das bases de dados.

Tabela 1: Triagem dos artigos bases "Web of Science" e "Scopus"

Triagem dos artigos	Scopus	Web of science
"queuing theory"	4222	5252
"queuing theory" AND "simulation"	1537	1686
"queuing theory" AND "logistics"	126	112
"queuing theory" AND "simulation" AND "Logistics"	52	46
"queuing theory" AND "simulation" AND "Logistics" (filtros: 2010-2023; inglês e português; artigo de periódico)	13	28

Fonte: Elaborado pelos autores

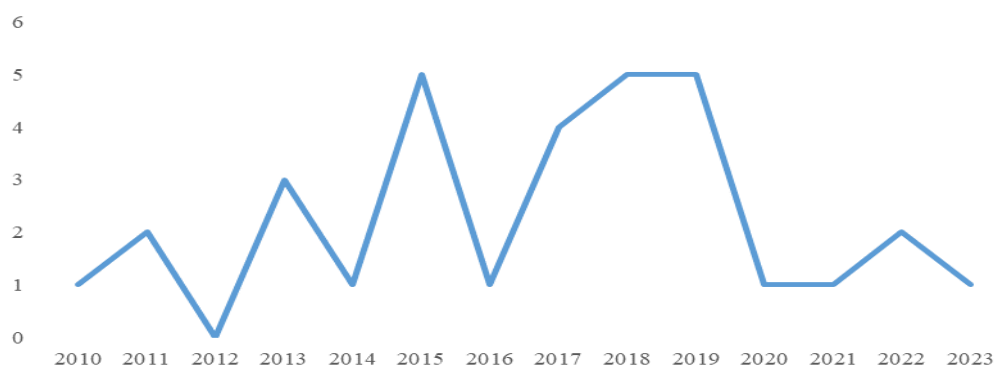
A predominância de artigos é notável quando se empregam os termos "queuing theory" e "queuing theory" AND "simulation". Isso reflete a abrangência e relevância dessas ferramentas em uma ampla gama de campos científicos. A riqueza de pesquisa nessas áreas sugere sua importância contínua no estudo de sistemas de espera e análise de processos.

No entanto, é interessante observar que ao incluir o termo "logistics" juntamente com os outros dois, a quantidade de artigos se reduz. Esse fenômeno sugere que, embora a incorporação da logística possa trazer foco e aplicabilidade específicos, ela também pode limitar a diversidade de tópicos abordados. Isso ilustra como a interseção entre esses termos influencia a amplitude das publicações e aponta para a influência direta que a logística exerce na contextualização e escopo da pesquisa.

Considerando essa situação, os conjuntos de dados provenientes do Scopus e da Web of Science foram unidos. Os dois arquivos de pesquisa foram combinados, e os artigos duplicados foram eliminados, o que resultou em um novo arquivo no formato de planilha. A base de dados da Web of Science apresentou 28 artigos, enquanto a base da Scopus continha 13 artigos. Após a remoção das duplicações, o total ficou em 32 artigos únicos. Estes 32 artigos vão compor a base de dados inicial para continuar o processo de ranqueamento proposto na Figura 1.

Na Figura 3, é visível a representação gráfica da quantidade de artigos abordando o tema estabelecido, publicados no intervalo de tempo entre 2010 e 2023. Pode-se apreciar que os anos com a maior quantidade de publicações são 2015, 2018 e 2019, cada um com um total de 5 artigos publicados. De maneira geral houve uma constância nos estudos sobre teoria das filas, simulação e logística ao longo dos anos. O decréscimo nos últimos anos, ainda pode ser um reflexo da pandemia. Isso mostra que o tema é atual com aplicações em diversos campos.

Figura 3: Quantidade de artigos por ano



Fonte: Elaborado pelos autores

A próxima etapa da proposta metodológica é definir os critérios de ranqueamento dos artigos e os pesos estabelecidos. Aqui será dividido por tópicos e critérios como mostra a Tabela 2.

Tabela 2: Tópicos e critérios para ranqueamento

Tópicos	Crítérios
<i>Journal</i> (Tem como objetivo identificar a relevância e excelência do jornal em que o artigo foi publicado)	Valor dos índices SJR/SNIP do jornal em que o artigo foi publicado.
Tema (Tem como objetivo identificar a relevância do artigo para a pesquisa. Para isso, avalia-se através dos resumos e palavras-chave dos artigos, o quanto que estão alinhados aos objetivos desta pesquisa acadêmica em questão).	Nota dada ao artigo em relação ao seu alinhamento à pesquisa em função dos seguintes temas: - Modelos de filas - Simulação de eventos discretos e teoria de filas aplicados de forma conjunta em operações logísticas - Vantagens e desvantagens do uso de cada ferramenta no mesmo processo - Modelos de filas e simulação associados a outras ferramentas de tomada de decisão

Fonte: Elaborado pelos autores

Outros critérios ainda podem ser definidos no transcurso da pesquisa. Para estabelecer as prioridades podem ser usados métodos de tomada de decisão com múltiplos critérios como o *Analytic Hierarchy Process* (AHP) ou matrizes de prioridades. Pode ser usada escala tipo Likert para a pontuação. A nota dada ao artigo em relação ao alinhamento do tema será estabelecida pela própria autora, todos os outros indicadores têm como fonte o motor de busca acadêmica das bases de dados usadas.

Com esse ranqueamento finalmente poderá ser feita uma avaliação qualitativa que pode ser auxiliada por *softwares* e decidir quais artigos vão compor a base de dados final que se usará como referência para a pesquisa em questão.

Segundo os autores pesquisados ambas as ferramentas se mostram pertinentes para resolver problemas logísticos e avaliar cenários para melhoria de processos com gargalos (FELIX; PEIXOTO; EDWIGES, 2023; FELIX; PEIXOTO; EDWIGES, 2023; DA SILVA et al. 2021). Vantagens e desvantagens podem ser analisadas para casos específicos e estabelecer modelos que permitam aproveitar as vantagens de cada ferramenta por separado ou implementar o uso de forma conjunta em projetos de pesquisa, mostra-se um assunto pertinente para um estudo mais detalhado.

Conclusões

Com base na proposta metodológica adotada neste estudo, foi realizada uma revisão dos temas de teoria das filas e simulação com aplicações na logística. A proposta do estudo estabelece a realização de três etapas, das quais somente foi concluída a primeira delas, a revisão bibliográfica.

O mapeamento inicial permitiu a construção da árvore de palavras-chave destacando as mais frequentes relacionadas ao tema. A estruturação dos termos com lógica booleana possibilitou a formulação de consultas alinhadas com os critérios de pesquisa das bases de dados. A aplicação de filtros foi importante para garantir a qualidade dos dados, eliminando informações desatualizadas e irrelevantes.

Como resultado inicial obteve-se uma base formada por 32 artigos que serão estudados e organizados segundo as etapas da proposta metodológica. Com o aprofundamento no estudo serão

identificadas lacunas do conhecimento e a pesquisa será dirigida a um setor importante da economia, o portuário. Isto permitirá dirigir o estudo para uma proposta sólida de tema de pesquisa que será usada como tema de dissertação.

Referências

ALVES, A.C.B.; MENEZES, M.A.F. Introdução à Pesquisa Operacional. 1 ed. – Goiânia: PUC Goiás, 2010.

BALLOU, R. Logística empresarial. Transportes. Administração de materiais. Distribuição física. São Paulo: Atlas, 2012.

BIOLCHINI, J. et al. Systematic review in software engineering. System engineering and computer science department COPPE/UFRJ, Technical Report ES, v. 679, n. 05, p. 45, 2005.

BRONSON, R. Pesquisa operacional. São Paulo: McGraw-Hill do Brasil, 1985.

CAMELO, G. R.; COELHO, A. S.; BORGES, R. M.; DE SOUZA, R. M. Teoria das filas e da simulação aplicada ao embarque de minério de ferro e manganês no terminal marítimo de ponta da madeira. Cadernos do IME-Série Estatística, v. 29, n. 2, p. 1, 2010.

CASTRO NETO, L. R. Modelagem e simulação da cadeia produtiva do minério de ferro. 2006. Tese de Doutorado. Universidade de São Paulo.

DA SILVA, M. R. F. et al. Proposta de otimização dos processos logísticos do porto de Itajaí utilizando simulação e indicadores da teoria das filas. Revista Produção Online, v. 21, n. 1, p. 231-258, 2021.

DE BRUNS, R.; SONCIM, S. P.; DE SINAY-PHD, M. C. F. Pesquisa operacional: uma aplicação da teoria das filas a um sistema de atendimento. 2001.

DE OLIVEIRA, G. T. A RELAÇÃO DOS INDICADORES DE DESEMPENHO DA LOGÍSTICA PORTUÁRIA COM OS INDICADORES DE DESEMPENHO DA LOGÍSTICA INTERNACIONAL. Revista Eletrônica de Estratégia & Negócios, v. 11, 2018.

FELIX, J. C.; PEIXOTO, C. S.; EDWIGES, C. A. O PLANEJAMENTO DE SUPRIMENTOS UTILIZANDO-SE A TEORIA DE FILAS E SIMULAÇÃO COMPUTACIONAL EM UMA OFICINA DE RECUPERAÇÃO DE COMPONENTES DE VAGÕES. Exacta, v. 21, n. 2, 2023.

HILLIER, F. S.; LIEBERMAN, G. J. Introdução à pesquisa operacional. McGraw Hill Brasil, 2013.

MACHADO, L. K. C.; DOS SANTOS, A. C. Índice de Desempenho Logístico (LPI): Uma análise da eficiência logística e da importância relativa dos seus indicadores. Ciências da Administração, v. 23, n. 60, p. 53-72, 2021.

TAHA, Hamdy A. Pesquisa operacional. Pearson Educación, 2008.

WIESNER, A. K. Modelagem e simulação de uma solução de integração para identificação de gargalos de desempenho baseadas em formalismo matemático: uma abordagem orientada à Teoria das Filas. 2016.

ROTEAMENTO DE MÉDICOS POR ESPECIALIDADE EM SITUAÇÃO DE PÓS - DESASTRE ATRAVÉS DE UM MODELO DE PROGRAMAÇÃO LINEAR INTEIRA MISTA

Raila de Medeiros Melo
Universidade Federal Fluminense
raila_medeiros@id.uff.br

Tiago de Araujo Neves
Universidade Federal Fluminense
tneves@id.uff.br

Eliane da Silva Christo
Universidade Federal Fluminense
elianechristo@id.uff.br

Resumo: O artigo apresenta um modelo matemático que pode auxiliar na tomada de decisão para garantir a minimização do atraso do tempo de atendimento das equipes médicas para distribuição de socorro na fase de resposta a um desastre, considerando as especialidades médicas da equipe e a demanda por estas especialidades. Para construção da abordagem exata é utilizado um modelo matemático similar ao usado para problemas de roteamento com janelas de tempo. Os resultados mostram que a abordagem exata entrega soluções viáveis aos decisores, porém, para grandes instâncias, é necessário um tempo computacional relativamente alto.

Palavras-chave: Logística Humanitária. Modelagem Matemática para Distribuição de Socorro Médico. Roteamento Médico. Modelagem Computacional. Programação Linear Inteira Mista.

Introdução

Furacões, enchentes, tempestades, rompimento de barragens entre muitos outros desastres afetam diversas pessoas anualmente e causam imensuráveis danos ao redor do mundo. De acordo com CRED (2021), desde 1990 houve 16.165 desastres naturais em todo o mundo, atingindo cerca de 8,3 bilhões de pessoas, gerando 32,4 milhões de mortes e um custo aproximado de 5,4 bilhões de dólares. O Brasil ocupa a 10ª posição no ranking de países afetados pelos desastres segundo o CRED (2021). Segundo a CNN Brasil (Janone, 2022), no início de 2022, aproximadamente 8 milhões de brasileiros foram afetados por desastres ambientais. Eventos recentes, como os rompimentos de barragens em Mariana (2015) e Brumadinho (2019) em Minas Gerais, deixaram um impacto duradouro na população e causaram a necessidade de auxílio humanitário. O

acidente de Brumadinho, por exemplo, levou mais de dez mil pessoas a deixarem suas casas (Folha de S.Paulo, 2022). Diante desses desastres devastadores, a logística humanitária (LH) surgiu como uma forma de gerenciar crises e prestar ajuda em momentos críticos. A LH busca organizar, coordenar e efetuar assistência de maneira eficaz, abrangendo evacuações, distribuição de alimentos, água, suprimentos de emergência, medicamentos, voluntários, equipes médicas e a gestão de resíduos pós-desastre. Com base nesse cenário, este estudo apresenta um modelo matemático para reduzir atrasos nos atendimentos, considerando as especialidades médicas da equipe e as demandas específicas por essas especialidades.

Características e Formulação do Problema

O problema tratará a distribuição de auxílio médico aos necessitados, buscando através da função objetivo minimizar o atraso global nos atendimentos de todos pacientes. Conforme Hezam et al.(2021), esta área tem sido pouco considerada nos estudos científicos. Na revisão literária sobre LH desenvolvida em Hezam et al. (2021), apenas dois estudos são apontados para distribuição de socorro através de auxílio médico utilizando modelos determinísticos, que são Wang et al. (2018) e Lei et al. (2015). Duas particularidades de ambos os estudo chamam a atenção. Nenhum deles aborda as demandas por especialidade e conjunto médico disponível. Mediante ao que foi exposto, esse estudo busca contribuir com a literatura abordando estas duas lacunas. O presente trabalho considera que cada paciente tem demanda por uma única especialidade. Caso o paciente precise ser atendido por mais de uma especialidade ele deve ser inserido múltiplas vezes (uma para cada especialidade demandada). Todas demandas devem ser atendidas. Considera-se que há uma central que possa concentrar as informações. Estas informações se referem à localização geográfica das demandas, suas necessidades médicas (em termos de especialidades e gravidade), a quantidade de médicos disponíveis e as especialidades (pediatria, ortopedia, etc) de cada médico (um médico pode possuir mais de uma especialidade), o tempo médio de atendimento por especialidade médica, o tempo limite para iniciar o atendimento de cada paciente (esse tempo indica a gravidade do paciente. Pacientes em situações críticas devem ter valores menores que pacientes em situações menos críticas), o tempo de deslocamento entre cada par de localidades, tempo de descanso médico entre rotas (este tempo pode variar em função de condições individuais de cada médico como, por exemplo: idade, restrições físicas, cuidados especiais, etc) e tempo limite para a duração das rotas (es-

tes valores serão usados para evitar a exaustão da equipe médica). Também considera-se que existe um centro médico, onde os profissionais se preparam para atender as demandas (reunem equipamento e suprimentos médicos) e que as rotas começam e terminam neste centro. De posse destas informações, a central determinará a rota a ser seguida por cada médico e repassar aos pontos de demanda os horários estimados para os atendimentos. A pesquisa construída pelos autores Kabcome e Mouktonglang (2015) foi utilizada como norteadora para a construção do modelo matemático proposto neste estudo. Abaixo as notações adotadas para a construção do modelo matemático:

Conjuntos(índices): Seja $G(N,A)$ um grafo direcionado onde N é um conjunto de vértices e A é um conjunto de arcos (i,j) representando a conexões entre cada par de localidades (incluindo o centro médico), onde $i \neq j$. Seja:

$N = \{o, d, 1, 2, 3, \dots, n\}$ seja o conjunto de nós (o a origem e d o destino. No modelo ambos serão o mesmo local físico, que é o centro médico). $N' = N \setminus \{o, d\}$. $A = \{(i,j) : i, j \in N, i \neq j\}$, para o conjunto de arcos. $K = \{1, 2, 3, \dots, k\}$, para todo conjunto de médicos. $P = \{1, 2, 3, \dots, p\}$, para todo conjunto de especialidade médica. $R = \{1, 2, 3, \dots, r\}$, para todo conjunto de rotas de cada médico.

Dados de entrada:

B_i é o tempo limite para início do atendimento do paciente i ; T_{ij} é o tempo de deslocamento entre as localidades onde estão os pacientes i e j ; S_i é o tempo de atendimento do paciente i ; C_k é o tempo necessário para descanso entre viagens do médico k ; Q_k é o limite de tempo para rotas do médico k ;

$$H_{kp} = \begin{cases} 1, & \text{se o médico } k \text{ possui habilitação na especialidade } p; \\ 0, & \text{caso contrário.} \end{cases}$$

$$D_{ip} = \begin{cases} 1, & \text{se o paciente } i \text{ demanda a especialidade } p; \\ 0, & \text{caso contrário.} \end{cases}$$

Variáveis de decisão:

$Y_{k,r,i}$ é o atraso no atendimento do paciente i na r -ésima viagem do médico k ; $W_{k,r,i}$ é o tempo de início do atendimento do paciente i pelo médico k em sua r -ésima rota;

$$X_{k,r,i,j} = \begin{cases} 1, & \text{se o paciente } j \text{ é atendido imediatamente após o paciente } i \text{ na } r\text{-ésima} \\ & \text{viagem do médico } k; \\ 0, & \text{caso contrário.} \end{cases}$$

Modelo matemático: O problema de distribuição de socorro médico por especialidade buscando atender diferentes localidades e suas demandas pode ser descrito matematicamente da seguinte forma:

$$\text{Min } Z = \sum_{i=1}^{N'} \sum_{k=1}^K \sum_{r=1}^R Y_{k,r,i} \quad (1)$$

s.a.

$$Y_{k,r,i} \geq W_{k,r,i} - B_i, \quad \forall i \in N', \forall k \in K, \forall r \in R; \quad (2)$$

$$W_{k,r,i} + S_i + T_{i,j} - W_{k,r,j} \leq (1 - X_{k,r,i,j})M, \quad \forall k \in K, \forall r \in R, \\ \forall i, j \in N, i \neq j; \quad (3)$$

$$W_{k(r+1),o} \geq W_{k,r,d} + C_k, \quad \forall k \in K, \forall r \in \{1, \dots, (|R| - 1)\}; \quad (4)$$

$$W_{k,r,d} - W_{k,r,o} \leq Q_k, \quad \forall k \in K, \forall r \in R; \quad (5)$$

$$W_{k,r,o} \leq W_{k,r,d}, \quad \forall k \in K, \forall r \in R; \quad (6)$$

$$\sum_{j=1}^{N \setminus \{d\}} X_{k,r,j,i} - \sum_{j=1}^{N \setminus \{o\}} X_{k,r,i,j} = 0, \quad \forall k \in K, \forall r \in R, \forall i \in N'; \quad (7)$$

$$\sum_{k=1}^K \sum_{r=1}^R \sum_{j=1}^{N \setminus \{d\}} X_{k,r,j,i} = 1, \quad \forall i \in N'; \quad (8)$$

$$\sum_{j=1}^{N \setminus \{o\}} X_{k,r,o,j} \leq 1, \quad \forall k \in K, \forall r \in R; \quad (9)$$

$$\sum_{j=1}^{N'} X_{k,r,o,j} = \sum_{j=1}^{N'} X_{k,r,j,d}, \quad \forall k \in K, \forall r \in R; \quad (10)$$

$$X_{k,r,i,j} = 0, \quad \forall i \in N', \forall p \in P, \forall k \in \{z : z \in K, D_{i,p} \neq 0, \\ H_{k,p} = 0, \forall j \in N, \forall r \in R\}; \quad (11)$$

$$Y_{k,r,i}, W_{k,r,i}, X_{k,r,i,j} \geq 0 \text{ e } Y_{k,r,i}, W_{k,r,i}, X_{k,r,i,j} \in \mathbb{Z}, \\ \forall i, j \in N, \forall p \in P, \forall k \in K, \forall r \in R. \quad (12)$$

A Equação (1) representa a função objetivo do problema. Ela visa minimizar o somatório dos atrasos em todas as localidades, a medida que as restrições (2) permitirão que o problema calcule o atraso. No caso do valor de B ser maior que W , ou seja não há atraso, as restrições de não negatividade farão com que Y assumam zero. Por sua vez as restrições (3) certificam que, se o paciente j é atendido pelo médico k em sua r -ésima viagem imediatamente após o paciente i , então o tempo de início do atendimento do paciente j será igual ao tempo de início do atendimento do paciente i somado com o tempo de atendimento de i e com o tempo de deslocamento de i até j bem como as restrições (4) garantem que o médico k só poderá sair para uma nova viagem após o descanso

obrigatório estipulado. As restrições (5) e (6) garantem que o tempo limite das rotas sejam respeitados o tempo de chegada ao destino seja maior (ou igual no caso de rota vazia) que o tempo de partida na origem. Em adição, as restrições (7) representam o equilíbrio de fluxo. Dizem que se um médico chegou a um paciente em uma viagem ele deve sair deste mesmo paciente na mesma viagem. Por outro lado as restrições (8) dizem que todos os pacientes devem ser visitados uma única vez, enquanto nas restrições (9) está limitando o número de partidas por rota a partir da origem. Já as restrições (10) certificam que se uma viagem foi iniciada, ela deve ser finalizada. Da mesma forma as restrições (11) garantirão que um paciente não seja atendido por um médico sem a especialidade requerida. Por fim, as restrições (12) são as restrições de não negatividade e integralidade do problema.

Para testar os algoritmos construídos para o problema foram criadas instâncias artificiais com dados de oferta e demanda. Contudo, para tornar os cenários mais realísticos, dados referente às posições geográficas dos pacientes, centro médico, limites de tempo das rotas e o tempo de descanso entre rotas são reais (Melo, 2023). É válido ressaltar que as instâncias apresentadas no trabalho são maiores que encontradas na literatura.

Há 4 possíveis tamanhos para o conjunto de pacientes: 5, 10, 15 e 20 pacientes. Cada paciente possui: 1) um tempo máximo para início de atendimento; 2) um tempo de atendimento, iguais aos usados por Lei et al. (2015); 3) uma posição geográfica (unidades básicas de saúde da cidade de Angra dos Reis, RJ-Brasil) e 4) uma demanda médica gerada de acordo com três distribuições estatísticas: Uniforme, Normal e Zipf. Os dados foram gerados usando Python (Numpy, 2022) e a ferramenta Google Colaboratory® (Borges, 2014). Também foram consideradas dois possíveis tamanhos para o conjunto de médicos: 4 e 6 médicos. Cada médico possui duas especialidades, escolhidas arbitrariamente para evitar a criação de instâncias inviáveis. Os tempos de descanso entre as rotas para cada médico, foram determinados com base em valores usados no dia-a-dia das equipes de resgate (entre 8 e 12 horas de descanso). O tempo máximo de cada rota também foi determinado com bases em dados reais (24 horas) e se estabeleceu que cada médico pode realizar, nas instâncias propostas, no máximo três rotas. Detalhes da criação das instâncias podem ser encontrados em Melo (2023). Forneceu-se ao *software* CPLEX ® o tempo limite de execução de 10.800 segundos. O resolvidor para voluntariamente em duas situações: gap (diferença relativa entre o melhor valor inteiro e o valor objetivo) igual a zero ou em casos em que não há memória suficiente no computador. Outra maneira é quando o usuário insere critérios de parada, tal como realizado

neste estudo. A Tabela 1 apresenta uma síntese global de todos os resultados obtidos.

Tabela 1: Síntese dos atrasos, gap do erro empírico e tempo de máquina para cada uma das distribuições de acordo com as instâncias

4 médicos									
Localidades	Uniforme			Normal			Zipf		
	Atraso	Tempo (s)	Gap (%)	Atraso	Tempo (s)	Gap (%)	Atraso	Tempo (s)	Gap (%)
5	87	0,06	0,00%	110	0,05	0,00%	140	0,19	0,00%
10	288	3,25	0,00%	288	3,55	0,00%	382	26,08	0,00%
15	742	10807,19	71,73%	834	10257,56	78,11%	903	10803,48	84,34%
20	1678	10806,72	98,62%	1761	10804,19	99,80%	1795	10817,33	91,42%
6 médicos									
Localidades	Uniforme			Normal			Zipf		
	Atraso	Tempo (s)	Gap (%)	Atraso	Tempo (s)	Gap (%)	Atraso	Tempo (s)	Gap (%)
5	87	0,09	0,00%	87	0,13	0,00%	126	0,05	0,00%
10	193	2,48	0,00%	170	1,44	0,00%	242	5,45	0,00%
15	411	445,42	0,00%	675	10811,58	57,17%	1252	10811,91	93,46%
20	932	10807,90	74,69%	1636	10801,83	94,91%	2796	10812,44	99,11%

Fonte: Elaborado pela autora.

Os resultados apresentados na Tabela 1 demonstram que 54% das simulações alcançaram o gap igual a zero. Observa-se também que o resolvidor demandou um intervalo de tempo relativamente significativo para se adequar às condições propostas em grande parte das simulações.

Conclusões

A LH é uma ferramenta eficaz nos momento caóticos gerados pelos desastres. Através dela a coordenação e gestão adequada podem ser implementadas, integrando os principais envolvidos no objetivo de auxílio adequado e eficaz aos necessitados e o restabelecimento da ordem.

O modelo matemático proposto foi amplamente testado, abordando diferentes instâncias (demandas, disponibilidade de especialidades, médicos especialistas, etc). Os resultados apresentados na Tabela 1 indicam que quanto maior a quantidade de pacientes, maior a dificuldade do resolvidor para encontrar a solução ótima. Também é possível notar que o resolvidor apresenta muita dificuldade para solucionar as maiores instâncias visto que, mesmo após o limite de tempo, o melhor valor para o gap foi maior que 50%. Com isto conclui-se que a abordagem alcançou soluções em um tempo computacional relativamente alto para as condições impostas para cenários de desastre.

Com base nos resultados obtidos nesta pesquisa, é possível concluir também que é necessário considerar propostas que exijam menos tempo computacional. Todos os valores alcançados pelo modelo proposto estão disponíveis como pontos de referência para comparações.

Referências

Borges, L. E. (2014). **Python para desenvolvedores: aborda Python 3.3**. Novatec Editora.

CRED (2021). **Database de desastres naturais**. <https://www.emdat.be>. Acessado em 13 novembro de 2021.

de Oliveira, E. F. e Soares, E. B. (2019). **Logística humanitária: o desafio da gestão diante de desastres**. Brazilian Journal of Business, 1(3):870–880.

de S.Paulo, J. F. (2022). **Cinco pessoas morrem em carro soterrado em Brumadinho (MG)**. <https://www1.folha.uol.com.br/cotidiano/2022/01/cinco-pessoas-morrem-em-carro-soterrado-em-brumadinho-mg.shtml>. Acessado em 06 agosto de 2022.

Hezam, I. M. et al. (2021). **A systematic literature review on mathematical models of humanitarian logistics**. Symmetry, 13(1):11.

Janone, L. (2022). **A cada desastre natural no brasil, em média, 3,4 mil pessoas são afetadas**. CNN Brasil. Acessado em 06 agosto de 2022.

Kabcome, P. e Mouktonglang, T. (2015). **Vehicle routing problem for multiple product types, com- partments, and trips with soft time windows**. International Journal of Mathematics and Mathe- matical Sciences, 2015.

Lei, L., Pinedo, M., Qi, L., Wang, S., e Yang, J. (2015). **Personnel scheduling and supplies provisi- oning in emergency relief operations**. Annals of Operations Research, 235(1):487–515.

Melo, R.M. **Modelo Heurístico e Matemático para Roteamento de Médicos por Especialidade em Situação de Pós-Desastre**. 2023. 95 f. Trabalho de Conclusão de Mestrado – Universidade Federal Fluminense, UFF, Rio de Janeiro, 2023.

Numpy (2022). **Random Generator Python**. <https://numpy.org/doc/stable/reference/random/generator.html>.2022.

Wang, S., Liu, F., Lian, L., Hong, Y., e Chen, H. (2018). **Integrated post disaster medical assistance team scheduling and relief supply distribution**. The International Journal of Logistics Manage- ment.

ANÁLISE DO ROMPIMENTO DA BARRAGEM DA UHE BARRA DO BRAÚNA UTILIZANDO O SOFTWARE DE SIMULAÇÃO HIDRODINÂMICA IBER

Sânzzia de Oliveira e Souza Figueira

Instituto Federal Fluminense

sanzziaf@gmail.com

Thiago Jordem Pereira

Universidade Federal Fluminense

tjordem@id.uff.br

Ricardo Silveira Sousa

Universidade Federal Fluminense

rsousa@id.uff.br

Tibério Borges Vale

Universidade Federal Fluminense

tiberio_vale@id.uff.br

Jader Lugon Junior

Instituto Federal Fluminense

jlugonjr@gmail.com

Wagner Rambaldi Telles

Universidade Federal Fluminense

wtelles@id.uff.br

Resumo: O número de acidentes decorrentes do rompimento de barragens é cada vez maior, principalmente quando estas se localizam nas proximidades das regiões urbanizadas. Nos últimos anos, especialmente no Brasil, muitos deles geraram tanto impactos ambientais e socioeconômicos quanto de perdas humanas. Antes dessas construções, vários estudos são realizados como meio de minimizar ou mesmo não permitir que ocorram desastres futuros. Desse modo, em virtude de uma provável catástrofe envolvendo as barragens, foi realizada uma simulação da hipotética ruptura da barragem da Usina Hidrelétrica Barra do Braúna, situada no rio Pomba, tendo a cidade de Santo Antônio de Pádua-RJ como estudo de caso. Os resultados obtidos, utilizando o software de simulação hidrodinâmica IBER, mostram que tal evento, caso venha a ocorrer, representa significativo risco para a população e construções localizadas a jusante à barragem.

Palavras-chave: Rompimento de barragem. Usina Hidrelétrica Barra do Braúna. Santo Antônio de Pádua. Software IBER

Introdução

A água é considerada um dos recursos naturais indispensáveis na manutenção da vida de todo o planeta, seja relacionada à vida humana ou ao meio ambiente. Quando vista em termos quantitativos, sua escassez prejudica no modo como é utilizada pelos seres vivos, bem como nos sistemas comportamentais e fisiológicos desenvolvidos a fim de se adequar às diversas condições ambientais trazidas pela gama da distribuição temporal e especial dos recursos hídricos (MENDONÇA, 2013).

Um dos maiores problemas enfrentados pelo mundo se baseia no planejamento e na execução desses recursos hídricos a fim de que a água não se torne escassa e chegue a toda população, sem exceção. Diante dessa situação, inúmeras barragens foram construídas, contendo diversas vantagens em seu uso, sendo a principal delas, o fornecimento de energia elétrica para toda a população.

Em contrapartida, devido à variedade de tamanhos, bem como seus usos, nem todas as barragens são mantidas em perfeitas condições, não atendendo a normas de segurança, o que pode resultar no rompimento da estrutura.

Diante disso, os estudos envolvendo a ruptura de barragens vêm aumentando, despertando o interesse de órgãos ambientais como forma de tornar eficaz as técnicas seguras de manutenção e de operação das mesmas (TSCHIEDEL, 2022).

Por meio dos avanços tecnológicos, Santos et al. (2003) estabelecem que os modelos matemáticos e físicos passaram a ser utilizados como modo de responder aos efeitos de ruptura da barragem. Assim, o uso de tais modelos simplifica no momento em que corre a simulação, o que, consequentemente, favorece no estudo das barragens.

De modo especial, a presente pesquisa tem como objeto de estudo a análise do rompimento hipotético da barragem da Usina Hidrelétrica (UHE) Barra do Braúna, localizada no estado de Minas Gerais, através do software IBER, tendo como estudo de caso a cidade de Santo Antônio de Pádua, situada no estado do Rio de Janeiro.

Para a estrutura do trabalho, são apresentados os resultados e discussões baseados na construção do modelo digital da região de interesse, da variação das malhas espaciais e da calibração e validação do modelo para o cenário de rompimento hipotético, assim como a análise dos resultados numéricos e suas respectivas conclusões.

Materiais e Métodos

Nesta seção é caracterizada a UHE Barra do Braúna, bem como o processo de obtenção de dados hidrometeorológicos da região de interesse e o software de simulação hidrodinâmica IBER.

UHE Barra do Braúna

A UHE situa-se no rio Pomba e o eixo de sua barragem localiza-se entre as cidades mineiras de Laranjal e Recreio, além de estar próxima de cidades fluminenses, como Santo Antônio de Pádua, Itaocara, Aperibé e Cambuci. As águas de seu reservatório também ocupam as cidades de Leopoldina e Cataguases, no estado de Minas Gerais. A máxima vazão de projeto dessa barragem é de 3.296 m³/s, o que significa que vazões excedentes representam o seu quesito de emergência.

A cidade escolhida como base para o estudo dos efeitos do rompimento hipotético da barragem da UHE Barra do Braúna, Santo Antônio de Pádua, deve-se ao fato da mesma ser a primeira cidade que se encontra a jusante da barragem e que possui o rio Pomba passando dentro de seus limites, dividindo-a ao meio. Um rompimento da barragem atingiria a área mais urbanizada da cidade, causando enormes danos a toda a população.

Santo Antônio de Pádua possui cinco pontes localizadas em sua área urbana, as quais ligam uma margem à outra do rio Pomba. Três delas são destinadas à passagem de veículos e duas pontes de pedestres. Além disso, há outras duas pontes situadas no trecho do rio Pomba entre a represa e a referida cidade, mais especificamente, nos distritos de Cisneiros e Itapiruçu, respectivamente, os quais pertencem ao município de Palma, estado de Minas Gerais. A localização das referidas pontes é mostrada na Tabela 1. Por outro lado, na Figura 1 são representadas as pontes da zona urbana da cidade de Santo Antônio de Pádua.

Como forma de comparação dos resultados, foi tomado como base o nível d'água em cada uma das 6 pontes que há no trecho analisado, ou seja, entre a UHE Barra do Braúna e Santo Antônio

de Pádua, as quais possibilitam o fluxo de veículos automotores. Nesse sentido, não foram consideradas as pontes Alberíades Gabry e Abel da Silva Malafaia, descrita na Tabela 1, uma vez que esta é destinada apenas a passagem de pedestres.

Tabela 1: Relação das pontes situadas entre a represa e a cidade de Santo Antônio de Pádua.

Descrição da Ponte	Cidade	Localização (UTM WGS84 23S)
Cisneiros	Palma-MG	773143,403; 7626548,468
Itapiruçu	Palma-MG	774530,326; 7624890,031
Paraquena-Campelo	Santo Antônio de Pádua-RJ	784268,138; 7620655,88
Badih Chicralla	Santo Antônio de Pádua-RJ	789371,786; 7617911,681
Alberíades Gabry	Santo Antônio de Pádua-RJ	791782,059; 7615496,07
Abel da Silva Malafaia	Santo Antônio de Pádua-RJ	792016,947; 7615189,114
Raul de Moraes Veiga	Santo Antônio de Pádua-RJ	792395,971; 7614420,389
Paulino Padilha	Santo Antônio de Pádua-RJ	773143,403; 7626548,468



Figura 1: Pontes localizadas na zona urbana de Santo Antônio de Pádua.

Para a análise do rompimento hipotético da barragem da UHE Barra do Braúna, neste trabalho, é utilizado o software IBER, o qual é um modelo que simula o fluxo turbulento, de superfície livre, em escoamento não permanente. Dentre alguns de seus campos de aplicação estão: ruptura de barragens; avaliação das zonas de inundação; cálculo hidráulico das redes de canais e de canalização; fluxo de marés em estuários; transporte de sedimentos; qualidade da água; e outros (SARTORI, 2018).

O IBER compõe-se em diferentes módulos ligados entre si, como por exemplo, o módulo hidrodinâmico, que determina características do escoamento, tais como a velocidade e a profundidade da água; módulo de turbulência; e módulo de transporte de sedimentos (CHARGEL, 2018).

Dados Hidrometeorológicos

O Brasil é um país que apresenta uma grande extensão territorial, todavia, os períodos de ocorrências das chuvas e estações de seca são bem definidos. Por isso, a enorme necessidade de se acompanhar todas as formas de chuva, realizando a gestão eficiente dos reservatórios, a fim de que os

mesmos garantam que haja recursos hídricos, mesmo em épocas mais secas. Nesse sentido, Sousa et al. (2012) definem o Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) como a Política Nacional de Recursos Hídricos, inscrito na Lei nº 9433/1997, que coleta, trata, armazena e recupera todos os recursos hídricos.

No Brasil, existem dois bancos de dados ambientais de suma importância: Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) e Portal Hidroweb, mantido pela Agência Nacional de Águas e Saneamento Básico (ANA). Ambos são de livre acesso e gratuito (ROSSETTI, 2008).

Por meio do aplicativo do Hidroweb, as estações hidrometeorológicas de interesse são selecionadas, sendo que cada uma delas apresenta um período diferente de funcionamento. É possível, por exemplo, fazer download dos dados das estações de interesse nos formatos do Access (.mbd), texto (.txt) e planilha eletrônica (.csv).

Ao selecionar uma estação de interesse é possível verificar o código, a bacia, a sub-bacia, o rio, o estado, o município, o responsável, a operadora e o tipo da estação. Quando há dados telemétricos disponíveis, é possível definir o período inicial e final da pesquisa, não excedendo 90 dias. Em particular, na região situada entre a barragem da UHE Barra do Braúna e a cidade de Santo Antônio de Pádua, há quinze estações, as quais dispõem de dados convencionais e/ou telemétricos.

No presente trabalho, como forma de calibrar e validar o modelo desenvolvido no software IBER, foram utilizados dados de duas estações de monitoramento fluviométrico localizadas a jusante da UHE Barra do Braúna, no período entre o dia 24 de janeiro de 2020 a 28 de janeiro de 2020:

- Estação 58788600: localizada próximo à barragem e utilizada como condição de contorno (dado de entrada) no modelo;
- Estação 58774000: localizada próximo à zona urbanizada de Santo Antônio de Pádua, de operacionalização do INEA, e utilizada para comparação com os resultados numéricos obtidos por meio das simulações no software IBER.

Resultados e Discussões

Para as simulações realizadas nesse trabalho, foi definido um domínio de interesse entre a barragem e o município de Santo Antônio de Pádua, sendo adotada uma malha não estruturada com elementos triangulares de diferentes tamanhos de acordo com o uso e ocupação do solo. Na região que delimita o reservatório e planície, foi utilizada uma malha não estruturada com elementos triangulares de tamanho 200 m enquanto na barragem, rio e área urbanizada, foram adotados, respectivamente, elementos de tamanhos iguais a 10 m, 100 m e 50 m. Já, para os coeficientes de rugosidade de Manning foram atribuídos os valores de $0,025 \text{ s/m}^{1/3}$ (reservatório, barragem e rio), $0,050 \text{ s/m}^{1/3}$ (planície) e $0,150 \text{ s/m}^{1/3}$ (área urbanizada). Na Figura 2 é mostrado domínio de interesse.

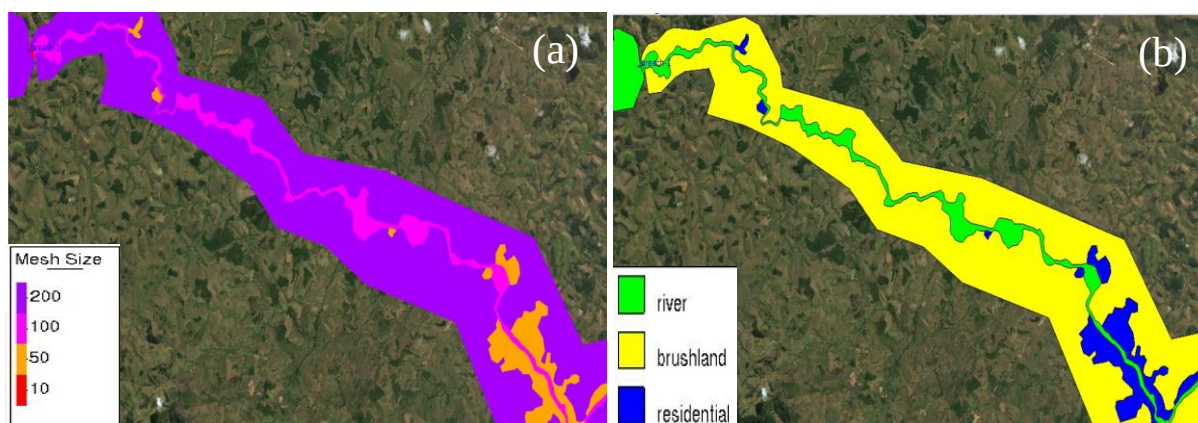


Figura 2: (a) Representação das superfícies e seus respectivos elementos e (b) caracterização das superfícies da região de estudos de acordo com o uso e ocupação do solo.

Calibração e Validação do Modelo

Para a validação e calibração do software IBER para a região de interesse, foram utilizados os dados das estações de monitoramento da ANA, a fim de caracterizar o evento de cheia do rio Pomba no período compreendido entre 24 e 28 de janeiro de 2020. Na Figura 03, é apresentada a comparação entre os resultados numéricos obtidos no software IBER e os dados experimentais provenientes da estação de monitoramento da ANA/INEA.

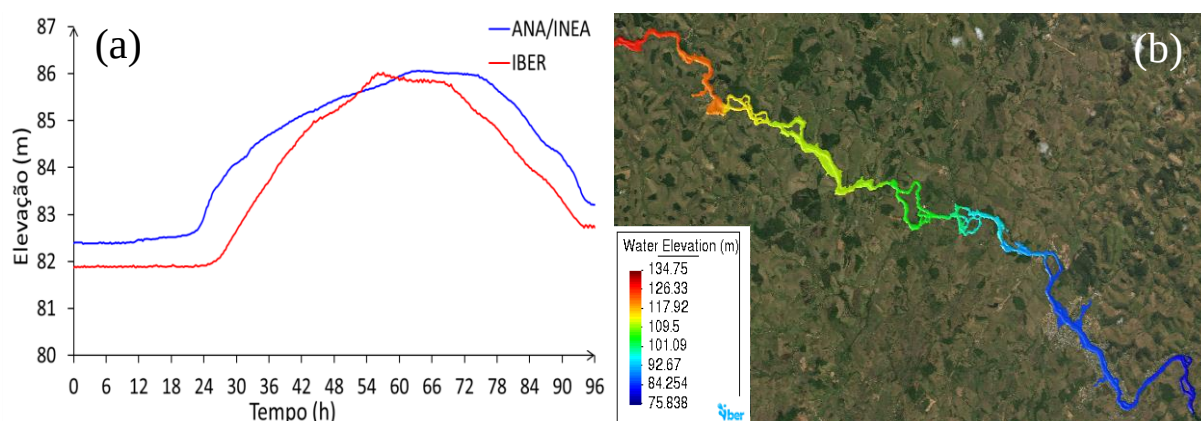


Figura 3: (a) Perfis da elevação da lâmina d'água na estação da ANA/INEA e (b) perfis da elevação máxima da lâmina d'água no rio Pomba na região analisada.

De acordo com a Figura 03, verifica-se que o modelo desenvolvido no software IBER simulou o pico da onda de cheia de forma satisfatória quando comparado com os dados experimentais obtidos no site da ANA/INEA, uma vez que representou o pior momento da onda de cheia na cidade de Santo Antônio de Pádua. Todavia, desconsiderou os níveis d'água antes e após o referido pico.

Afere-se, ainda, que a maior profundidade da lâmina d'água atingida tanto no software, quanto nos dados obtidos na estação da ANA/INEA, localizada próxima à zona urbanizada de Santo Antônio de Pádua, esteve próxima a 6 m, a qual ficou estabelecida entre um dado instantâneo de tempo de 55 h a 65 h, aproximadamente, com elevação em torno de 86 m.

Simulação de Cenários Hipotéticos de Rompimento de Barragem

Baseado na configuração descrita anteriormente, foi simulado um cenário para o rompimento hipotético da barragem da UHE Barra do Braúna, para o qual foi tomado como referência os dados disponíveis no trabalho de Daru et al. (2013), o qual contém as seguintes informações: (i) largura média da brecha (m): 91,5; (ii) tempo de ruptura da barragem (h:min): 1:00; (iii) elevação da lâmina d'água no reservatório (m): 154,5.

Para a simulação desse cenário, foi considerado um tempo total de simulação igual a 24 horas. Já o rompimento total variou conforme o tempo de ruptura, contados a partir do instante inicial da simulação, o qual também foi considerado como o início da formação da brecha. Na Figura 4, são mostrados os perfis das elevações da lâmina d'água para o cenário simulado.

Além dos perfis das elevações da lâmina d'água nos locais onde as seis pontes de passagem de veículos ao longo do trecho analisado estão inseridas, levando em consideração o cenário simulado, também foi realizada uma análise dos mapas de máximos para elevações da lâmina d'água, representada na Figura 05.

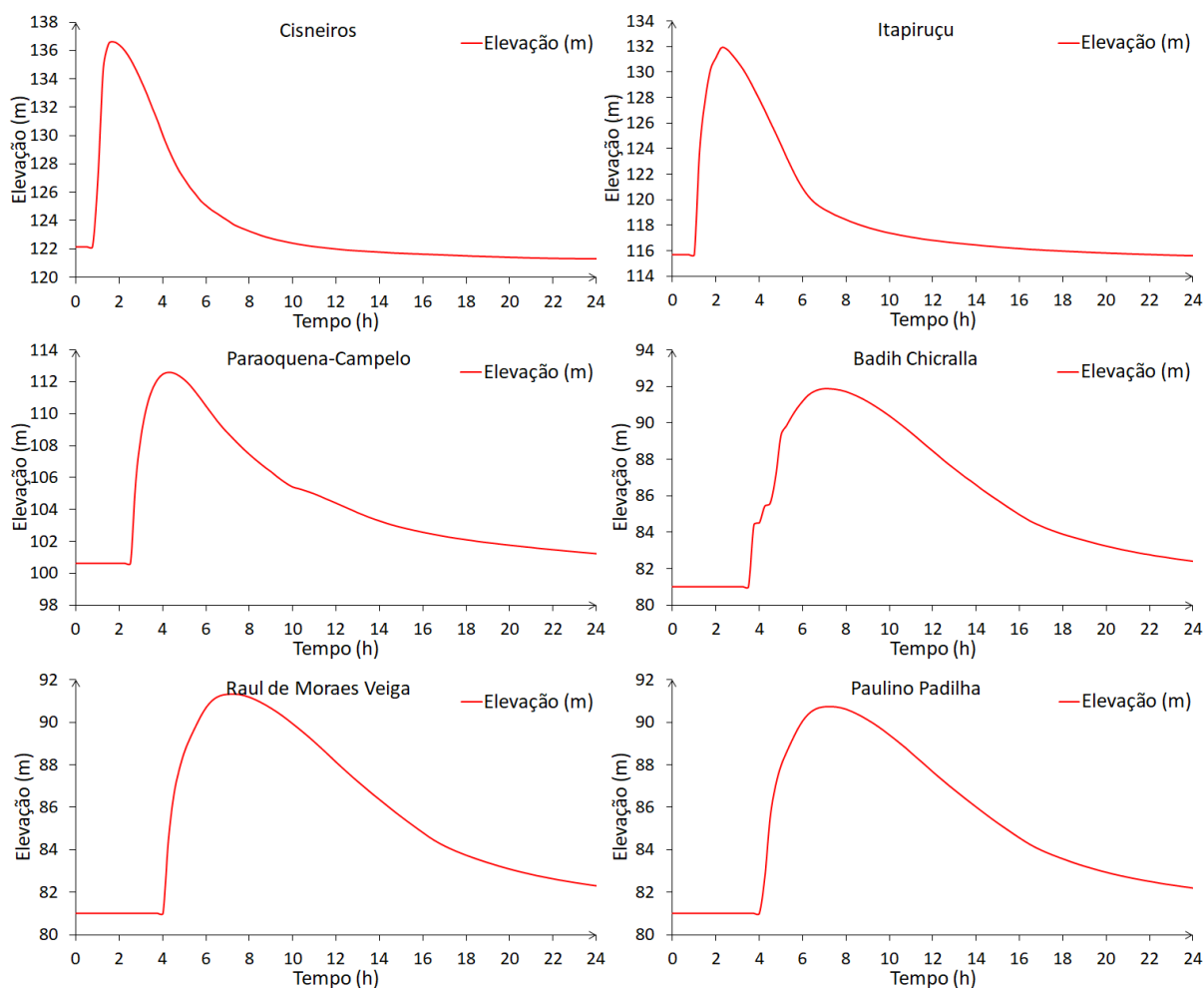


Figura 4: Perfis das elevações da lâmina d'água no cenário simulado com base nas 6 pontes ao longo de 24 h.

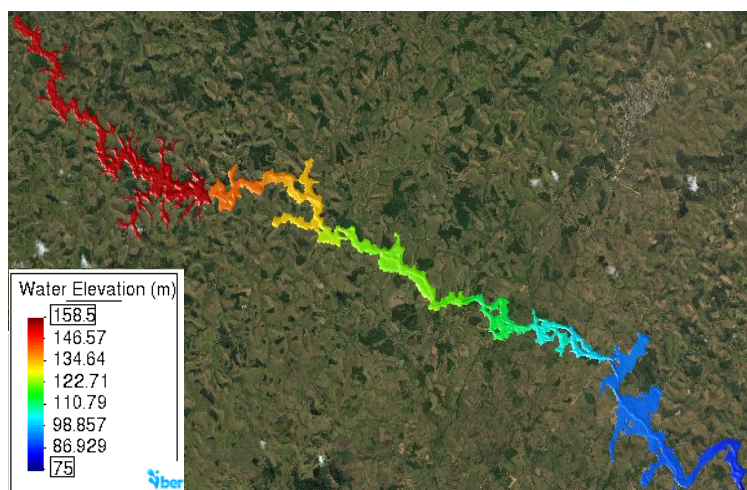


Figura 5: Mapa de máximo da elevação da lâmina d'água referente ao cenário simulado.

Conclusões

A análise do rompimento hipotético da barragem UHE Barra do Braúna mostrou a necessidade indispensável da fiscalização dos responsáveis técnicos no decorrer e, posteriormente, à sua construção, bem como por órgãos competentes, visto que a sua ocorrência gera problemas irreversíveis, na maior parte dos casos, em especial, à zona urbana de Santo Antônio de Pádua.

Além disso, baseando-se no evento de inundação real simulado no software IBER (dados das estações da ANA/INEA), constata-se que o modelo o representou de modo satisfatório, o que indica uma boa robustez da metodologia aqui proposta para a região investigada

Perante as simulações realizadas no software IBER, é possível assegurar a viabilidade do modelo utilizado para o estudo, considerando-o de grande valia de cunho ambiental, em virtude de promover a simulação dos diferentes cenários e, com isso, fazer do seu uso uma ferramenta auxiliar no processo para a tomada de decisões, em se tratando de eventos de ordem catastrófica, como as ondas de cheias, ocasionadas pelo volume de água oriundo do rompimento de uma barragem.

Agradecimentos

Esse estudo foi apoiado pela Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (E-47/2021-SEI260003/016517/2021-E-26/210.107/2022).

Referências

CHARGEL, L. T. **Avaliação do uso de modelos hidrodinâmicos, suas limitações, efeitos de escala e da representação topográfica**. 2018. 296 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, 2018.

DARU, R. L.; SILVEIRA, R. M.; HAINOSZ, F. S.; BURAS, M. Propagação de ondas de ruptura e plano de ações emergenciais - UHE Barra do Braúna. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, Bento Gonçalves (RS), 2013. Anais do **XX Simpósio Brasileiro de Recursos Hídricos**. 2013. p. 1-8.

MENDONÇA, N. S. **Sustentabilidade e usos múltiplos da água da barragem Poção da Ribeira-SE**. 2013. 225 f. Dissertação (Mestrado em Desenvolvimento e Meio Ambiente) – Universidade Federal de Sergipe, Sergipe, 2013.

ROSSETTI, E. K. **Gestão ambiental: ecoeficiência a caminho da sustentabilidade: o caso do Segmento Moveleiro da Serra Gaúcha**. 2008. 120 f. Dissertação (Mestrado em Administração) – Universidade de Caxias do Sul, Caxias do Sul, 2008.

SANTOS, R.P.; FERREIRA, W.V.F.; PEREIRA, P.N.; MARTINS, J.R.S; FADIGA JR, F.M.; SANTOS, R.C.P. Modelação da propagação de cheias ocasionadas por rompimento de barragens na cadeia de geração do rio Paranapanema. In: II CITENEL, Porto Alegre (RS), 2003. Anais do **II Congresso de Inovação Tecnológica em Energia Elétrica**. 2003. p. 1-4.

SARTORI, R. Z. **Avaliação comparativa de modelos hidrodinâmicos para previsão de inundações: um estudo de caso do município de Getúlio Vargas-RS**. 2018. 148 f. Dissertação (Mestrado em Ciência e Tecnologia Ambiental) – Universidade Federal da Fronteira do Sul, Erechim, Brasil, 2018.

SOUSA, F. M. L.; NETO, V. S. C.; PACHECO, W. E.; BARBOSA, S. A. Sistema nacional de informações sobre recursos hídricos: Sistematização conceitual e modelagem funcional. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, Brasília (DF), 2012. Anais do **XVIII Simpósio Brasileiro de Recursos Hídricos**. Bento Gonçalves: Associação Brasileira de Recursos Hídricos, 2012. p. 1-20.

TSCHIEDEL, A. F. **Abordagem de grande escala para simulações de cheias geradas por rompimentos de barragens de armazenamento de água**. 2022. 255 f. Tese (Doutorado em Recursos Hídricos e Saneamento Ambiental) – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, 2022.

Análise do Transporte de Contaminantes em um Rio Utilizando Problemas Inversos

Anna Luisa Soares Castro Coimbra
Universidade Federal Fluminense (UFF)
anna_luisa@id.uff.br

Fábio Freitas Ferreira
Universidade Federal Fluminense (UFF)
fabiofreitasferreira@id.uff.br

Wagner Rambaldi Telles
Universidade Federal Fluminense (UFF)
wtelles@id.uff.br

Resumo: É notável a existência de algum tipo de poluição na maioria dos rios no Brasil e, diante de situações como essa, o controle do comportamento de contaminantes nesses corpos hídricos é extremamente útil para processos de tomada de decisão, entre elas, determinar o melhor local para descarte de resíduos domésticos e industriais, evitando, assim, uma contaminação do meio ambiente e contribuindo para um melhor monitoramento da qualidade da água. Neste trabalho é analisado um modelo computacional para simular o transporte de contaminantes lançados em um trecho do rio São Pedro, localizado no município de Nova Friburgo, estado do Rio de Janeiro, visando estimar os valores ótimos para a velocidade média e o coeficiente de dispersão utilizando um método de otimização determinístico, mais especificamente, o método Levenberg-Marquardt.

Palavras-chave: Transporte de Contaminantes. Modelagem Computacional. Problemas Inversos. Método de Levenberg-Marquardt.

Introdução

A água é um recurso natural essencial para a sobrevivência humana, desde criança aprendemos que nosso corpo é 70% água, e que precisamos nos manter hidratados e beber água. No entanto, por lidar diariamente com a água em abundância presente nos mares, rios, lagos e represas, a sociedade continua banalizando sua importância e usando-a como um bem descartável.

Entretanto, nem toda essa água presente no mundo é um recurso hídrico. Conforme a Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA, 2022), recurso hídrico é toda água proveniente da superfície ou subsuperfície da Terra, que pode ser empregada em um determinado uso ou atividade, podendo também passar a ser um bem econômico. Todo recurso hídrico é água, mas nem toda água é recurso hídrico. Logo, a água precisa possuir alguma finalidade de uso para a sociedade.

Assim, a poluição hídrica é causada por atividades humanas que geram algum impacto sobre esse recurso natural em qualidade e quantidade. A poluição se dá devido à alteração das características biológicas, físicas, químicas e sedimentar.

No Brasil, a Constituição da República Federativa do Brasil de 1988, documento oficial de maior importância nos países, prevê em seu art. nº 225 que todos têm direito ao meio ambiente ecologicamente equilibrado, bem de uso comum do povo e essencial à sadia qualidade de vida, impondo-se ao Poder Público e à coletividade o dever de defendê-lo e preservá-lo para as presentes e futuras gerações. Contudo, segundo o relatório da Fundação SOS Mata Atlântica divulgado em 2022, apenas 6,9% dos rios observados estão com qualidade da água em boa condição; 72,6% apresentaram qualidade da água regular; 17,8%, ruim; e 2,7%, péssima. O levantamento foi realizado em 146 pontos de coletas de 90 rios e corpos d'água de 65 municípios, em 16 estados do bioma Mata Atlântica, por 106 grupos de monitoramento da qualidade da água do Programa Observando os Rios (SOSMA, 2022).

Portanto, torna-se evidente a necessidade de ferramentas para auxiliar no monitoramento da qualidade das águas, uma vez que os órgãos gestores do meio ambiente esbarram em dificuldades logísticas, financeiras e operacionais: coletas e análises de água são caras, além de requererem material e pessoal especializado (Telles, 2009).

A relevância do presente trabalho se dá devido à grande mobilização mundial para um melhor controle da qualidade de água, em especial a água enquanto recurso hídrico presente nos rios, lagos, lençóis freáticos superficiais e atmosfera, a qual vem sendo atingida cada vez mais pelos crescentes índices de poluição. Diante deste cenário, é importante buscar subsídios para monitorar a qualidade da água de cursos d'água naturais.

Nesse sentido, a modelagem matemática aplicada ao problema de transporte de contaminantes em rios se mostra uma opção mais viável financeira e logisticamente ao desenvolver e aprimorar modelos computacionais que auxiliem no controle da poluição dos rios, pois ajudam a prever o comportamento e estimar o nível de danos que podem ser causados pela poluição despejada na água.

Área de Estudo – Rio São Pedro

A área de estudo para a qual se aplicou a simulação analisada neste trabalho está situada na região serrana do Estado do Rio de Janeiro, dentro da bacia hidrográfica do rio Macaé, a qual possui uma área de drenagem de aproximadamente 1.765 km^2 e engloba praticamente toda a área dos limites territoriais do município de Macaé, onde se localiza sua foz.

Em particular, o rio São Pedro, região escolhida para a realização do experimento, está situada em São Pedro da Serra, 7º Distrito de Nova Friburgo (RJ), localização ($2219'17.7''\text{S}$ $4219'38.0''\text{W}$), local que foi afetado por um rápido processo de transformação espacial, social e econômica, com fortes impactos ambientais (LimaVerde, 2005).

A região tem como característica uma pequena profundidade e margens repletas de rochas conforme mostrado na Figura 1. Além de conter trechos sinuosos e grande biodiversidade, tem como principais atividades econômicas a agricultura familiar e o turismo, dado que a área engloba atrativos como as cachoeiras e a mata, possibilitando a prática de trilhas.



Figura 1: Vista parcial do trecho do rio São Pedro.
Fonte: Sousa (2009).

Modelagem e Solução do Problema Proposto

O modelo matemático para o problema proposto neste trabalho leva em consideração o fato de o trecho do rio estudado ser suficientemente estreito e raso, fazendo com que as variações nas direções

transversal e vertical sejam desprezíveis em relação às variações longitudinais. Logo, um modelo unidimensional que considera apenas as variações na referida direção, sendo expresso pela Equação de Advecção-Dispersão, Equação (1), que descreve o comportamento de um constituinte em relação ao seu deslocamento e sua dispersão ao longo do rio, com condições inicial e de contorno, descritas pelas Equações (2) e (3), respectivamente:

$$\frac{\partial C}{\partial t} + U \frac{\partial C}{\partial x} = E_L \frac{\partial^2 C}{\partial x^2}, \quad -\infty < x < \infty, \quad t > 0 \quad (1)$$

$$C(x, 0) = C_0 + \frac{M}{A} \delta(x), \quad -\infty < x < \infty \quad (2)$$

$$C(\pm\infty, t) = C_0, \quad t > 0 \quad (3)$$

em que C é a concentração (mg/L), C_0 é a concentração no domínio, exceto no ponto de lançamento do poluente (mg/L), t é a variável temporal (s), x é a variável espacial (m), U é a velocidade no sentido longitudinal do escoamento (m/s), E_L é o coeficiente de dispersão longitudinal (m^2/s), M é a massa do constituinte (mg), A é a área da seção transversal do rio (m^2) e $\delta(x)$ é a função *Delta de Dirac*.

A solução analítica para o modelo matemático descrito na Equação (1), com condições de contorno e inicial, dadas pelas Equações (2) e (3), é descrita na Equação (4) (Sousa, 2009):

$$C(x, t) = C_0 + \frac{M}{A\sqrt{4\pi E_L t}} \exp\left(-\frac{(x - Ut)^2}{4E_L t}\right), \quad -\infty < x < \infty, \quad t > 0 \quad (4)$$

A estimativa dos valores ótimos para a velocidade (U) e o coeficiente de dispersão longitudinal (E_L), de maneira que permita o melhor ajuste entre os dados experimentais e os obtidos analiticamente, Equação (4), é feita utilizando o método Levenberg-Marquardt, cujos resultados são apresentados na próxima seção.

Resultados e Discussões

Os resultados apresentados foram obtidos usando o método Levenberg-Marquardt, em que o mesmo foi executado 100 vezes com o intuito de estimar e analisar a velocidade (U) e o coeficiente de dispersão (E_L) presentes no problema. Em cada uma dessas execuções, foi calculado o número de avaliações da função objetivo ($S(\vec{Z})$), imprescindível para que o algoritmo encontre os parâmetros de interesse, tomando como critério de parada o erro relativo entre duas iterações consecutivas expresso por $\max|z_i^{k+1} - z_i^k| < 10^{-1}$. Para as estimativas iniciais $\vec{Z}^0$ referentes a cada uma das 100 execuções, utilizou-se um valor tomado aleatoriamente nos intervalos $0, 1 \leq U \leq 1,0$ e $1,0 \leq E_L \leq 2,0$. Os resultados foram gerados utilizando o software scilab em um computador Intel Core i5 com 8 GB de memória RAM,

Ademais, para uma análise dos resultados obtidos pelo método Levenberg-Marquardt, foram calculados indicadores estatísticos, levando em conta o total de execuções, sendo esses: melhor resultado (menor $S(\vec{Z})$), pior resultado (maior $S(\vec{Z})$), média aritmética simples e desvio padrão da velocidade (U) e do coeficiente de dispersão longitudinal (E_L), assim como para o valor e número de avaliações da função objetivo ($S(\vec{Z})$).

Estimativa da Velocidade (U)

A seguir, são salientados os resultados para a estimativa da velocidade (U) usando o método Levenberg-Marquardt. Na Tabela 1, é apresentado o resumo das informações estatísticas para o parâmetro de interesse e, na Figura 2, é mostrado o gráfico das curvas obtidas computacionalmente para o perfil de

concentração ao longo do tempo, levando em consideração o melhor, pior e média dos valores obtidos para U após as 100 execuções do algoritmo. O coeficiente de dispersão longitudinal (E_L) foi mantido fixo e igual a $1,82 \text{ m}^2/\text{s}$, da mesma maneira que os demais parâmetros citados nessa seção.

Tabela 1: Resultados obtidos para a velocidade (U).

Parâmetros	$U(\text{m/s})$	$S(\bar{Z})$
Melhor	0,590224	323,587388
Pior	0,223002	5425,668459
Médio	0,516569	1231,910495
Desvio Padrão	0,152267	1941,398293

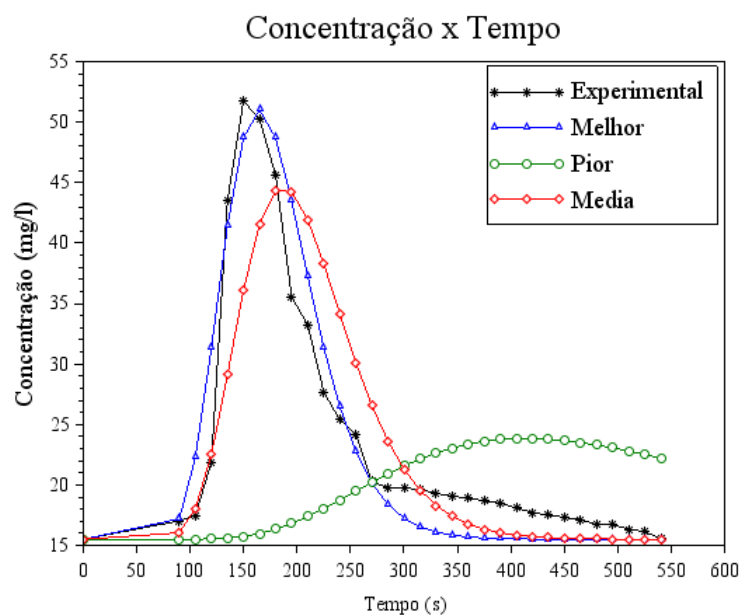


Figura 2: Perfis das concentrações obtidas com a estimativa do parâmetro U pelo método Levenberg-Marquardt no melhor caso, pior caso, média e dados experimentais.

O baixo desvio padrão (0,152267) para o valor da velocidade (U), encontrado ao longo das 100 execuções, evidencia a boa precisão e desempenho satisfatório do método de otimização Levenberg-Marquardt para esta série de problemas.

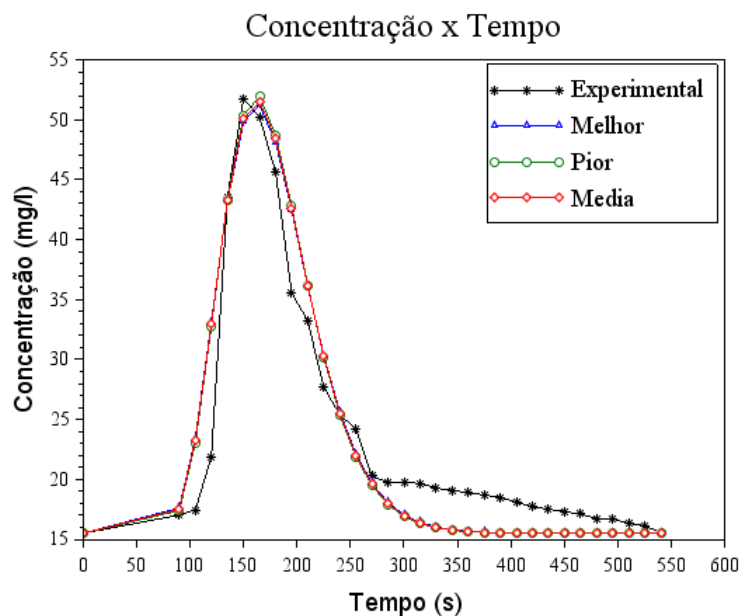
Estimativa do Coeficiente de Dispersão Longitudinal (E_L)

Na sequência, são evidenciados os resultados para a estimativa do coeficiente de dispersão longitudinal (E_L). Assim, na Tabela 2 são mostrados os resultados obtidos pelo método Levenberg-Marquardt, tal como o gráfico, Figura 3, contendo as curvas para o perfil da concentração ao longo do tempo feitas com o melhor, pior e média dos valores obtidos após as 100 execuções do algoritmo. A velocidade (U) permaneceu fixa e igual a $0,60 \text{ m/s}$, do mesmo modo que os demais parâmetros abordados nesta seção.

Novamente, é observado um baixo desvio padrão (0,023844), achado no decorrer das 100 execuções, para o valor do coeficiente de dispersão Longitudinal (E_L), o que evidencia o bom desempenho do método de otimização Levenberg-Marquardt para este tipo de problemática.

Tabela 2: Resultados obtidos para coeficiente de dispersão longitudinal (E_L).

Parâmetros	$E_L(m^2/s)$	$S(\vec{Z})$
Melhor	1,879151	336,417366
Pior	1,804638	337,893279
Médio	1,846805	336,839099
Desvio Padrão	0.023844	0,404773


 Figura 3: Perfis das concentrações obtidas com a estimativa do parâmetro E_L pelo método Levenberg-Marquardt no melhor caso, pior caso, média e dados experimentais.

Estimativa da Velocidade (U) e do Coeficiente de Dispersão Longitudinal (E_L)

Nessa seção são mostrados os resultados obtidos pelo método Levenberg-Marquardt para a estimativa da velocidade (U) e do coeficiente de dispersão longitudinal (E_L) simultaneamente. As informações estatísticas envolvendo as estimativas dos parâmetros de interesse são descritas na Tabela 3, já o gráfico, Figura 4, exibe as curvas para o perfil da concentração ao longo do tempo feitas com o melhor, pior e média dos valores obtidos após as 100 execuções do algoritmo. Os outros parâmetros abordados no início desta seção foram mantidos constantes.

 Tabela 3: Resultados obtidos para a estimativa da velocidade (U) e do coeficiente de dispersão longitudinal (E_L).

Parâmetros	$U(m/s)$	$E_L(m^2/s)$	$S(\vec{Z})$
Melhor	0,589341	1,876560	322,994827
Pior	0,988958	2,000000	14232,904809
Médio	0,593822	1,930021	1726,264413
Desvio Padrão	0,153576	0,071352	3497,073157

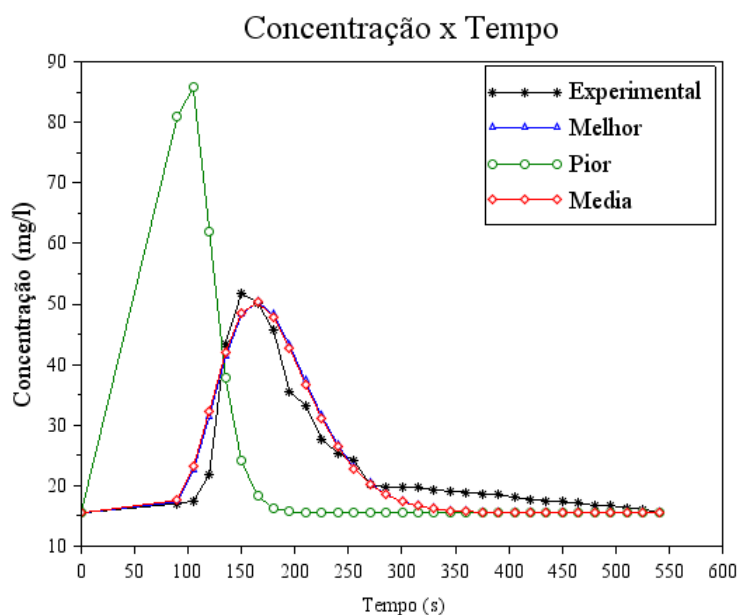


Figura 4: Perfis das concentrações obtidas com a estimativa dos parâmetros U e E_L pelo método Levenberg-Marquardt no melhor caso, pior caso, média e dados experimentais.

Apesar de ser constatado um baixo desvio padrão (0,153576 e 0,071352), ao longo das 100 execuções, para as estimativa da velocidade (U) e do coeficiente de dispersão longitudinal (E_L), o que enfatiza, outra vez, o bom desempenho do método de otimização Levenberg-Marquardt para este tipo de problemática.

Conclusões

O presente trabalho buscou simular e investigar o transporte e comportamento de poluentes em corpos hídricos e sua consequente modelagem, utilizando conceitos matemáticos. Em seguida, para tal fim, buscou-se analisar a influência da variação do coeficiente de dispersão longitudinal e de velocidade no do rio São Pedro, afluente do rio Macaé, que está situado em São Pedro da Serra, 7º Distrito de Nova Friburgo, estado do Rio de Janeiro.

Por fim, foi realizada a estimativa dos parâmetros U e E_L usando o método de otimização irrestrita Levenberg-Marquardt. Onde, tanto de forma individual, quanto simultânea, pode-se concluir que, ao utilizar o método de otimização Levenberg-Marquardt com o objetivo de analisar o comportamento de uma pluma de contaminante em corpos hídricos ao longo do tempo, a partir de um ponto de interesse, foi possível representar o perfil numérico da concentração ao longo do tempo de maneira próxima aos dados experimentais, demonstrando um bom desempenho do referido método.

Agradecimentos

Ao programa Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia, ofertado pela Escola de Engenharia Industrial Metalúrgica de Volta Redonda (EEMIMVR/UFRJ) em parceria com o Instituto do Noroeste Fluminense de Ensino Superior (INFES), campus regional da Universidade Federal Fluminense (UFF), localizado em Santo Antônio de Pádua, estado do Rio de Janeiro.

Esse estudo foi apoiado pela Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) (E-47/2021-SEI260003/016517/2021-E-26/210.107/2022).

Referências

CAMPBELL, M. G. C. **Estimativa de parâmetros em corpos hídricos utilizando modelagem matemática e o método Levenberg-Marquardt**. 2019. Trabalho de Conclusão de Curso (Licenciatura em Matemática) – Instituto do Noroeste Fluminense de Educação Superior, Universidade Federal Fluminense, Santo Antônio de Pádua, RJ, 2019.

Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). **Manejo de recursos hídricos**. Educação e ensino de ciências e matemática: Pesquisa, Aplicação e Novas Tendências, v. 2, n. 1, 2022. Disponível em: <<https://www.embrapa.br/tema-manejo-de-recursos-hidricos/perguntas-e-respostas#:~:text=Recurso%20h%C3%ADdrico%2C%20por%20sua%20vez,toda%20%C3%A1gua%20%C3%A9%20recurso%20h%C3%ADdrico.>>. Acesso em: 17 de Fev. de 2023.

PLANALTO. **Constituição da República Federativa do Brasil de 1988**. 1988. Disponível em: <https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 17 de Fev. de 2023.

SOS Mata Atlântica (SOSMA). **Observando os rios - o retrato da qualidade da água nas bacias hidrográficas da Mata Atlântica**. 2022. Disponível em: <https://cms.sosma.org.br/wp-content/uploads/2022/03/SOSMA_Observando-os-Rios_2022.pdf>. Acesso em: 17 de Fev. de 2023.

SOUSA, E. P. **Avaliação de mecanismos dispersivos em rios através de problemas inversos**. 2009. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2009.

TELLES, W. R. **Simulação do transporte horizontal bidimensional de substância conservativa**. 2009. Dissertação (Mestrado em Modelagem Computacional) – Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2009.

MÉTODO ASSINTÓTICO APLICADO AO MODELO DE TRÁFEGO DE GREENBERG

Mariana Silva

*Universidade Federal Fluminense
silvamariana@id.uff.br*

Panthers Rodriguez Bermudez

*Universidade Federal Fluminense
panthersrb@id.uff.br*

Resumo: Neste trabalho foi aplicado o método assintótico no modelo de tráfego de Greenberg, que é representado por uma lei de conservação, para obter soluções do tipo choque. O texto inicia apresentando a função polinomial que é uma aproximação da função de fluxo do modelo de Greenberg, a partir dela se constrói a cadeia de Hugoniot-Maslov, que é o método assintótico. A base matemática para construção do próprio é a subálgebra de Colombeau. Tal cadeia é um sistema de equações diferenciais ordinárias. Esse sistema foi resolvido numericamente pelo método numérico de Runge-Kutta, com isso se obteve a solução numérica pelo método assintótico. Por fim, foi feita uma comparação com dois métodos de diferenças finitas, para mostrar a eficácia do método assintótico.

Palavras-chave: Método Assintótico. Modelo de Greenberg. Cadeia de Hugoniot-Maslov. Álgebra de Colombeau. Runge-Kutta.

Introdução

As equações de leis de conservação possuem alguns tipos de soluções, uma delas são as soluções do tipo choque, que é a que foi focada nesse artigo. Soluções do tipo choque são descontinuidades, e um dos primeiros obstáculos foi obter a mesma com significado físico, assim surgiu as condições de entropia para resolver tal problema, isso claro para condições iniciais do tipo Riemann. Já, numericamente há o problema de achar um método que represente o choque de forma semelhante a analítica, já que os métodos numéricos utilizados sofrem algum tipo de desvio (podendo ser uma difusão, oscilação, etc) perto da descontinuidade. A partir disso criou-se o interesse de criar um outro método, que é o método assintótico. Neste artigo, a base matemática para construir o próprio foi a subálgebra de Colombeau.

Neste trabalho foi aplicado o método assintótico no modelo de tráfego de Greenberg. Tal modelo foi criado em 1959 a partir de experimentos feitos nos túneis e pontes de Nova York, modelando assim uma faixa única sem entradas e nem saídas laterais. Ele é do tipo macroscópico, ou seja, nesse modelo não é considerado o movimento de um só veículo, e sim utiliza-se de variáveis, como a densidade, velocidade e fluxo, para modelar o comportamento do tráfego. Modelos desse tipo descrevem o tráfego como um fluxo contínuo, assim consegue-se deduzir uma equação, que é a lei de conservação, que é dada por

$$\rho_t + \left[v_{\max} \ln \left(\frac{\rho_{\max}}{\rho} \right) \rho \right]_x = 0. \quad (1)$$

onde ρ é a densidade, $\rho_{\max}$ é a densidade máxima, $0 < \rho \leq \rho_{\max}$, $v_{\max}$ é a velocidade máxima obtida através dos experimentos e a função velocidade dependendo de ρ é $v(\rho) = v_{\max} \ln \left(\frac{\rho_{\max}}{\rho} \right)$.

Esse modelo tem um melhor comportamento para situações do tipo congestionamento (altas densidades), pois para baixas densidades, ou seja, $\rho \rightarrow 0$, com isso $v(\rho)$ se aproxima de infinito, o que não é permitido fisicamente. Fazendo $u = \rho / \rho_{\max}$, obtém-se que $0 < u \leq 1$ e $v(u) = -v_{\max} \ln(u)$.

Portanto, a equação de lei de conservação para o modelo de Greenberg é

$$u_t - v_{\max} (u \ln(u))_x = 0, \quad 0 < u \leq 1. \quad (2)$$

A função de fluxo da Equação (2) é dada por

$$f(u) = -v_{\max} u \ln(u). \quad (3)$$

Para o que se interessa em fazer, precisa-se que a função de fluxo seja um polinômio, o que não ocorre na Equação (3). Para isso aproxima-se a função de fluxo da Equação (3) por uma função polinomial, cujo teorema a seguir garante isso.

Teorema 1: Considere a função $f(u) = -v_{\max} u \ln(u)$, que representa o fluxo do modelo de Greenberg. A expansão em série de Taylor no ponto $u_0 = 1/2$ é dada por

$$f(u) = \frac{v_{\max}}{2} + v_{\max} u (\ln(2) - 1) + \sum_{k=2}^{\infty} \frac{(-1)^{k+1} v_{\max} 2^{(k-1)}}{k(k-1)} (u - 1/2)^k. \quad (4)$$

Pode-se verificar as seguintes propriedades

- (i) $f(u)$ é analítica no intervalo $0 \leq u \leq 1$, ou seja, a série de Taylor na Equação (4) converge nesse intervalo.
- (ii) o polinômio de grau N obtido do truncamento da série de Taylor na Equação (4) tem a mesma concavidade (côncava) da função de fluxo de Greenberg no intervalo $(0, 1)$, qualquer que seja o grau de truncamento $N \geq 2$.

Esse teorema e sua prova se encontram em Silva (2022).

Por fim, foi feita simulações obtendo o resultado numérico utilizando o método assintótico e foi feito também comparações com dois métodos numéricos de diferenças finitas para mostrar a eficácia do método assintótico.

Método Assintótico

A ideia do método assintótico, que foi criado pelo Maslov (1972), é escrever a solução da equação de lei de conservação como um somatório de funções generalizadas da subálgebra de Colombeau, e tais funções vêm da resolução do sistema infinito de equações diferenciais ordinárias acopladas, que é conhecido como cadeia de Hugoniot-Maslov. E para resolver esse sistema de equações diferenciais ordinárias acopladas utiliza-se o método numérico de Runge-Kutta para sistemas, que é visto na seção seguinte. Consequentemente resolvendo o sistema, resolve-se a lei de conservação.

Considerando a Equação (4) como a equação de fluxo do modelo de Greenberg se faz $\mu = u - 1/2$, lembrando que $0 \leq u \leq 1$, assim $-1/2 \leq \mu \leq 1/2$. Com isso $\mu_x = u_x$, então

$$f(\mu) = \frac{v_{\max}}{2} + v_{\max} (\mu + 1/2) (\ln(2) - 1) + \sum_{k=2}^{\infty} \frac{(-1)^{k+1} v_{\max} 2^{k-1}}{k(k-1)} \mu^k. \quad (5)$$

Tal substituição é feita para facilitar a obtenção da cadeia de Hugoniot-Maslov.

Para obter soluções do tipo choque pelo método assintótico, tem-se que a solução tem que ser da seguinte forma

$$\mu(x, t) = A(x, t) + B(x, t) H(x - X(t)) \quad (6)$$

onde H é uma função generalizada de Heaviside em $G_S(\mathbb{R})$, A , B e X são funções C^∞ de todos os seus argumentos, e a curva $x - X(t) = 0$ descreve a trajetória da singularidade. Supondo que a solução da Equação (5) e a Equação (6), tem-se então que a solução da Equação (4) é dada por $u(x, t) = A(x, t) + B(x, t) H(x - X(t)) + \frac{1}{2}$.

Com Equação (6), obtém-se que

$$\begin{aligned} f(\mu) &= \frac{v_{max}}{2} + v_{max}[A(x, t) + B(x, t)H(x - X(t)) + 1/2](\ln(2) - 1) \\ &+ \sum_{k=2}^{\infty} \frac{(-1)^{k+1} v_{max} 2^{k-1}}{k(k-1)} [A(x, t) + B(x, t)H(x - X(t))]^k. \end{aligned} \quad (7)$$

Assim,

$$\begin{aligned} \mu_t + (f(\mu))_x &= A_t + B_t H - BX' \delta(\tau) + v_{max}[A_x + B_x H(\tau) + B \delta(\tau)](\ln(2) - 1) \\ &+ \sum_{k=2}^{\infty} \frac{(-1)^{k+1} v_{max} 2^{k-1}}{k-1} [A + BH(\tau)]^{k-1} [A_x + B_x H(\tau) + B \delta(\tau)]. \end{aligned} \quad (8)$$

Portanto,

$$\begin{aligned} A_t + B_t H - BX' \delta(\tau) + v_{max}[A_x + B_x H(\tau) + B \delta(\tau)](\ln(2) - 1) \\ + \sum_{k=2}^{\infty} \frac{(-1)^{k+1} v_{max} 2^{k-1}}{k-1} [A + BH(\tau)]^{k-1} [A_x + B_x H(\tau) + B \delta(\tau)] \approx 0. \end{aligned} \quad (9)$$

Logo,

$$\mu(t, x) = \sum_{k=0}^{\infty} A_k(t)(x - X(t))^k + \left(\sum_{k=0}^{\infty} B_k(t)(x - X(t))^k \right) H(x - X(t)) \quad e \quad (10)$$

$$u(t, x) = \sum_{k=0}^{\infty} A_k(t)(x - X(t))^k + \left(\sum_{k=0}^{\infty} B_k(t)(x - X(t))^k \right) H(x - X(t)) + \frac{1}{2}. \quad (11)$$

Para obter a cadeia de Hugoniot-Maslov, foi utilizado o seguinte teorema, que se encontra em Rodriguez-Bermudez et al. (2007).

Teorema 2: Se existe uma solução fraca da equação $u_t + f'(u)u_x = 0$ na álgebra $G_S(\mathbb{R}^2)$ da forma da Equação (10), então as funções $X, A_l, B_l, l = 0, 1, 2, \dots$, satisfazem o seguinte sistema infinito de EDOs:

$$X' = \sum_{k=0}^n \frac{1}{k+1} \sum_{i=k}^n a_i f_i(k) A_0^{i-k} B_0^k, \quad (12)$$

$$A'_l = (l+1)A_{l+1}(X' - a_0) - \sum_{j=0}^l \left[(j+1)a_1 A_{l-j} A_{j+1} + \sum_{i=2}^n a_i \tilde{A}_{i,j,l} (j+1) A_{j+1} \right], \quad (13)$$

$$\begin{aligned} B'_l &= (l+1)B_{l+1}(X' - a_0) - \sum_{j=0}^l \left\{ (j+1)a_1 (A_{l-j} B_{j+1} + B_{l-j} B_{j+1} + B_{l-j} A_{j+1}) + 2a_2 (j+1) [A_{j+1} \right. \\ &+ B_{j+1}] \sum_{m=0}^{l-j} A_{l-j-m} B_m + \sum_{i=2}^n [(j+1)a_i (B_{j+1} \tilde{A}_{i,j,l} + A_{j+1} \tilde{B}_{i,j,l} + B_{j+1} \tilde{B}_{i,j,l})] + \sum_{i=2}^{n-1} [(i+ \\ &+ 1)a_{i+1} A_{l-j} \sum_{m=0}^j (m+1)(A_{m+1} B_{m+1}) \tilde{B}_{i,m,j}] + \left(\sum_{i=3}^n i a_i \tilde{A}_{i-1,j,l} \right) \left(\sum_{m=0}^j (m+1) B_{j-m} [A_{m+1} + B_{m+1}] \right) \\ &\left. + \sum_{k=2}^n \left(\sum_{i=k+2}^n a_i f_i(k) \tilde{A}_{i-k,j,l} \right) \left(\sum_{m=0}^j (m+1) \tilde{B}_{k,m,j} (A_{m+1} + B_{m+1}) \right) \right\} \end{aligned} \quad (14)$$

O sistema infinito de EDOs escrito desta forma não tem solução. Para conseguir resolver esse sistema foi feito um truncamento, ou seja, fixa-se l , no caso $l = 2$, lembrando que l começa no 0. Com isso, foi obtido um sistema com sete EDOs para ambos os modelos, que já é o suficiente para adquirir bons resultados. E admiti-se que $A_i = B_i = 0$, para quaisquer A_i e B_i que aparecerem com $i > 2$.

Usando o Teorema (2), observando que $a_0 = v_{\max}(\ln(2) - 1)$, $a_i = [(-1)^i v_{\max} 2^i]/i$, para todo $i = 1, 2, 3, \dots$ e fazendo o truncamento, obtém-se o seguinte sistema

$$\left\{ \begin{array}{l} X' = v_{\max}(\ln(2) - 1) + \sum_{i=1}^n a_i A_0^i + \sum_{k=1}^n \frac{1}{k+1} \sum_{i=k}^n a_i f_i(k) A_0^{i-k} B_0^k \\ A'_0 = A_1[X' - v_{\max}(\ln(2) - 1)] + 2v_{\max} A_0 A_1 - \sum_{i=2}^n a_i \tilde{A}_{i,0,0} A_1 \\ A'_1 = 2A_2[X' - v_{\max}(\ln(2) - 1)] + 2v_{\max} A_1^2 - \sum_{i=2}^n a_i \tilde{A}_{i,0,1} A_1 + 4v_{\max} A_0 A_2 - \sum_{i=2}^n a_i \tilde{A}_{i,1,1} 2A_2 \\ A'_2 = 6v_{\max} A_1 A_2 - \sum_{i=2}^n a_i \tilde{A}_{i,0,2} A_1 - \sum_{i=2}^n a_i \tilde{A}_{i,1,2} 2A_2 \\ B'_0 = B_1[X' - v_{\max}(\ln(2) - 1)] - \{(-2v_{\max})(A_0 B_1 + B_0 B_1 + B_0 A_1) + 4v_{\max}[A_1 + B_1]A_0 B_0 \\ + \sum_{i=2}^n [a_i(B_1 \tilde{A}_{i,0,0} + A_1 \tilde{B}_{i,0,0} + B_1 \tilde{B}_{i,0,0})] + \sum_{i=2}^{n-1} [(i+1)a_{i+1}A_0(A_1 + B_1)\tilde{B}_{i,0,0}] \\ + \left(\sum_{i=3}^n i a_i \tilde{A}_{i-1,0,0}\right)(B_0[A_1 + B_1]) + \sum_{k=2}^n \left(\sum_{i=k+2}^n a_i f_i(k) \tilde{A}_{i-k,0,0}\right)(\tilde{B}_{k,0,0}(A_1 + B_1))\} \\ B'_1 = 2B_2[X' - v_{\max}(\ln(2) - 1)] - \sum_{j=0}^1 \{(j+1)(-2v_{\max})(A_{1-j}B_{j+1} + B_{1-j}B_{j+1} + B_{1-j}A_{j+1}) \\ + 4v_{\max}(j+1)[A_{j+1} + B_{j+1}] \sum_{m=0}^{1-j} A_{1-j-m}B_m + \sum_{i=2}^n [(j+1)a_i(B_{j+1}\tilde{A}_{i,j,1} + A_{j+1}\tilde{B}_{i,j,1} + B_{j+1}\tilde{B}_{i,j,1})] \\ + \sum_{i=2}^{n-1} \left[(i+1)a_{i+1}A_{1-j} \sum_{m=0}^j (m+1)(A_{m+1} + B_{m+1})\tilde{B}_{i,m,j}\right] + \left(\sum_{i=3}^n i a_i \tilde{A}_{i-1,j,1}\right) \left(\sum_{m=0}^j (m+1)B_{j-m}[A_{m+1} + B_{m+1}]\right) \\ + \sum_{k=2}^n \left(\sum_{i=k+2}^n a_i f_i(k) \tilde{A}_{i-k,j,1}\right) \left(\sum_{m=0}^j (m+1)\tilde{B}_{k,m,j}(A_{m+1} + B_{m+1})\right)\} \\ B'_2 = - \sum_{j=0}^2 \{(j+1)(-2v_{\max})(A_{2-j}B_{j+1} + B_{2-j}B_{j+1} + B_{2-j}A_{j+1}) \\ + 4v_{\max}(j+1)[A_{j+1} + B_{j+1}] \sum_{m=0}^{2-j} A_{2-j-m}B_m + \sum_{i=2}^n [(j+1)a_i(B_{j+1}\tilde{A}_{i,j,2} + A_{j+1}\tilde{B}_{i,j,2} + B_{j+1}\tilde{B}_{i,j,2})] \\ + \sum_{i=2}^{n-1} \left[(i+1)a_{i+1}A_{2-j} \sum_{m=0}^j (m+1)(A_{m+1} + B_{m+1})\tilde{B}_{i,m,j}\right] + \left(\sum_{i=3}^n i a_i \tilde{A}_{i-1,j,2}\right) \left(\sum_{m=0}^j (m+1)B_{j-m}[A_{m+1} + B_{m+1}]\right) \\ + \sum_{k=2}^n \left(\sum_{i=k+2}^n a_i f_i(k) \tilde{A}_{i-k,j,2}\right) \left(\sum_{m=0}^j (m+1)\tilde{B}_{k,m,j}(A_{m+1} + B_{m+1})\right)\}. \end{array} \right.$$

Para resolver o sistema de EDOs precisa-se de valores para A_i e B_i em $t = 0$, com $i = 0, 1$ ou 2 , tais valores vêm da condição inicial $u_0(x)$. Tem-se que as funções suaves $u_\ell(x)$ e $u_r(x)$ podem ser formalmente expressas como a série de Taylor na forma

$$u_\ell(x) = \sum_{k=0}^{+\infty} \frac{u_\ell^{(k)}(0)}{k!} x^k, \quad u_r(x) = \sum_{k=0}^{+\infty} \frac{u_r^{(k)}(0)}{k!} x^k.$$

Aqui não se está preocupado com o raio de convergência dessas séries. Aplicando na Equação (11) para $t = 0$, tem-se

$$u(0, x) = \left[\sum_{k=0}^{+\infty} \frac{u_\ell^{(k)}(0)}{k!} x^k \right] + \left[\sum_{k=0}^{+\infty} \frac{u_r^{(k)}(0) - u_\ell^{(k)}(0)}{k!} x^k \right] H(x) + \frac{1}{2}. \quad (15)$$

Tomando a equação solução do modelo de Greenberg (11) em $t = 0$, com $X(0) = 0$ e igualando a Equação (15), foi obtido que

$$A_0(0) = \frac{u_\ell^{(0)}(0)}{0!} - \frac{1}{2}, \quad A_k(0) = \frac{u_\ell^{(k)}(0)}{k!}, \quad k = 1, 2. \quad (16)$$

$$B_0(0) = \frac{u_r^{(0)}(0)}{0!} - \frac{1}{2}, \quad B_k(0) = \frac{u_r^{(k)}(0)}{k!}, \quad k = 1, 2. \quad (17)$$

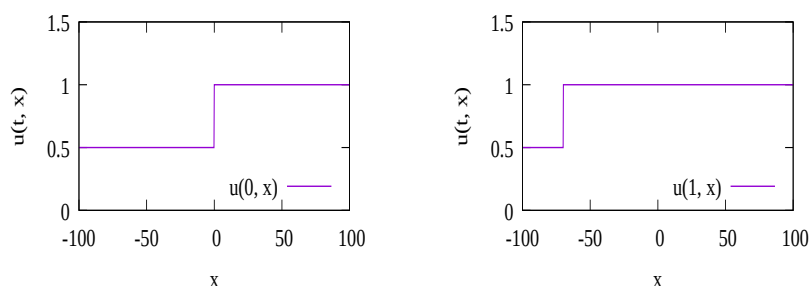
Simulações

Aqui vê-se as soluções numéricas variando as condições iniciais, fazendo do problema de Riemann clássico até condições iniciais não constantes, problema de Riemann generalizado. Tais simulações foram feitas para o seguinte dado de velocidade máxima, $v_{max} = 100 \text{ km/h}$. Como dito anteriormente, foi utilizado o método de Runge-Kutta para sistema para resolver a cadeia de Hugoniot-Maslov. Todos os exemplos foram resolvidos pelo método assintótico, e além disso é feita a comparação com outros métodos numéricos de diferenças finitas, que são os métodos de Lax-Wendroff e Lax-Friedrichs. Foram realizadas 1000 iterações, para o intervalo de tempo de 0,0 à 1,0, e para o intervalo espacial de -100 à 100 , para ambos os métodos numéricos.

Nesta seção pode-se ver que as soluções numéricas do modelo de Greenberg utilizando a função aproximada truncada em $n = 9$, que como foi observado em Silva (2022), já é uma ótima aproximação.

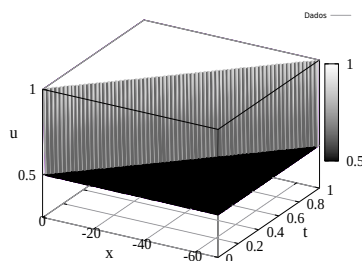
Exemplo 1: Seja a seguinte condição inicial, que é um problema de Riemann clássico,

$$u(0, x) = \begin{cases} 0,5, & \text{se } x < 0 \\ 1,0, & \text{se } x > 0 \end{cases}. \quad (18)$$



(a) Condição inicial do Exemplo 1.

(b) Solução numérica no tempo $t = 1$, Exemplo 1.



(c) Gráfico 3D do Exemplo 1.

Figura 1: Perfis da descontinuidade Exemplo 1.

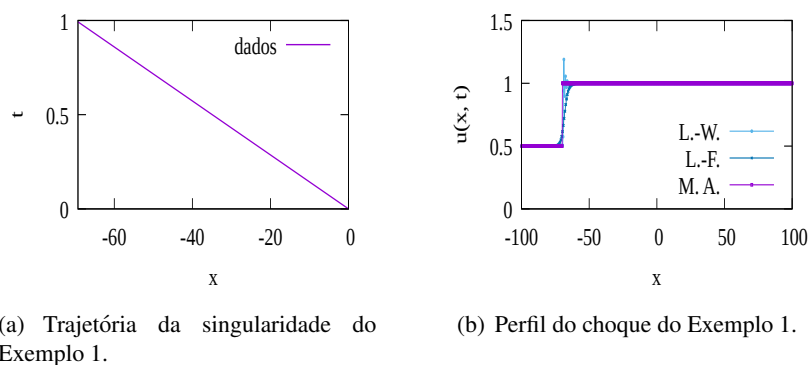


Figura 2: Gráficos do Exemplo 1.

Isto é, respeitando a condição de entropia. Então, se obtém o resultado numérico, via método assintótico, para $t = 1$, onde o choque está localizado em $x \approx -69,315$, o que pode ser observado na Figura (1(b)). Na Figura (1(c)), pode-se observar que a altura do choque e seu comportamento, ambos permanecem iguais ao longo do tempo. A trajetória da singularidade é dada por $x = -69,315t$, que é encontrada pela condição de Rankie-Hugoniot, só que neste caso com a função aproximada sendo a função de fluxo. E é exatamente o que foi encontrado numericamente, como pode ser visto na Figura (2(a)). Para este exemplo, como pode ser visto na Figura (2(b)), o desempenho do método assintótico é o que mais parece com o choque, numericamente, onde L-W é o Lax-Wendroff, L-F é o Lax-Friedrichs e M. A. é o método assintótico.

Exemplo 2: Dada a condição inicial

$$u(0, x) = \begin{cases} 0,25 + 0,001x, & \text{se } x < 0 \\ 0,75 + 0,001x, & \text{se } x > 0 \end{cases} . \quad (19)$$

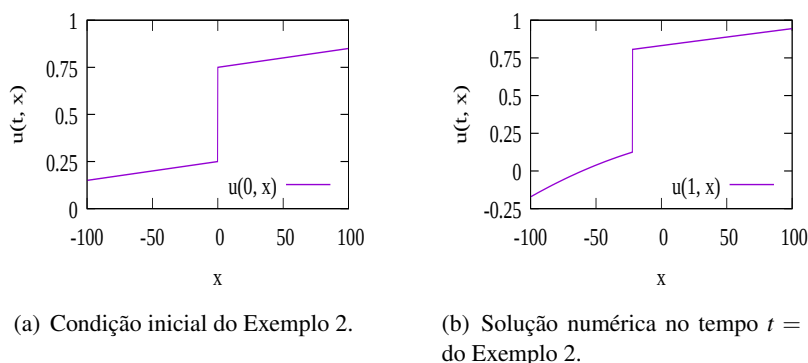


Figura 3: Perfis da descontinuidade do Exemplo 2.

Para $t = 1$ se pode ver o perfil do choque obtido pelo método assintótico na Figura (3(b)), cuja a localidade do choque é em $x \approx -21,988$. No gráfico 3D da Figura (4(a)) pode-se notar que a altura do choque aumenta ao longo do tempo, e além disso nota-se que o comportamento permanece o mesmo. Nesse caso temos que a velocidade da descontinuidade varia ao longo do tempo, Figura (4(b)). Na Figura (5(a)) estamos no tempo $t = 0,5$ e a Figura (5(b)) se encontra em $t = 0,9$, em ambos se tem que o método assintótico é o que se aproxima melhor do comportamento do choque analiticamente.

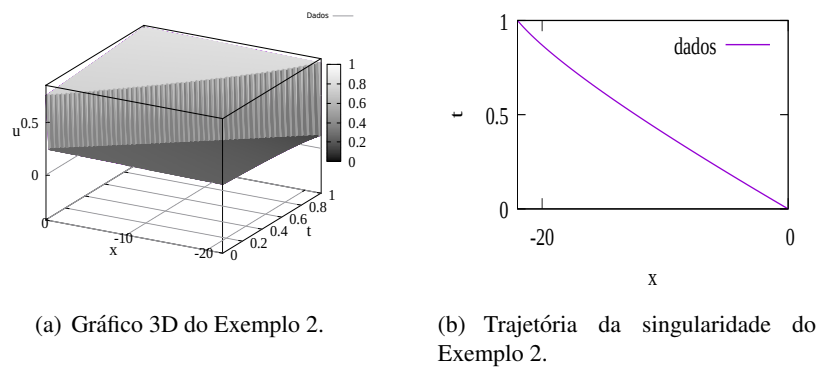


Figura 4: Gráficos do Exemplo 2.

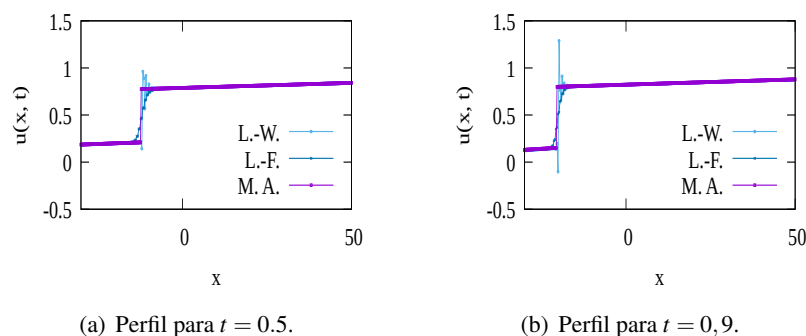


Figura 5: Perfis do choque do Exemplo 2.

Conclusões

Neste trabalho foi aplicado o método assintótico para o modelo de tráfego de Greenberg, e consequentemente obteve-se a cadeia de Hugoniot-Maslov para este modelo, utilizando uma aproximação da função de fluxo de Greenberg. Além disso, mostrou-se que de fato o método assintótico é o que melhor representa a solução do tipo onda de choque.

Referências

- MASLOV, V. P. **Théorie des perturbations et méthodes asymptotiques**. Dunod; 1972.
- RODRIGUEZ-BERMEUDEZ, P.; ALONSO, B. V. **Hugoniot-Maslov chains of a shock wave in conservation law with polynomial flow**. *Mathematische Nachrichten*, v. 280, n.8, p. 907-915, 2007.
- SILVA, M.; RODRÍGUEZ-BERMEUDEZ, P. Aproximação Polinomial da Função de Fluxo do Modelo de Tráfego de Greenberg. In: ANAIS DO ENCONTRO NACIONAL DE MODELAGEM COMPUTACIONAL, ENCONTRO DE CIÊNCIA E TECNOLOGIA DE MATERIAIS, CONFERÊNCIA SUL EM MODELAGEM COMPUTACIONAL E SEMINÁRIO E WORKSHOP EM ENGENHARIA OCEÂNICA, 2022, Brasil. **Anais...** Pelotas(RS) UFPel / FURG / UNIPAMPA, 2022. Disponível em: <https://www.even3.com.br/anais/enmcmscusemengo2022/527147-aproximacao-polinomial-da-funcao-de-fluxo-do-modelo-de-trafego-de-greenberg>. Acesso em: 10/07/2023 16:36
- SILVA, M. **Método Assintótico Aplicado ao Modelo de Tráfego de Greenberg**. 2023. Dissertação de Mestrado– MCCT, UFF, Volta Redonda, 2023.

O USO DA INTELIGÊNCIA ARTIFICIAL PARA AUXILIAR NA PREDIÇÃO DE ABANDONO DE ALUNOS EM UMA ESCOLA DE ENSINO REGULAR

Bianca Santana da Silva

Universidade Federal do Rio de Janeiro - UFRRJ

bia20santanadasilva@gmail.com

Joyce de Limeira Caldeira Rangel

Universidade Federal do Rio de Janeiro - UFRRJ

joylcrangel@ufrj.br

Vanessa Alexandre Silva

Universidade Federal Rural do Rio de Janeiro- UFRRJ

vanessa.mec.ime@gmail.com

Marcelo Dib Cruz

Universidade Federal Rural do Rio de Janeiro- UFRRJ

madibcruz@gmail.com

Ronaldo Malheiros Gregório

Universidade Federal Rural do Rio de Janeiro- UFRRJ

rgregor@ufrj.br

Resumo: A evasão escolar é um desafio que afeta sistemas educacionais em todo o mundo. Existem diversas causas relacionadas a esse fenômeno, como problemas familiares, necessidade de contribuir financeiramente com as despesas de casa, baixa perspectiva, dentre outros. Compreender e prever os fatores que levam os alunos a abandonar os estudos é essencial para desenvolver estratégias eficazes de prevenção e intervenção. A Inteligência Artificial (IA), é um campo da ciência que visa a realização de tarefas de forma similar a de seres humanos. O aprendizado de máquina é um ramo da Inteligência Artificial, onde permite que sistemas computacionais aprendam a partir de dados. O objetivo deste trabalho é criar um modelo de classificação, utilizando as Redes Neurais Artificiais, para fazer a predição de abandono em uma escola do sistema Estadual de Ensino.

Palavras-chave: Evasão Escolar. Inteligência Artificial. Aprendizado de Máquina. Modelo de Classificação. Aprendizado Supervisionado. Redes Neurais Artificiais.

Introdução

A evasão escolar é um desafio que afeta sistemas educacionais em todo o mundo. Existem diversas causas relacionadas a esse fenômeno, dentre várias razões, podem ser destacadas: problemas familiares, necessidade de contribuir financeiramente com as despesas de casa, gravidez na adolescência, baixa perspectiva, dentre outros. Compreender e prever os fatores que levam os alunos a abandonar seus estudos é essencial para desenvolver estratégias eficazes de prevenção e intervenção. (MENDES; SOUZA, 2021).

Para o desenvolvimento do trabalho foi preciso a colaboração dos alunos do Educação de Jovens e Adultos (EJA), que recentemente passou a adotar a sigla EJA-NEM por conta do Novo Ensino Médio. Essa modalidade é ofertada no turno da noite visando o resgate acadêmico de jovens e adultos, que distribui todos os conteúdos do ensino médio ao longo de quatro módulos, o que totaliza o período de dois anos. O EJA-NEM desperta a atenção para esse projeto por ser representado por um grupo de estudantes com trajetórias educacionais não convencionais. Muitos desses alunos, em algum momento de suas vidas, interromperam seus estudos no Ensino Regular. Essa valiosa fonte de dados proporciona

uma oportunidade de explorar as razões subjacentes à evasão escolar e identificar padrões que contribuem para esse fenômeno. Dessa forma, para a realização do estudo, o algoritmo de previsão de abandono foi alimentado com informações coletadas dos alunos do EJAEM.

O trabalho utiliza um conjunto de dados contendo informações sobre esses alunos, incluindo fatores socioeconômicos, histórico da vida escolar, características familiares e outras variáveis relevantes. Ao aproveitar essa riqueza de dados, o objetivo foi desenvolver um modelo capaz de classificar os alunos com base em suas características e prever se estão em risco de evasão. Ao adotar essa abordagem, se torna possível fornecer informações para educadores, formuladores de políticas e profissionais da área de educação. A análise dos resultados obtidos do modelo de previsão de evasão contribuirá para o entendimento mais profundo das circunstâncias que levam ao problema, permitindo a implementação de medidas preventivas mais direcionadas.

Nas seções subsequentes deste artigo, será descrito em detalhes a metodologia utilizada para a coleta e preparação dos dados, a implementação do algoritmo de previsão e a análise dos resultados obtidos. Acredita-se que a aplicação do aprendizado de máquina supervisionado a este contexto específico pode fornecer contribuições valiosas para a compreensão das causas da evasão escolar e, conseqüentemente, para a promoção de um ambiente educacional mais inclusivo e de sucesso.

Objetivo Geral

Desenvolvimento de um modelo computacional utilizando Redes Neurais Artificiais para auxiliar na previsão e prevenção da evasão escolar.

Materiais e Métodos

Para obter dados relevantes para a pesquisa, foi proposto um levantamento dos alunos que evadiram ao longo de um período de 10 anos. No entanto, devido à falta de informações disponíveis sobre esses alunos, tornou-se necessário adotar uma estratégia alternativa. A decisão foi de buscar informações com os alunos do Ensino de Jovens e Adultos do Novo Ensino Médio (EJAEM), pois muitos deles foram anteriormente alunos do Ensino Regular e tiveram que interromper seus estudos por algum motivo. Ao coletar informações com os alunos do EJAEM, foi possível obter uma visão mais completa e aprofundada das causas da evasão escolar. Compreender as circunstâncias específicas que levaram à interrupção dos estudos é fundamental para desenvolver estratégias e políticas eficazes de prevenção da evasão escolar e fornecer suporte adequado aos alunos em situação de risco. É importante destacar que essa abordagem não apenas contribui para a compreensão do problema, mas também permite que os alunos do EJAEM sejam ouvidos e tenham a oportunidade de compartilhar suas experiências e perspectivas. Essa inclusão promove um ambiente de participação e valorização, no qual suas vozes são consideradas na busca por soluções.

Inicialmente, a ideia era disponibilizar esse formulário online por meio de um e-mail do Google, permitindo que os alunos respondessem no momento e local mais convenientes para eles. No entanto, logo surgiram os primeiros contratempos. Foi constatado que muitos alunos enfrentaram dificuldades para compreender as perguntas do formulário ou simplesmente não demonstraram interesse em respondê-lo. Essa abordagem online não atingiu o objetivo esperado de obter um número significativo de respostas completas e precisas.

Surgiu, então, a necessidade de elaborar um formulário impresso para ser distribuído de sala em sala. (MENDES; SOUZA, 2021), (SILVA; SANTOS; PEREIRA, 2019), (CARVALHO, 2014.), (PAKENAS; FILHO, 2017.), (BRANCO; ADRIANO; BRANCO, 2020.), (SILVA; SANTOS; PEREIRA, 2019). O formulário possui os seguintes atributos: *sexo*, *idade*, *renda mensal*, *num moradores (incluindo você)*, *turno*, *serie*, *responsavel por alguem* e *nucleo familiar afeta*. O método presencial permitiu alcançar um maior número de alunos e os formulários foram preenchidos imediatamente.

Com os formulários impressos em mãos, o próximo passo foi cadastrar os dados. Esse processo envolve a inserção das informações coletadas nos sistemas ou bancos de dados adequados, garantindo que os dados estejam organizados e prontos para serem analisados.

Embora a mudança do formato online para o impresso tenha sido necessária devido às dificuldades encontradas, essa nova abordagem permitiu um maior engajamento dos alunos e uma coleta mais eficiente de dados. Com os formulários preenchidos, foi possível realizar uma análise mais aprofundada e exploratória dos dados, identificando padrões, tendências e informações relevantes relacionadas ao problema.

Além disso, também foram realizadas pesquisas semelhantes com alunos do Ensino Regular e Ensino Médio (alunos que não evadiram), utilizando um conjunto de dados correlacionados. Esses alunos são os que estão cumprindo regularmente o Ensino Fundamental e Médio, na mesma escola e nos períodos da manhã ou tarde. O objetivo ao fazer isso, era de estabelecer relações entre os alunos do EJAEM e os alunos que possam estar em situação de risco. Essa abordagem permite identificar semelhanças e padrões entre esses grupos, contribuindo para uma compreensão mais abrangente dos fatores que levam à interrupção dos estudos. Tais pesquisas foram concluídas por meio de formulários com as mesmas perguntas que foram distribuídas aos alunos do EJAEM, dessa forma, passou a ser viável correlacionar os dados dos dois grupos.

Inteligência Artificial

A Inteligência Artificial (IA), é um campo da ciência que visa a realização de tarefas de forma similar a de seres humanos. O aprendizado de máquina é um ramo da Inteligência Artificial, onde permite que sistemas computacionais aprendam a partir de dados. Esse método baseia-se na ideia de que os algoritmos aprendam com o passado para fazer a predição do futuro. Existem vários modelos de aprendizado de máquina, dentre eles, está presente as Redes Neurais, que são modelos inspirados no funcionamento dos neurônios do cérebro humano. Ela consiste em camadas de unidades interconectadas, chamadas neurônios artificiais. Essas técnicas são uma ferramenta poderosa para prever e solucionar vários problemas relacionados ao cotidiano. (RUSSELL; NORVIG, 2013).

O aprendizado de máquina é dividido em várias subáreas incluindo aprendizado por reforço, não supervisionado e supervisionado, o qual será abordado nesse trabalho. Os seguintes procedimentos são necessários para implementar uma rede neural artificial.

A **coleta de dados**, onde é preciso reunir um conjunto relevante de dados para o problema que se deseja resolver. Esses dados podem incluir exemplos, informações ou características relevantes. Outro processo necessário, é o **pré-processamento de dados**, onde os mesmos são preparados antes de alimentar o algoritmo de aprendizado de máquina. Isso inclui a transformação de valores categóricos em numéricos, dessa maneira, foi feita uma tradução dos dados para valores numéricos, possibilitando prosseguir com o trabalho.

A próxima etapa é o **treinamento**, onde os parâmetros internos são ajustados com base nos padrões e relações presentes nos dados. Nessa fase o algoritmo busca minimizar os erros a fim de fazer previsões ou classificações mais precisas.

Após o treinamento é preciso avaliar o desempenho do modelo para saber o quão bem o modelo generaliza para novos dados. Com base nos resultados da avaliação, podem ser feitos alguns ajustes necessários que podem ser desde a troca de parâmetros até a troca de algoritmo caso o desempenho não seja satisfatório.

A categorização de dados é uma tarefa essencial em muitos domínios, desde a medicina até a análise de sentimentos nas redes sociais. Modelos de classificação oferecem uma abordagem sistemática para automatizar esse processo, permitindo a identificação precisa das classes a que os dados pertencem.

Um modelo de classificação é uma representação matemática que aprende a relação entre as características dos dados e as classes a que pertencem. Aprendizado supervisionado é a abordagem mais comum, onde um conjunto de dados de treinamento é usado para ensinar o modelo a fazer previsões. O

objetivo é criar um mapeamento entre as características de entrada e as saídas de classe correspondentes. Os modelos de classificação encontram aplicações em uma ampla gama de domínios, desempenhando um papel significativo em diversas situações. No âmbito educacional, o modelo de classificação pode ajudar a prever a possibilidade de evasão escolar por parte dos alunos, o que possibilita a implementação de intervenções direcionadas para melhorar a retenção e o sucesso educacional. Neste trabalho foi utilizado Redes Neurais Artificiais com o modelo de classificação.

Uma vez que o modelo foi treinado e otimizado adequadamente, ele está apto a ser implantado para tomar decisão e fazer previsões. À medida que novos dados se tornam disponíveis, o modelo pode ser atualizado e retreinado para melhorar o seu desempenho.

Resultados

A amostra utilizada na pesquisa foi de 253 formulários, nos quais 92 são dos alunos que evadiram, e 161 dos alunos do Ensino Regular que **não** evadiram. Desses dados, 70% foram destinados para treino, o que corresponde a 177 formulários, e 30% para teste, correspondendo a 76 formulários. O modelo de **classificação** foi escolhido como objeto de estudo e a implementação foi desenvolvida em Python. (VANDERPLAS, 2016).

Dessa forma, foram utilizadas duas classes: **abandonou** e **não abandonou**. Para executar foram utilizadas bibliotecas poderosas dessa linguagem, como *pandas*, *numpy*, *seaborn* e *sklearn*, sendo as duas últimas essenciais para a geração de gráficos que facilitaram visualizar a correlação entre as colunas do dataset. Tais bibliotecas são amplamente utilizadas para análise, manipulação e visualização dos dados, e foram de suma importância para uma desenvoltura mais aprimorada do projeto.

Foram feitos alguns ajustes, como a transformação categórica em variáveis numéricas. O atributo "serie", em especial, diz respeito à série que o aluno se encontra atualmente para o Ensino Regular, e à série abandonada para os alunos do EJAENEM. Abaixo são mostradas essas transformações:

- **sexo:**

Masculino = 1;

Feminino = 2;

- **idade:**

Número inteiro correspondente.

- **renda_mensal:**

inferior a um salário mínimo = 1;

um salário mínimo = 2;

entre um e dois salários mínimos = 3;

entre dois e três salários mínimos = 4;

entre três e quatro salários mínimos = 5;

acima de quatro salários mínimos = 6;.

- **num_moradores (incluindo você):**

Número inteiro correspondente.

- **turno:**

Manhã = 1;

Tarde = 2;

Noite = 3.

- **serie**

Fundamental;

6º ano = 1;

7º ano = 2;

8º ano = 3;

9º ano = 4;

Ensino Médio;

1º ano = 5;

2º ano = 6;

3º ano = 7;

antes do 6º ano = 8.

- **responsavel_por_alguem:**

sim = 1;

não = 2;

- **nucleo_familiar_afeta:**

sim = 1;

não = 2;

Após o pré processamento dos dados, foi desenvolvida a etapa de tratamento. Nessa etapa, é importante destacar o uso da técnica de preenchimento dos campos faltantes, facilitada pelo método **fillna()**, disponibilizado pela biblioteca pandas do Python. O número negativo (-1) foi utilizado para o preenchimento automático dos dados numéricos faltantes, representando uma nova categoria: “*não sabe/não preencheu*”.

As Redes Neurais Artificiais foram executadas com validação cruzada e a acurácia obtida foi de 0,79. A validação cruzada (*cross-validation*) é uma técnica estatística usada para calcular o desempenho de um modelo de Aprendizado de Máquina. A ideia dessa técnica, é dividir o conjunto dos dados em folders (dobras) e testar com diferentes dobras, estimando sua capacidade de generalização em dados não vistos. Nesse caso, foram divididos em 3 folders e a média obtida foi de 0,797. Isso significa que o algoritmo acerta em aproximadamente 79% dos casos.

Conclusões

É possível concluir que o algoritmo auxilia a prever a possibilidade de ocorrer evasão por parte de algum(a) aluno(a). Nesse projeto, foi preciso lidar com dados insuficientes e obstáculos na fase de coleta de dados. Esses cenários foram revertidos com soluções viáveis à medida que tais imprevistos foram apresentados.

Um marco significativo foi alcançado ao desenvolvermos um modelo de classificação que demonstrou uma acurácia notável de 79% na previsão da evasão escolar. Isso reforça a eficácia desses modelos em fornecer informações para auxiliar instituições educacionais na tomada de decisões. No entanto, ressaltamos a necessidade contínua de refinamento e melhoria do modelo, bem como a consideração de outros indicadores e variáveis relevantes que podem aprimorar ainda mais sua capacidade preditiva.

Referências

- BISHOP, Christopher M. **Neural Networks for Pattern Recognition**. Oxford University Press, 1995.
- BRANCO, E.; ADRIANO, G.; BRANCO, A.; IWASSE, L. **Evasão escolar: desafios para permanência dos estudantes na educação básica**, 2020.
- CARVALHO, M. **Evasão Escolar no Brasil: Uma Perspectiva Multidimensional**, 2014.
- HAYKIN, S. **Neural Networks and Learning Machines**. Prentice Hall, 2008.
- MENDES, L. S.; SOUZA, F. C. (2021). **Análise das Causas da Evasão Escolar no Ensino Fundamental**. Educação em Foco, 25(2), 89-102. DOI: 10.5752/P.2317-788X.2021v25n2p89
- OLIVEIRA, F.; NÓBREGA, L. **Evasão escolar: um problema que se perpetua na educação brasileira**. Revista Educação Pública, 2021.
- PAKENAS, H.; FILHO, J. **Evasão e Abandono no Ensino Médio** - Revista Internacional de Debates de Administração Pública (UNIFESP), 2017.
- RUSSEL, S.; NORVIG, P. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Editora Elsevier, 2013.
- VANDERPLAS, J. **Python Data Science Handbook: Essential Tools for Working with Data**. Editora O'Reilly Media, 2016.
- SILVA, J.; SANTOS, M.; PEREIRA, A. **Evasão Escolar no Ensino Médio Brasileiro: Análise dos Fatores e Possíveis Soluções**. Educação em Revista, v. 25, n. 3, 2019.

SEGMENTAÇÃO DE IMAGENS DE RESSONÂNCIA MAGNÉTICA DE TUMOR CEREBRAL UTILIZANDO MODELO 3D U-NET

Erika Rocha de Araujo
Universidade Federal Fluminense
erikaa@id.uff.br

Vanessa da Silva Garcia
Universidade Federal Fluminense
vanessagarcia@id.uff.br

Wesley Luiz da Silva Assis
Universidade Federal Fluminense
wesleyassis@id.uff.br

Resumo: A partir das Imagens de Ressonância Magnética (IRM) ocorreram avanços significativos na medicina diagnóstica. A utilização do modelo 3D U-Net, tem se destacado na detecção precoce de tumores em IRM, como é o caso dos tumores cerebrais. Esta pesquisa tem por objetivo o desenvolvimento inicial de uma aplicação de um conjunto de imagens de Ressonância Magnética ao modelo 3D U-Net para detecção de tumores cerebrais nas imagens e avaliar seus resultados. Para tanto, utiliza-se IRM no formato NIFTI, no qual os dados são da Competição BraTS2020, um total de 369 grupos de imagens com 5 imagens cada grupo, que são utilizados no treinamento do modelo considerado. As imagens são geradas com o auxílio da linguagem python e a biblioteca NiBabel para realizar a interação com esse formato de dados. As imagens iniciais geradas possibilitam a visualização do tumor realçando a localização tumoral do cérebro. Após o treinamento do modelo 3D U-Net, as previsões geradas possuem relevante correspondência com as máscaras reais.

Palavras-chave: Segmentação de IRM. Modelo 3D U-NET . IRM. Identificação de tumor cerebral

Introdução

A Ressonância Magnética (RM), uma técnica pioneira na medicina diagnóstica, oferece imagens detalhadas dos tecidos moles do corpo humano. Na vanguarda dessa revolução, os avanços na visão computacional e na aprendizagem profunda levaram à criação do modelo 3D U-Net (KRIZHEVSKY, A. et al. 2012), que integra a inteligência artificial a medicina diagnóstica. Por meio dessa tecnologia, é possível acelerar o processo de detecção de tumores e melhorar a precisão das análises, trazendo benefícios significativos aos pacientes.

A arquitetura do 3D U-Net consiste em uma rede neural profunda com uma estrutura de "U". Ela possui uma seção de codificação (encoder) que captura características das imagens em várias escalas e uma seção de decodificação (decoder) que reconstrói a segmentação em alta resolução (ÇIÇEK, et al. 2016). O modelo utiliza convoluções tridimensionais e camadas de pooling tridimensionais para processar volumes de imagem completos.

Este artigo tem por objetivo realizar um estudo inicial para desenvolver uma aplicação de um conjunto de imagens de Ressonância Magnética ao modelo 3D U-Net para detecção mais rápida e precisa de tumores cerebrais nas IRM e avaliar seus resultados. Para tanto, utiliza-se a biblioteca NiBabel para manipular os arquivos de imagens de IRM na linguagem python. Tais imagens estão no formato Neuroimaging Informatics Technology Initiative (NIFTI), que é utilizado para armazenar e compartilhar dados de imagens médicas no campo da neuroimagem. Estes dados são da Competição BraTS2020, um total de 369 grupos de imagens com 5 imagens do mesmo cérebro com imagens de T1, T1ce, T2, Flair e uma última sendo uma imagem do tumor segmentado para o treinamento do modelo.

As imagens são públicas e utilizadas em campeonatos de pesquisas para as equipes participantes desenvolverem seus estudos contribuindo assim com a evolução do conhecimento científico.

Fundamentação Teórica:

- IRM (Imagem de Ressonância Magnética)

A Imagem de Ressonância Magnética é uma técnica de imagem médica não invasiva que utiliza campos magnéticos e ondas de rádio para gerar imagens detalhadas do corpo humano.

Segundo KNOLL, F. et al. (2019), durante a imagem, uma sequência de campos magnéticos variáveis, espacial e temporalmente, chamada "sequência de pulsos", é aplicada pela máquina de RM. Isso induz o corpo a emitir campos de resposta eletromagnética ressonantes, que são medidos pela bobina receptora. As medições geralmente correspondem a pontos ao longo de um caminho prescrito através da representação de Fourier multidimensional do corpo imageado.

- Formato DICOM

O Formato DICOM (Digital Imaging and Communications in Medicine) é um padrão amplamente adotado para aquisição, armazenamento, troca e visualização de imagens médicas. Ele foi desenvolvido para garantir a interoperabilidade e a compatibilidade entre diferentes sistemas e dispositivos médicos, permitindo a comunicação eficiente e confiável de informações médicas (DICOM, 2023).

- NIFTI

O formato NIFTI consiste em um arquivo binário que contém informações sobre a aquisição e as características da imagem, juntamente com os próprios dados de imagem. Uma das principais vantagens do formato NIFTI é a capacidade de armazenar dados em três dimensões (3D) ou quatro dimensões (4D) para imagens dinâmicas, como sequências temporais (LAROBINA, Michele; MURINO, Loredana, 2013). Ele permite que os pesquisadores acessem e analisem facilmente os dados de imagem usando várias ferramentas de processamento e análise.

A imagem é de padrão no formato DICOM e é pré-processada para o formato NIFTI para ser trabalhada. No caso deste trabalho, as imagens já estão no formato NIFTI.

- 3D UNET

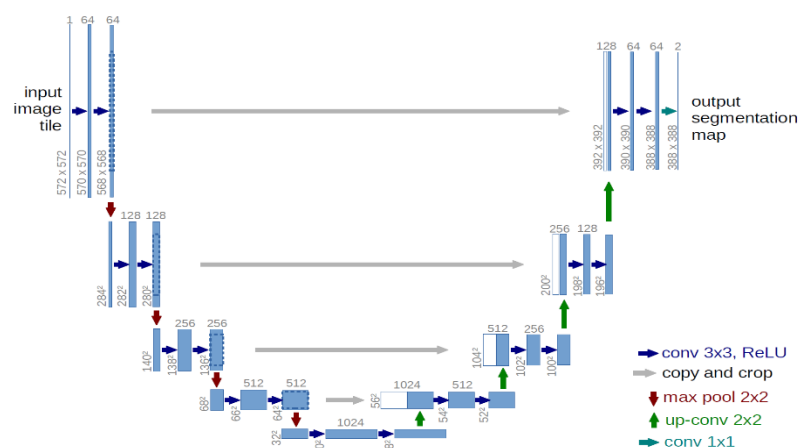


Figura 1 - 3D U-NET- Fonte: (RONNEBERGER, et al. 2015).

Este modelo, inicialmente concebido para segmentação de imagens 2D, foi expandido para trabalhar em três dimensões, tornando-o ideal para processar volumes completos de imagem, como os obtidos em exames de IRM (RONNEBERGER, et al. 2015).

A U-Net 3D, arquitetura representada na Figura 1, mantém a estrutura codificador-decodificador da sua contraparte 2D, porém, opera em três dimensões. Isso significa que ela pode analisar informações tridimensionais em vez de apenas fatias bidimensionais das imagens. Essa capacidade é fundamental para a detecção precisa de estruturas complexas em imagens volumétricas, como as obtidas em exames de IRM cerebral. Uma característica notável da U-Net é a utilização de conexões residuais, que permitem a preservação de detalhes de alta resolução durante o processo de decodificação (RONNEBERGER, et al. 2015).

Para compreender o funcionamento da U-Net 3D, é crucial entender o conceito de "encoder" e "decoder". O "encoder" inicialmente reduz as dimensões da imagem e extrai características importantes, enquanto o "decoder" reverte esse processo, aumentando as dimensões e restaurando a resolução, permitindo a segmentação precisa. As conexões residuais mencionadas anteriormente permitem que informações detalhadas, que normalmente seriam perdidas durante a redução de dimensões, sejam reintegradas no processo de decodificação (ÇIÇEK, et al. 2016).

A aplicação mais comum da U-Net 3D é a segmentação de estruturas médicas, como tumores cerebrais em imagens de IRM. Sua habilidade de aprender a forma tridimensional das estruturas torna-a inestimável para identificar regiões de interesse em imagens médicas complexas (ÇIÇEK, et al. 2016).

Etapas do Método:

Pré-processamento dos dados

Para serem utilizadas no modelo 3D U-Net as imagens necessitam ainda de outros pré-processamentos.

- Gerar subvolumes: Nesta etapa é necessário gerar “patches” dos dados. Devido a necessidade de grande processamento, realizou-se a divisão em subvolumes na região do tumor para ser possível extrair dados dessa região.
- Padronizar os dados: Para facilitar o aprendizado da rede neural artificial aplica-se a normalização para padronizar a imagem (média 0 e desvio padrão 1).

Métricas e Perda

O Coeficiente Dice (Dice Similarity Coefficient - DSC) é essencialmente uma medida de sobreposição entre duas amostras.

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

A Equação (1) é uma primeira avaliação para calcular o Coeficiente Dice de uma única camada, porém como são 4 tipos imagens referentes ao mesmo cérebro: T1, T1ce, T2 e Flair, pode-se utilizar a média entre elas.

$$DSC(f, x, y) = \frac{1}{4} (DSC1(f, x, y) + DSC2(f, x, y) + DSC3(f, x, y) + DSC4(f, x, y)) \quad (2)$$

A Equação (2) realiza a média, onde o índice varia de 0 (incompatibilidade completa) para 1 (correspondência perfeita).

O Coeficiente Dice (DSC) é uma métrica comum usada na avaliação de segmentações médicas. Ela mede a sobreposição entre a máscara segmentada predita pela rede neural e a máscara de referência, segundo DICE, L. R. (1945). Quanto maior a sobreposição, melhor a segmentação. No entanto, o Coeficiente Dice (DSC) padrão é rígida e não leva em consideração a incerteza nas previsões. Isso pode ser problemático, pois as bordas das estruturas podem ser difíceis de definir nas imagens médicas.

Coeficiente de Similaridade de Dados (Dice Similarity Coefficient - DSC)

```
def dice_coef(y_true, y_pred, smooth=1.0):
    class_num = 4
    for i in range(class_num):
        y_true_f = K.flatten(y_true[:, :, :, i])
        y_pred_f = K.flatten(y_pred[:, :, :, i])
        intersection = K.sum(y_true_f * y_pred_f)
        loss = ((2. * intersection + smooth) / (K.sum(y_true_f) + K.sum(y_pred_f) + smooth))
        if i == 0:
            total_loss = loss
        else:
            total_loss = total_loss + loss
    total_loss = total_loss / class_num
    return total_loss
```

Figura 2 - Imagem Coeficiente de Similaridade de Dados (DSC) - Fonte: Imagem gerada pelos autores

Na Figura 2 é apresentada a aplicação em linguagem Python da Coeficiente de Similaridade de Dados (DSC), a qual é utilizada na aplicação do modelo.

Criando e treinando o Modelo 3D U-Net

O modelo 3D U-Net é treinado com pares de imagens e rótulos. Os rótulos são máscaras que indicam a localização dos tumores nas imagens. Durante o treinamento, o modelo aprende a mapear os patches de imagem para as máscaras de segmentação correspondentes.

Neste estudo é adotado a separação de três subconjuntos: de dados para treinamento, validação e teste deste modelo, separados em 65/20/15.

Avaliação dos Resultados

Para avaliação do desempenho dos resultados será analisado:

- Função de Perda: avaliar como ele está se ajustando aos dados de treinamento
- Acurácia: verificar a proporção de previsões corretas em relação ao número total de previsões.
- Coeficiente Dice (DSC): avaliar a sobreposição entre a máscara prevista e a máscara de referência.
- Índice de Jaccard (IoU): verificar a sobreposição entre a máscara prevista e a máscara de referência para cada classe (ou região) de interesse.

Resultados

Os resultados apresentados a seguir são referentes ao processamento de imagens de ressonância magnética de tumores cerebrais pré-processadas no formato NIFTI, disponíveis pela Competição BraTS2020, um total de 369 grupos de imagens com 5 imagens do mesmo cérebro com imagens de T1, T1c, T2, Flair e uma última sendo uma imagem do tumor segmentado para o treinamento do modelo.

As visualizações a seguir foram geradas a partir do arquivos de treinamento do grupo de imagens BraTS20_Training_009.

Conhecendo os dados criando visualizações e examinando:

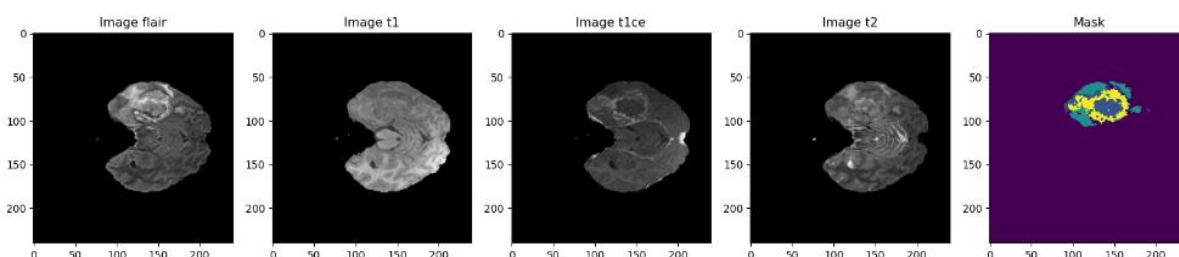


Figura 3 – 5 tipos de imagens do mesmo cérebro - Fonte: Imagem gerada pelos autores

A Figura 3 apresenta as imagens geradas do grupo de imagens BraTS20_Training_009 onde podemos observar 5 tipos imagens referentes ao mesmo cérebro com a localização tumoral através da imagem chamada Mask. Foi utilizada a biblioteca NiBabel para realizar a interação com esse formato de dados NIFTI e a visualização das 5 tipos da esquerda para a direita: Flair, T1, T1ce, T2 e Mask.

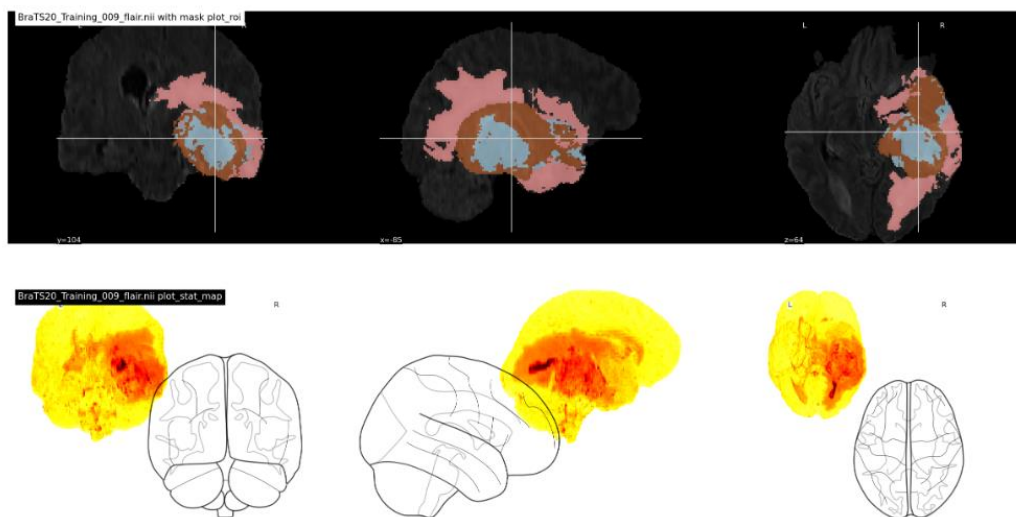


Figura 4 - Imagens de angulos diferentes - Fonte: Imagem gerada pelos autores

Na Figura 4 destaca-se por meio de cores diferentes realçando a região tumoral do cérebro. Pode-se observar nas imagens que é possível “girar” o ângulo para análise de outras perspectivas, aqui apresentadas, da esquerda para direita, como: Coronal, Sagital e Transversal. Foi utilizado a biblioteca Nilearn que possui um conjunto de funções de plotagem para plotar volumes cerebrais. No caso da Figura 4, utilizado plot_glass_brain e plot_roi.

Treinando o Modelo 3D U-NET:

Foi realizado o treinamento do modelo consecutivas vezes com o objetivo de alcançar o modelo com melhor desempenho.

O tempo é extremamente longo e consome considerável processamento. Foi observado que epoch=35 trazia resultados relevantes, com um tempo apesar de alto (39h), então foi adotado como parâmetro nos demais treinamentos em busca de um bom desempenho.

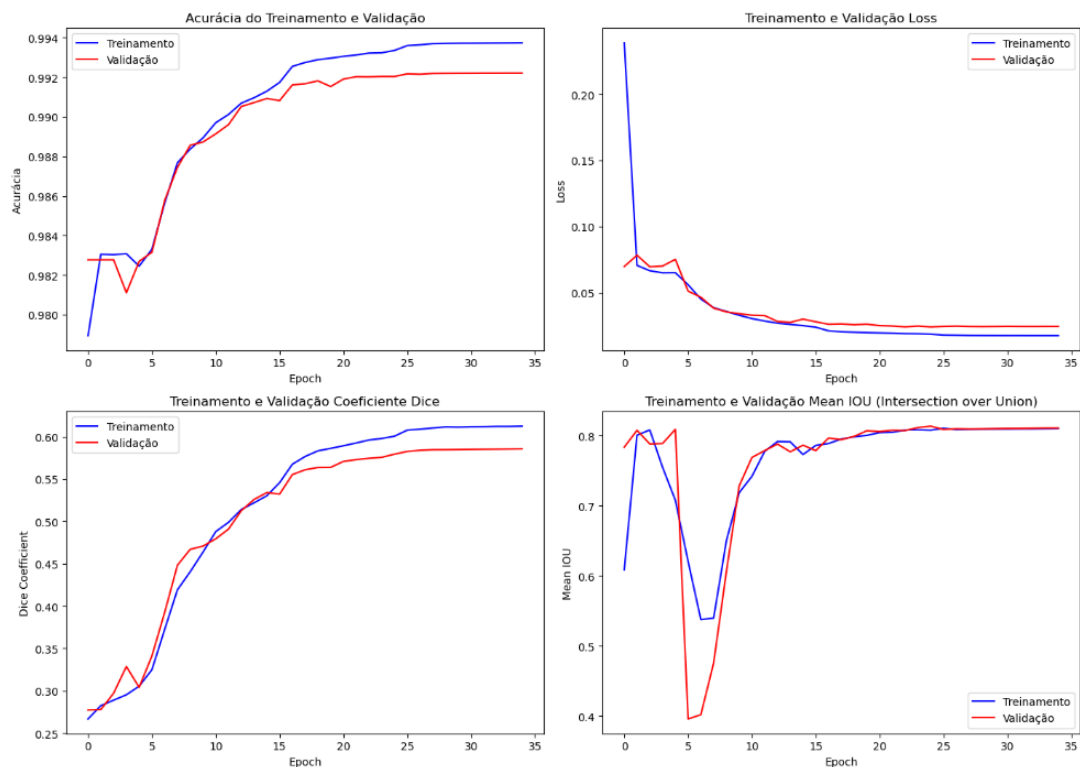


Figura 5 – Métricas de desempenho do modelo - Fonte: Imagem gerada pelos autores

Na Figura 5 temos as métricas como o resultado para avaliação do desempenho deste modelo 3D U-Net treinado. Este resultado foi obtido de um modelo que apresentou melhor desempenho avaliando as métricas.

Realizando Previsões com o Modelo:

Foram utilizadas imagens de ressonância magnética Flair e T1ce, as quais foram redimensionadas e enviadas para o modelo previamente treinado para fazer previsões. O resultado é uma matriz de probabilidades para cada pixel na imagem. Em seguida são exibidas uma série de imagens em uma única figura, incluindo a imagem original, a máscara ground truth e as previsões para as diferentes classes. Isso permite visualizar como o modelo está se saindo na segmentação em comparação com a máscara ground truth.

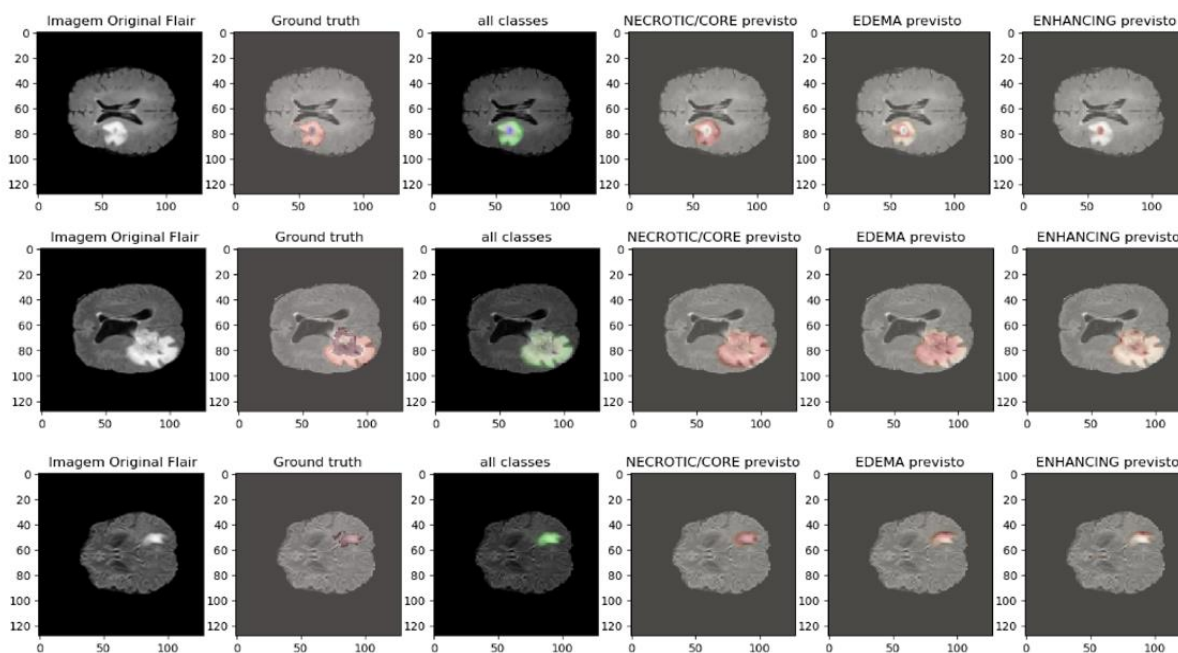


Figura 6 – Previsões do modelo 3D U-Net - Fonte: Imagem gerada pelos autores

Na figura 6 podemos observar as previsões do modelo 3D U-Net treinado para as imagens de treinamento dos grupos BraTS20_Training_006, BraTS20_Training_001 e BraTS20_Training_005, somente as imagens Flair e T1ce desses grupos de imagem. Como saída temos a imagem original Flair, a máscara ground truth, uma contendo todas as classes, o tumor com realce (necrotic/core) previsto, o edema previsto e o tumor sem realce (enhancing) previsto.

Conclusões

Esta pesquisa teve por objetivo realizar um estudo inicial para desenvolver uma aplicação de um conjunto de imagens de Ressonância Magnética ao modelo 3D U-Net para detecção mais rápida e precisa de tumores cerebrais e avaliar seus resultados.

A avaliação do modelo revelou métricas de desempenho bastante relevantes. A acurácia, tanto no treinamento (0.9937) quanto na validação (0.9922), indica que o modelo está fazendo previsões precisas em ambos os conjuntos de dados. No entanto, é importante lembrar que uma alta acurácia não é garantia de um modelo excelente, e, portanto, a avaliação deve ser complementada com outras métricas.

O Índice de Jaccard (IoU), que mede a sobreposição entre as máscaras previstas e as verdadeiras, também revelou resultados relevantes. Com um valor de 0.8101 de treinamento e 0.8107 na validação, o modelo demonstrou uma capacidade sólida de segmentação de objetos nas imagens.

Além disso, o Coeficiente de Dice (DSC), obteve no treinamento 0.6123. Embora ligeiramente mais baixo na validação (0.5856), ainda é considerado razoável. Isso sugere que o modelo está conseguindo uma boa concordância entre as máscaras segmentadas e as máscaras verdadeiras.

Em relação à perda (Loss), os resultados foram igualmente promissores. Com uma baixa perda de treinamento (0.0175), indica que o modelo está se ajustando bem aos dados. A ligeira diferença entre a perda de treinamento e a perda de validação (0.0245) não parece indicar superajuste (overfitting).

O modelo revelou previsões bastante relevantes quando comparadas com as máscaras ground truth, conforme evidenciado pelas métricas e imagens de previsão na Figura 6. No entanto, há espaço para otimização, tanto em termos de eficiência computacional quanto em relação ao coeficiente Dice, que poderia ser aprimorado com ajustes no código. Em resumo, os resultados destacam a eficiência do modelo, mas também indicam oportunidades para futuras melhorias.

Referências

- ÇIÇEK, Ö.; ABDULKADIR, A.; LIENKAMP, S. S.; BROX, T.; RONNEBERGER, O. **3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation**. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2016, pp. 424-432.
- DICE, L. R. **Measures of the amount of ecologic association between species**. In: Ecology, vol. 26, no. 3, 1945, pp. 297-302.
- KNOLL, F. et al. **Fastmri: An open dataset and benchmarks for accelerated mri**, 2019.
- KRIZHEVSKY, A. et al. **ImageNet Classification with Deep Convolutional Neural Networks**. In: Advances in Neural Information Processing Systems, 2012.
- LAROBINA, Michele; MURINO, Loredana. **Medical Image File Formats**. PMC – PubMed Central, 2013.
- RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. **U-Net: Convolutional Networks for Biomedical Image Segmentation**. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
- DICOM Standard**. Disponível em: <https://www.dicomstandard.org/>. Acesso em 28 de maio de 2023.

UM FRAMEWORK PARA SOLUÇÃO DO PROBLEMA INVERSO DO ESPALHAMENTO BASEADO NA FORMULAÇÃO VARIACIONAL ULTRA FRACA

Fernanda Lúcia Sá Ferreira

*CEFET-RJ/Nova Iguaçu e aluna de Doutorado no PPGI/UFRJ
fernanda.ferreira@cefet-rj.br*

Julius Monteiro de Barros Filho

*CEFET-RJ/Nova Iguaçu e aluno de Doutorado no PPGI/UFRJ
julius.filho@cefet-rj.br*

Amaury Alvarez Cruz

*PPGI/UFRJ
amaury@dcc.ufrj.br*

Daniel Gregório Alfaro Vigo

*PPGI/UFRJ
dgalfaro@ic.ufrj.br*

Resumo: Este estudo concentra-se nos problemas de espalhamento inverso em duas dimensões, nos quais informações sobre uma região inacessível são inferidas a partir de medições realizadas em áreas acessíveis. O foco principal está na resolução iterativa do coeficiente de refração-difração na equação de Helmholtz, utilizando um problema de mínimos quadrados regularizado para lidar com a natureza mal-posta do problema. Para resolver o problema direto correspondente, empregamos o Método de Formulação Variacional Ultra Fraca (UWVF), conhecido por sua eficiência computacional, demandando menos recursos e permitindo cálculos analíticos. Além disso, aplicamos com sucesso essa metodologia para determinar a batimetria em regiões costeiras a partir do conhecimento das ondas em águas profundas e na linha costeira, obtendo resultados confiáveis e precisos.

Palavras-chave: Problemas Inversos. Equação de Declive Suave (Mild-Slope Equation - MME). Formulação Variacional Ultra Fraca (Ultra Weak Variational Formulation - UWVF)

1 Introdução

Problemas de espalhamento inverso envolvem a obtenção de informações sobre regiões inacessíveis com base em medições realizadas em áreas acessíveis (PASTORINO, 2008; WANG, 2012; GILBERT; XU, 1993). Este estudo concentra-se no problema inverso associado à equação de Helmholtz em duas dimensões, com o objetivo de determinar iterativamente o coeficiente de refração-difração, enfrentando o desafio da natureza mal-posta do problema por meio de um método de mínimos quadrados regularizado (BORGES; GILLMAN; GREENGARD, 2016).

O alto custo computacional dos problemas inversos é notável devido a vários fatores, incluindo o volume de dados observados, a quantidade de parâmetros a serem estimados e a necessidade de resolver repetidamente o problema direto durante o processo iterativo de minimização (HUTTUNEN T.; KAIPIO, 2002). Portanto, é crucial empregar um método direto eficiente que leve em consideração as principais características do meio, garantindo rapidez e precisão na resolução.

O método numérico escolhido para resolver o problema direto associado foi a Formulação Variacional Ultra Fraca (Ultra-Weak Variational Formulation - UWVF) (CESSENAT; DESPRÉS, 1998), que utiliza ondas planas como base. Esse método se destaca pela eficiência computacional, exigindo menos recursos e permitindo o cálculo analítico das integrais envolvidas. Para a implementação do método, foram utilizadas as ferramentas de otimização e resolução de sistemas lineares do MATLAB.

Neste artigo, destacamos a principal novidade: a aplicação do método de recuperação para resolver o problema inverso de determinar a forma do fundo em águas costeiras. Isso é realizado utilizando informações de entrada das características das ondas em águas profundas (condição de contorno de Dirichlet) e das ondas ao longo da linha costeira (condição de contorno de Robin). Problemas relacionados à detecção da batimetria surgem naturalmente em estudos de dinâmica costeira (COLLINS, 1995), e sua resolução é de extrema importância para a construção de obras de engenharia costeira. Vale destacar que trabalhos similares foram previamente realizados por (VASAN; DECONINCK, 2013; FASSIEH, 2013).

Para isso, utilizamos a Equação de inclinação suave (Mild Slope Equation - MSE), que descreve fenômenos de reflexão, difração e refração de ondas em zonas costeiras (BERKHOFF, 1972). Essa equação modela a propagação de ondas aquáticas à medida que elas viajam através de águas de profundidade variável e interação com limites laterais, como falésias, praias, paredões e quebra-mares. Os resultados numéricos apresentados mostram a confiabilidade e precisão da metodologia proposta.

O artigo está estruturado da seguinte maneira: na Seção 2, apresentamos uma descrição do problema direto com base na equação de Helmholtz. Na Seção 3, delineamos o problema inverso genérico a ser abordado, seguido pela descrição do método aplicado ao caso específico da equação MSE. Em seguida, exploramos como obter informações sobre o leito marinho com base no coeficiente de difração-refração previamente determinado. A Seção 4 detalha o método Ultra-Weak utilizado para resolver numericamente o problema direto. Na Seção 5, oferecemos um resumo do algoritmo desenvolvido para obter a profundidade em função do coeficiente de difração-refração, e, finalmente, na Seção 6, apresentamos um exemplo ilustrativo. Por fim, na Seção 7, destacamos as conclusões do estudo.

2 Problema Direto

Neste caso, o problema direto consiste na solução da equação de Helmholtz com condições de fronteira híbrida de Dirichlet, Neumann e Robin. A seguir, apresentaremos a equação MSE e mostraremos como ela pode ser transformada na equação de Helmholtz. Em seguida, explicaremos a solução do problema inverso de forma genérica. Este procedimento pode ser aplicado ao problema em águas rasas.

2.1 A equação de inclinação suave (MSE)

Fenômenos ondulatórios, como reflexão, refração e difração de ondas, podem ser analisados por meio da Equação de Inclinação Suave (MSE) quando se trabalha em regiões de topografias com declives suaves. Esta equação, originalmente desenvolvida por (BERKHOFF, 1972), sofreu várias modificações para abranger casos mais gerais (ALVAREZ; GARCÍA; SARKIS, 2017).

A Equação de Inclinação Suave (MSE) é expressa da seguinte forma:

$$\nabla \cdot (CC_g \nabla u) + k^2 CC_g u = 0. \quad (1)$$

Nesta equação, $u(x, y) = u_1(x, y) + iu_2(x, y)$ representa a função potencial de velocidade para um fluxo de onda harmônico simples em que (x, y) indica a posição horizontal. Adicionalmente, $k = k(x, y)$, $C = C(x, y)$ e $C_g = C_g(x, y)$ denotam o número de onda, a velocidade de fase e a velocidade de grupo no ponto (x, y) , respectivamente. Estes parâmetros dependem da relação de dispersão e são deduzidos da teoria das ondas de Airy, como se segue:

$$\omega^2 = gk \tanh(kh), \quad C = \frac{\omega}{k}, \quad C_g = \frac{1}{2}C \left[1 + kh \frac{1 - \tanh^2(kh)}{\tanh(kh)} \right],$$

onde, $\omega = 2\pi/T$ representa a frequência angular da onda, T o período da onda e g a aceleração devido à gravidade. Para uma frequência angular dada ω , o número de onda k é determinado pela equação de dispersão que relaciona essas quantidades à profundidade da água $h = h(x, y)$ em cada ponto da região.

A metodologia escolhida para resolver o problema direto neste estudo é a formulação variacional ultra-fracca (UWVF), que está amplamente estabelecida para abordar o problema de contorno de Helmholtz. A transformação $\psi = u\sqrt{CC_g}$ permite que a Equação de Inclinação Suave para águas rasas, conforme a equação (1), seja reescrita como uma Equação de Helmholtz com coeficiente variável (consulte (RADDER, 1979) para uma derivação completa):

$$\Delta\psi + k_c^2\psi = 0, \quad \text{onde} \quad k_c^2 = k^2 - \frac{\Delta\sqrt{CC_g}}{\sqrt{CC_g}}. \quad (2)$$

Dessa forma, a partir deste ponto, descreveremos o método de uma maneira genérica para facilitar a apresentação dos métodos numéricos diretos para resolver essa equação. A Equação (2) é uma forma da equação de Helmholtz homogênea em duas dimensões e pode ser reescrita genericamente como:

$$\Delta\psi + a^2\psi = 0 \text{ em } \Omega, \psi = g_D \text{ em } \partial\Omega_D, \quad \frac{\partial\psi}{\partial n} = g_N \text{ em } \partial\Omega_N, \quad \frac{\partial\psi}{\partial n} + i\alpha a\psi = g_R \text{ em } \partial\Omega_R. \quad (3)$$

Aqui, Δ é o operador laplaciano, $a = a(x, y)$ representa o coeficiente de refração-difração, e $\psi = \psi(x, y)$ é a função de onda que estamos procurando. Supomos que a fronteira de Ω seja composta por $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N \cup \partial\Omega_R$. As condições de Dirichlet são representadas por $g_D = g_D(x, y)$, as condições de Neumann por $g_N = g_N(x, y)$, e as condições de Robin por $g_R = g_R(x, y)$. Cada uma dessas funções é conhecida, e elas são aplicadas, respectivamente, nas partes da fronteira $\partial\Omega_D$, $\partial\Omega_N$ e $\partial\Omega_R$. Além disso, o coeficiente α também é conhecido.

2.2 Condições de contorno para MSE

As regiões onde o MSE é resolvido estão delimitadas por fronteiras fechadas (representadas por linhas costeiras e estruturas que penetram na superfície, como paredes de píeres ou pernas de píeres, quebra-mares, muros de contenção) e uma fronteira aberta, que representa uma fronteira artificial entre a área sendo modelada e a região marítima externa. A fronteira aberta é representada como um tipo de fronteira de Dirichlet. A amplitude e a direção das ondas conformam a condição de contorno na costa marítima:

$$g_D = \psi_o \exp(i(kx \cos(\theta) + ky \sin(\theta))) \quad (x, y) \in \partial\Omega_D, \quad (4)$$

onde ψ_o é a amplitude em águas profundas, θ é o ângulo de direção entre as ondas que se aproximam da fronteira e o vetor normal da fronteira. Os parâmetros k e θ são constantes em águas profundas.

Ao longo da linha costeira e das estruturas que produzem superfícies, a seguinte condição de contorno tem sido tradicionalmente utilizada. No lado interior da geometria, é usada a fronteira do tipo Robin como $CC_g \nabla u \cdot \vec{n} - iqu = p$, onde $\vec{n}$ é a normal exterior do domínio físico em questão, $q(x, y)$ e $p(x, y)$ são parâmetros que dependem das características das ondas (ver por exemplo (TSAY; LIU, 1983; CHEN; HOUSTON, 1987))

3 Problema Inverso

No problema inverso, o objetivo é determinar o coeficiente $a = a(x, y)$ na equação de Helmholtz, considerando que as funções $g_D(x, y)$, $g_N(x, y)$ e $g_R(x, y)$ são conhecidas, como representado a seguir:

$$\Delta\psi + a\psi = 0, \quad \frac{\partial\psi}{\partial n} + i\alpha a\psi = g_R \text{ em } \partial\Omega_R. \quad (5)$$

Partimos do princípio de que, entre todas as soluções obtidas, o coeficiente da equação de Helmholtz que buscamos é aquele que melhor se ajusta à solução na fronteira, satisfazendo as condições de Dirichlet. Isso nos leva a resolver o problema de minimização:

$$\min_{a, \alpha} \mathcal{A}(a, \alpha), \quad \text{onde} \quad \mathcal{A}(a, \alpha) = \sum_{i=1}^N \left\| \frac{\partial\psi_{\omega_i}^a}{\partial n} + i\alpha a\psi_{\omega_i}^a - g_R \right\|_{L^2(\partial\Omega_N)}^2 + \mathcal{J}(a), \quad (6)$$

Nessa equação, $\mathcal{J}(a)$ é um termo de regularização, e $\psi_{\omega_i}^a$ é a solução do problema direto (3) para o coeficiente a e as frequências ω_i para $i = 1, \dots, N$. Para resolver esse problema, aplicamos um método de regularização, que busca obter uma solução aproximada suave e consistente com os dados observados, conforme sugerido por (BORGES; GILLMAN; GREENGARD, 2016; TARANTOLA, 2005). A evidência de que o problema inverso estabelecido está bem colocado pode ser encontrada em (GYLYS-COLWELL, 1996).

4 Solução numérica do problema direto

Para abordar o problema de otimização (6), é crucial resolver o problema direto em cada iteração e encontrar a solução $u^a(x, y)$ para um coeficiente de refração-difração a , a fim de avaliar a função objetivo. Para isso, o método de formulação fraca ou ultra fraca (Ultra Weak Variational Formulation - UWVF) é utilizado. Neste trabalho, o Método Ultra Fraco é usado com base de ondas planas, pois dessa forma as fórmulas para o cálculo da matriz de montagem do método são explícitas. Isso permite obter a solução de maneira rápida e até mesmo possibilita a paralelização de algumas partes do algoritmo.

O método UWVF na equação Hemholtz pode se resumir como segue. Seja a equação

$$\nabla \cdot (\nabla \psi) + k_c^2 \psi = 0 \quad \text{em } \Omega \quad \frac{\partial \psi}{\partial n} - i\sigma \psi = Q \left(-\frac{\partial \psi}{\partial n} - i\sigma \psi \right) - g_r \quad \text{em } \Gamma, \quad (7)$$

onde $k_c^2 = k^2 + r(h)$ e $\sigma = k_c$. Esta é uma formulação mais conveniente para a implementação numérica. Aqui, Ω é um domínio poligonal limitado em $\mathbb{R}^2$ com fronteira Γ e vetor normal unitário n . Em (7), $Q \in \mathbb{C}$, $|Q| \leq 1$, define a condição de contorno, por exemplo, para Dirichlet $Q = -1$, para Neumann $Q = 1$, e para condições de contorno do tipo Robin $Q \neq -1, 1$. O termo de fonte na fronteira é dado por g_r . Considera-se uma partição do domínio em elementos finitos disjuntos, de modo que $\Omega = \cup_{k=1}^K \Omega_k$ e Σ_{kj} denota a interface entre os elementos Ω_k e Ω_j (veja detalhes em (HUTTUNEN T.; KAIPIO, 2002)). A fim de apresentar o método introduzido em (CESSENAT; DESPRÉS, 2003), as equações (7)-(7) são decompostas em sub-problemas para cada elemento Ω_k , $k = 1, \dots, K$

$$\Delta u_k + k_c^2 u_k = 0 \quad \text{em } \Omega_k \quad (8)$$

$$u_k = u_j, \quad \frac{\partial u_k}{\partial n_k} = -\frac{\partial u_j}{\partial n_j} \quad \text{em } \Sigma_{kj} \quad \left(\frac{\partial u_k}{\partial n_k} - i\sigma_k u_k \right) = Q \left(-\frac{\partial u_k}{\partial n_k} - i\sigma_k u_k \right) + g_r \quad \text{em } \Gamma_k \quad (9)$$

Uma nova função é definida em cada fronteira do elemento $\partial\Omega_k$

$$\chi_k = \left(-\frac{\partial u_k}{\partial n_k} - i\sigma_k u_k \right) |_{\partial\Omega_k}, \quad 1 \leq k \leq K. \quad (10)$$

Da decomposição do problema (8)-(9) e integração por partes, segue-se que χ_k verifica

$$\begin{aligned} \sum_{k=1}^K \int_{\partial\Omega_k} \frac{1}{\sigma} \chi_k \overline{\left(-\frac{\partial v_k}{\partial k} - i\sigma v_k \right)} - \sum_{k=1}^K \sum_{j=1}^K \int_{\Sigma_{kj}} \frac{1}{\sigma} \chi_j \overline{\left(\frac{\partial v_k}{\partial k} - i\sigma v_k \right)} \\ - \sum_{k=1}^K \int_{\Gamma_k} \frac{Q}{\sigma} \chi_k \overline{\left(\frac{\partial v_k}{\partial k} - i\sigma v_k \right)} = \sum_{k=1}^K \int_{\Gamma_k} \frac{1}{\sigma} g_r \overline{\left(\frac{\partial v_k}{\partial k} - i\sigma v_k \right)}, \end{aligned} \quad (11)$$

onde v_k é uma função de teste que satisfaz a equação de Helmholtz adjunta local $\Delta v_k + \bar{k}_c^2 v_k = 0$ em Ω_k . Aqui, as barras representam o conjugado complexo. Encontrar $\chi_k \in L^2(\partial\Omega_k)$, $k = 1, \dots, K$, que satisfaça (11) é chamado de UWVF para a equação de Helmholtz. Este problema pode ser discretizado escolhendo uma superposição de ondas planas propagantes para aproximar a função χ_k da seguinte forma:

$$\chi_k^a = \sum_{l=1}^{p_k} \chi_{k,l} \left(-\frac{\partial \phi_{k,l}}{\partial k} - i\sigma \phi_{k,l} \right), \quad (12)$$

onde $\{\chi_{k,l}\}_{l=1}^{p_k}$ são as funções desconhecidas a serem calculadas. As funções de base de ondas planas são dadas por

$$\varphi_{k,l} = \begin{cases} e^{ik(x \cos \theta_l + y \sin \theta_l)} & \text{em } \Omega_k \\ 0 & \text{em outro lugar} \end{cases} \quad (13)$$

onde θ_l é a direção de propagação das ondas que pode ser construída usando direções angulares igualmente espaçadas no círculo unitário, ou seja, $\theta_l = 2\pi(l-1)/p_k$. É importante notar que o número de direções de base p_k pode variar entre os elementos Ω_k , $k = 1, \dots, K$, mas isso não é usado neste artigo.

Substituindo χ_k^a em (11) e escolhendo a função de teste $v_k = \varphi_{k,l}$, a forma discreta do UWVF pode ser escrita na forma de matriz como (ver (CESSENAT; DESPRÉS, 1998; HUTTUNEN T.; KAIPIO, 2002))

$$(D - C)X = b, \quad (14)$$

com os pesos desconhecidos $X = (\chi_{1,1}, \chi_{1,2}, \dots, \chi_{1,p_k}, \chi_{2,1}, \dots)^T$. A matriz D é a diagonal de blocos Hermitianos dada por

$$D_k^{l,m} = \int_{\partial\Omega_k} \frac{1}{\sigma} \left(-\frac{\partial \varphi_{k,m}}{\partial_k} - i\sigma \varphi_{k,m} \right) \overline{\left(-\frac{\partial \varphi_{k,l}}{\partial_k} - i\sigma \varphi_{k,l} \right)} \quad (15)$$

onde o sobrescrito de D mostra os elementos em cada bloco k . As entradas em C são

$$C_{k,j}^{l,m} = \int_{\Sigma_{kj}} \frac{1}{\sigma} \left(\frac{\partial \varphi_{j,m}}{\partial_k} - i\sigma \varphi_{j,m} \right) \overline{\left(\frac{\partial \varphi_{k,l}}{\partial_k} - i\sigma \varphi_{k,l} \right)} + \int_{\Gamma_k} \frac{Q}{\sigma} \left(-\frac{\partial \varphi_{k,m}}{\partial_k} - i\sigma \varphi_{k,m} \right) \overline{\left(\frac{\partial \varphi_{k,l}}{\partial_k} - i\sigma \varphi_{k,l} \right)}, \quad (16)$$

ambas as matrizes têm estrutura de bloco esparsa. O lado direito do sistema é o seguinte

$$b_{k,l} = \int_{\Gamma_k} \frac{1}{\sigma} g_r \overline{\left(\frac{\partial \varphi_{k,l}}{\partial_k} - i\sigma \varphi_{k,l} \right)} \quad (17)$$

Pela a escolha das funções bases como planas como (13) as integrais em (15), (16) e (17) podem ser calculadas de um jeito exato facilitando os cálculos e a paralelização do método.

5 Recuperação da profundidade $h(x, y)$ a partir do coeficiente de refração - difração $k_c(x, y)$

Nesta subseção se mostra como obter a profundidade ou forma do fundo marinho a partir do coeficiente de refração $k_c(x, y)$. Suponha que o coeficiente k_c tenha sido obtido por meio do processo de otimização, conforme descrito na seção anterior mas aplicado na equação (2). Em seguida, se resolve uma Equação Diferencial Parcial (EDP) para determinar o valor da função CC_g e, posteriormente, obteremos a função $h(x, y)$.

Denotando por $v = \sqrt{CC_g}$, da equação (2b) segue a seguinte EDP:

$$\nabla v + (k^2 - k_c^2)v = 0. \quad (18)$$

Suponha que a solução v tenha sido obtida aplicando o método dos elementos finitos à equação (18). Agora, pode-se recuperar a função de profundidade $h(x, y)$ e o número de onda $k(x, y)$ resolvendo o sistema no lineal com duas incógnitas e duas equações para cada ponto (x, y) :

$$\omega^2 = gk(x, y) \tanh(k(x, y)h(x, y)), \quad v(x, y)^2 = \frac{1}{2} \left(\frac{\omega}{k(x, y)} \right) \left[1 + kh \frac{1 - \tanh^2(k(x, y)h(x, y))}{\tanh(k(x, y)h(x, y))} \right],$$

dado que g e ω são constantes conhecidas.

Em resumo, os passos fundamentais para a determinação da batimetria são os seguintes: disponibilizar os dados das ondas em águas profundas e ao longo da linha costeira, aplicar o método descrito na Seção 3 para calcular o coeficiente de difração-refração $k_c(x, y)$, utilizar o método de quase-Newton para resolver o sistema delineado na Seção 4 e, finalmente, determinar a batimetria $h(x, y)$.

6 Exemplo Numérico

Como exemplo de aplicação, consideremos a equação:

$$\nabla^2 \psi(x, y) + (\omega^2 / h(x, y)^2) \psi(x, y) = 0,$$

sobre a geometria $\Omega = [0, 1] \times [0, 1]$, com condições de contorno $\partial\Omega_D = \{(x, y) : x = 0\}$, $\partial\Omega_R = \{(x, y) : x = 1\}$ e $\partial\Omega_N = \{(x, y) : y = 1 \text{ ou } y = 0\}$. As condições de fronteira são definidas como g_D (conforme a equação (4)) e $g_N = 0$. O parâmetro $\alpha = -1$.

A estratégia para obter dados de fronteira Robin envolve a definição de uma função $h(x, y) = 1 + a \cdot \exp(-((x - x_0)^2 + (y - y_0)^2)/b)$, onde $a = 0.5$, $b = 0.05$, $x_0 = 0.6$ e $y_0 = 0.5$. Em seguida, resolvemos o problema direto utilizando o método descrito na seção (4) e determinamos a função g_R , que será usada como entrada no problema inverso. Posteriormente, expressamos a função $h(x, y)$ de forma paramétrica, dependendo dos parâmetros a, x_0, y_0 e b , e iniciamos o processo de recuperação conforme descrito na seção (3).

Para as simulações, escolhemos uma onda com amplitude de 0.02 e testamos três frequências: 0.2, 0.3 e 0.38. Além disso, no método UWVF, selecionamos quatro ondas planas, o que foi suficiente para representar a isotropia dos padrões de difração-refração neste exemplo (ver Figura 1). Adicionalmente, optamos por uma malha com 5000 elementos. Na Figura 2, temos a representação da forma da batimetria, sendo o caso a) os dados reais assumidos ($a = 0.5$, $b = 0.05$, $x_0 = 0.8$ e $y_0 = 0.5$), enquanto o caso b) representa o fundo recuperado ($a = 0.44$, $b = 0.01$, $x_0 = 0.72$ e $y_0 = 0.46$). Casos simulados com diferentes condições iniciais resultaram em boas recuperações, com erro absoluto dos parâmetros menores que 8%. Nos exemplos simulados, o método proposto demonstrou robustez e precisão.

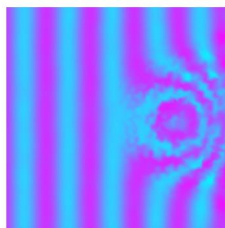


Figura 1: Padrões de difração refração

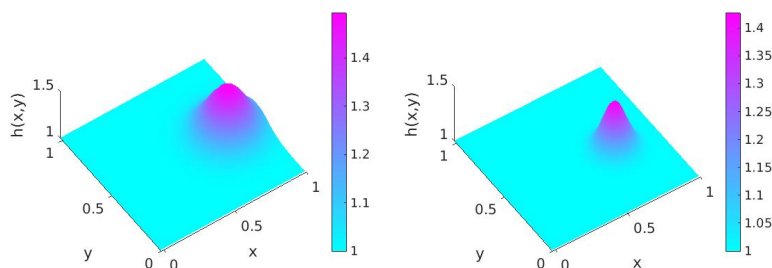


Figura 2: Representação de $h(x, y) = 1 + a \cdot \exp(-((x - x_0)^2 + (y - y_0)^2)/b)$ para a) dados reais (sintéticos), b) recuperados. Se observa boa recuperação, com a maior diferença no parâmetro y_0 .

7 Conclusões

Neste artigo, desenvolvemos um método de recuperação robusto que permite obter o coeficiente de refração-difração com base no conhecimento das características das ondas na fronteira. Esse método

pode ser aplicado em várias aplicações práticas, como na área de sismologia, acústica e engenharia costeira. No artigo, aplicamos esse método para determinar a topografia do fundo do mar com base nas propriedades das ondas em águas profundas e ao longo da linha costeira. O exemplo numérico apresentado demonstra a eficácia do método descrito no artigo.

Referências

- ALVAREZ, A. C.; GARCÍA, G. C.; SARKIS, M. **The ultra weak variational formulation for the modified mild-slope equation**. [S.l.]: Elsevier, 2017. v. 52. 28–41 p.
- BERKHOFF, J. **Computation of combined refraction-diffraction**. [s.n.], 1972. v. 1. 23 p. Disponível em: <<https://journals.tdl.org/icce/index.php/icce/article/view/2767>>.
- BORGES, C.; GILLMAN, A.; GREENGARD, L. **High resolution inverse scattering in two dimensions using recursive linearization**. [S.l.: s.n.], 2016. v. 10. 641-664 p.
- CESSENAT, O.; DESPRÉS, B. **Application of an ultra-weak variational formulation of elliptic PDEs to the two-dimension Helmholtz problem**. [S.l.: s.n.], 1998. v. 35. 255-299 p.
- CESSENAT, O.; DESPRÉS, B. **Using plane waves as base functions for solving time harmonic equations with the ultra weak variational formulation**. [S.l.: s.n.], 2003. v. 11. 227–238 p. Medium-frequency acoustics.
- CHEN, H. S.; HOUSTON, J. R. **Calculation of Water Level Oscillations in Harbors**. [S.l.]: Instructional Rept. CERC-87-2, Waterways Expt. Stn., Vicksburg, Mississippi, 1987.
- COLLINS, R. **Nondestructive testing of materials**. [S.l.]: IOS Press, 1995. v. 8.
- FASSIEH, K. **A Numerical Technique to Estimate Water Depths from Remotely Sensed Water Wave Characteristics**. [S.l.]: Hindawi, 2013. v. 2013.
- GILBERT, R. P.; XU, Y. **An inverse problem for harmonic acoustics in stratified oceans**. [S.l.]: Elsevier, 1993. v. 176. 121–137 p.
- GYLYS-COLWELL, F. **An inverse problem for the Helmholtz equation**. [S.l.]: IOP Publishing, 1996. v. 12. 139 p.
- HUTTUNEN T., M. P.; KAIPIO, J. **Computational aspects of the ultra-weak variational formulation**. [S.l.: s.n.], 2002. v. 182. 27–46 p.
- PASTORINO, M. **Medical and industrial applications of inverse scattering based microwave imaging techniques**. [S.l.], 2008. 34–38 p.
- RADDER, A. C. **On the parabolic equation method for water wave transformation**. [S.l.: s.n.], 1979. v. 95(1). 159-176 p.
- TARANTOLA, A. **Inverse problem theory and methods for model parameter estimation**. [S.l.]: SIAM, 2005.
- TSAY, T.; LIU, P. L. **A Finite Element Model for Wave Refraction and Diffraction**. [S.l.: s.n.], 1983. v. 5. 30-37 p.
- VASAN, V.; DECONINCK, B. **The inverse water wave problem of bathymetry detection**. [S.l.]: Cambridge University Press, 2013. v. 714. 562–590 p.
- WANG, Y. **Regularization for inverse models in remote sensing**. [S.l.]: Sage Publications Sage UK: London, England, 2012. v. 36. 38–59 p.

MÉTODO DA HOMOGENEIZAÇÃO ASSINTÓTICA EM MEIOS ELÁSTICOS UNIDIMENSIONAIS

Jessica Moura da Fonseca
Universidade Federal Fluminense
jemoura@id.uff.br

Reinaldo Rodriguez Ramos
Universidad de La Habana
reinaldo@matcom.uh.cu
Universidade Federal Fluminense
reinaldorr@id.uff.br

Panthers Rodriguez Bermudez
Universidade Federal Fluminense
panthersrb@id.uff.br

Resumo: O Método da Homogeneização Assintótica, se trata da transformação do problema com coeficientes rapidamente oscilantes de um meio heterogêneo, em outro, homogêneo, assintoticamente equivalente ao heterogêneo, mediante uma sequência recorrente de problemas. Este trabalho contribui descrevendo brevemente a teoria e a aplicação do Método da Homogeneização Assintótica. Também contribui e trás como novidade, a exemplificação do método através da implementação computacional na linguagem *Python* e a solução por meio do método de Diferenças Finitas Centradas, para uma comparação gráfica com o resultado do MHA. Para tal exemplo, é utilizado um meio elástico unidimensional, continuamente heterogêneo e periódico e então, são calculados o coeficiente efetivo, problema homogeneizado solução analítica e as soluções assintóticas formais com truncamentos $u^{(1)}$ e $u^{(2)}$. Os resultados, em grandes partes são ilustrados através de gráficos para diferentes valores de α . E ainda, há a contribuição das observações acerca das soluções se aproximarem da solução do problema homogeneizado conforme os valores de α se aproximam de zero. E através disso, foi concluído que o método é uma boa escolha para resolver problemas não homogêneos. Ficou evidente que as soluções dos problemas homogeneizados encontradas foram boas aproximações para as soluções dos problemas originais, onde encontrou-se o comportamento efetivo do meio.

Palavras-chave: Método da Homogeneização Assintótica. Materiais Compósitos. Matemática Aplicada.

Introdução

Materiais compósitos são predominantes tanto na natureza quanto entre os materiais de engenharia, é importante estudá-los porque o que aprendemos com o campo dos compósitos pode ter implicações de longo alcance em muitos campos da ciência. Há um progresso significativo na melhoria da compreensão de como o comportamento microscópico influencia o comportamento macroscópico. E ainda, o estudo de compósitos tem uma longa história, que tem atraído o interesse de alguns dos maiores cientistas (MILTON, 2004).

Na prática, de acordo com Persson et al. (1993), a distribuição real das diferentes fases do compósito é desconhecida. No entanto, em diversos casos, presume-se que a estrutura do material é periódica. Tal suposição é amplamente utilizada e os resultados numéricos se ajustam bem aos dados experimentais. A partir desta suposição de periodicidade, o modelo matemático consiste em equações diferenciais com coeficientes que oscilam periodicamente descrevendo a estrutura espacial do material (PARTON; KUDRYAVTSEV, 1993).

Neste sentido, para a solução de um problema com estrutura compósita, normalmente substitui-se a complexa microestrutura heterogênea do compósito por um material homogêneo com propriedades equivalentes e, em seguida, procedendo com o cálculo dos campos locais na escala estrutural usando técnicas computacionais padrão sem consideração explícita dos detalhes microestruturais. Esta substituição alivia a carga computacional quando estão envolvidas estruturas compostas de grande escala, ao mesmo tempo que garante, até certo ponto, a precisão. É disto que se trata o Método da Homogeneização (HE; PINDERA, 2020; PARTON; KUDRYAVTSEV, 1993; SILVA, 2014).

O Método da Homogeneização Assintótica, também conhecido pelas siglas MHA ou AHM, em inglês, que irá ser tratado neste trabalho, pode facilitar a formulação dos modelos que descrevem o comportamento físico e geométrico do meio, possibilitando a análise da estrutura como um todo. Este método consiste na transformação do problema de um meio heterogêneo, com coeficientes rapidamente oscilantes, chamado problema original, em outro sobre um meio homogêneo, equivalente ao heterogêneo, chamado problema homogeneizado, que normalmente não depende da variável rápida (LIMA et al., 2016; SILVA, 2014).

Segundo Lima et al. (2016), a técnica de modelagem mostra-se uma ferramenta significativa na simulação de fenômenos físicos em meios heterogêneos, alguns exemplos do usos são: 1) descrever a degradação de materiais porosos por Peter (2007); 2) estudo de ondas sísmicas por Capdeville, Guillot e Marigo (2010); 3) determinação de propriedades eletromecânicas efetivas de estruturas ósseas por Silva (2014); 4) análise de propriedades termoelásticas de materiais compósitos por Pinho-da-Cruz et al. (2009).

Tendo em vista a importância deste método e da teoria da elasticidade nos estudos de mecânica dos sólidos, o presente trabalho contribui descrevendo brevemente a teoria e a aplicação do Método da Homogeneização Assintótica. Este artigo também contribui e trás como novidade, a exemplificação do método através da implementação computacional na linguagem *Python* e a solução por meio do método de Diferenças Finitas Centradas, para uma comparação gráfica com o resultado do MHA. Para este exemplo, é utilizado um meio elástico unidimensional, continuamente heterogêneo e periódico e então, são calculados o coeficiente efetivo, função local N , problema homogeneizado solução analítica e as soluções assintóticas formais com truncamentos $u^{(1)}$ e $u^{(2)}$. Os resultados, em grandes partes são ilustrados através de gráficos para diferentes valores de α . E ainda, há a contribuição das observações acerca das soluções se aproximarem da solução do problema homogeneizado conforme os valores de α se aproximam de zero.

Método e Aplicação

Consideremos uma barra elástica não homogênea com módulo elástico e comprimento L , sob a ação da força volumétrica. A barra é um material compósito com estrutura periódica, ou seja, a barra é composta por n barras pequenas feitas de diferentes materiais homogêneos, a figura 1 exemplifica. Os n materiais formam uma célula periódica de comprimento l e supõe-se que $l \ll L$.

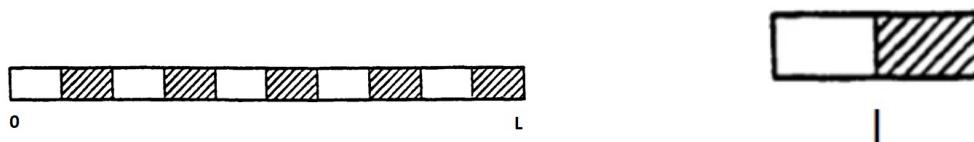


Figura 1: Barra não homogênea por Persson et al. (1993) à esquerda e célula periódica à direita.

Portanto, é uma função 1-periódica, e então o Módulo de Elasticidade $E(y)$ depende da variável rápida y . Assim, o problema consiste em encontrar a solução de $u(x, y)$, que é o deslocamento total, e é 1-periódica em relação à variável rápida. Para isso, declara-se x pertencente ao domínio $\Omega = [0, 1]$, o

Módulo de Elasticidade $E^\alpha(x) \in C^1([0, 1])$, onde $E^\alpha(x)$ é uma representação que indica que o Módulo de Elasticidade depende de α , e é positivo, limitado e α -periódico com relação à x , $X^\alpha(x) \in C^0([0, 1])$, e $u^\alpha(x) \in C^2([0, 1])$, onde $X^\alpha(x)$ e $u^\alpha(x)$ são representações que indicam a dependência de α . Assim, chega-se no seguinte problema original:

$$\begin{cases} \frac{d}{dx} \left[E^\alpha(x) \frac{du}{dx} \right] = X^\alpha(x), & x \in (0, 1) \\ u^\alpha(0) = 0, & u^\alpha(1) = 1. \end{cases} \quad (1)$$

onde, $E(y)$ é a função de oscilação rápida, g_1 é o deslocamento inicial, g_2 é o deslocamento final da barra, $y = \frac{x}{\alpha}$ é a variável rápida, no qual x a variável lenta e $\alpha = \frac{l}{L}$.

Para determinar a Solução Assintótica Formal, conhecida como SAF, de segunda ordem para o problema original, considera-se a aproximação assintótica abaixo com truncamento em $u^{(2)}$:

$$u^{(2)}(x, \alpha) = u_0(x, y) + \alpha u_1(x, y) + \alpha^2 u_2(x, y). \quad (2)$$

Substitui-se (2) em (1), utilizando a regra da cadeia e o operador (3) para simplificar.

$$L_{\alpha\beta} = \frac{\partial}{\partial \alpha} \left(E(y) \frac{\partial}{\partial \beta} \right), \text{ com } \alpha, \beta \in (x, y). \quad (3)$$

Posteriormente, reorganiza-se a expressão e encontra-se uma sequência de equações em potências de α :

$$\begin{aligned} \alpha^{-2} : L_{yy}u_0 &= 0, \\ \alpha^{-1} : L_{yy}u_1 &= -L_{yx}u_0 - L_{xy}u_0 \\ \alpha^0 : L_{yy}u_2 &= -L_{xx}u_0 - L_{yx}u_1 - L_{xy}u_1 + X(x, y). \end{aligned} \quad (4)$$

Para determinar a equação do problema local e do problema homogeneizado são solucionadas as equações (4). Para resolver esses problemas é necessário enunciar o Lema a seguir, que fornece uma condição necessária e suficiente para a existência e unicidade de soluções 1-periódicas em y dos problemas mencionados.

Lema: Considere $X(y)$ e $E(y)$ funções diferenciáveis e 1-periódicas, com $E(y)$ positiva. Uma condição necessária e suficiente para existir uma solução 1-periódica para a equação $L_{yy}N = X$ é que $\langle X(y) \rangle = 0$. Onde o operador $\langle \cdot \rangle$ representa o valor médio, dado por $\langle \cdot \rangle \equiv \int_0^1 (\cdot) dy$. Além disso, a solução é única e da forma $N(y) = \tilde{N}(y) + C$, onde C é uma constante arbitrária e $\tilde{N}$ é a solução 1-periódica de $L_{yy}N = X$ com a condição de que $\tilde{N}(0) = 0$ (BAKHVALOV; PANASENKO, 1984). A demonstração do lema pode ser conferida em obras como Bakhvalov e Panasenko (1984) e Leitzke (2017).

Depois, pode-se calcular o coeficiente efetivo, representado por $\hat{E}$, através de (5).

$$\hat{E} = \langle E^{-1} \rangle^{-1}. \quad (5)$$

E segue-se para a solução N_1 , mencionada no Lema.

$$N_1 = \int_0^y \left(\frac{\hat{E}}{E(s)} - 1 \right) ds. \quad (6)$$

Com o N_1 , pode-se encontrar $u_1(x, y)$, pois:

$$u_1(x, y) = N_1(y) \frac{du_0}{dx}. \quad (7)$$

Utilizando os conhecimentos até aqui, e através de manipulações, pode-se chegar na fórmula de N_2 :

$$N_2(y) = \int_0^y \frac{1}{E(s)} \left(\int_0^s (-\hat{E} + f(r)) dr - \left\langle \int_0^y (-\hat{E} + f(s)) ds \right\rangle + \langle N_1(y) \rangle \hat{E} \right) ds$$

$$- \int_0^y N_1(s) ds. \quad (8)$$

Com o N_2 , pode-se encontrar $u_2(x, y)$:

$$u_2(x, y) = N_2 \frac{d^2 u_0}{dx^2} \quad (9)$$

A partir de (2), a SAF com truncamento em $u^{(1)}(x, \alpha)$ será:

$$u^{(1)}(x, \alpha) = u_0(x) + \alpha N_1(y) \frac{du_0}{dx}, \quad (10)$$

e a SAF com truncamento em $u^{(2)}(x, \alpha)$ será:

$$u^{(2)}(x, \alpha) = u_0(x) + \alpha N_1(y) \frac{du_0}{dx} + \alpha^2 N_2(y) \frac{d^2 u_0}{dx^2}. \quad (11)$$

A partir de tais conhecimentos, pode-se seguir para exemplos e análises.

Exemplo e Análises

A partir do problema (1), serão consideradas condições de contorno não homogêneas, dadas por $u(0) = 0$ e $u(1) = 1$. Para o coeficiente original, $E(y) = 2 + 0.75 \cos(\pi y)$, e, para a fonte, $X(y) = -2 \cos(\pi y)$ (LEITZKE, 2017). É construído então, o problema de valor de contorno a seguir:

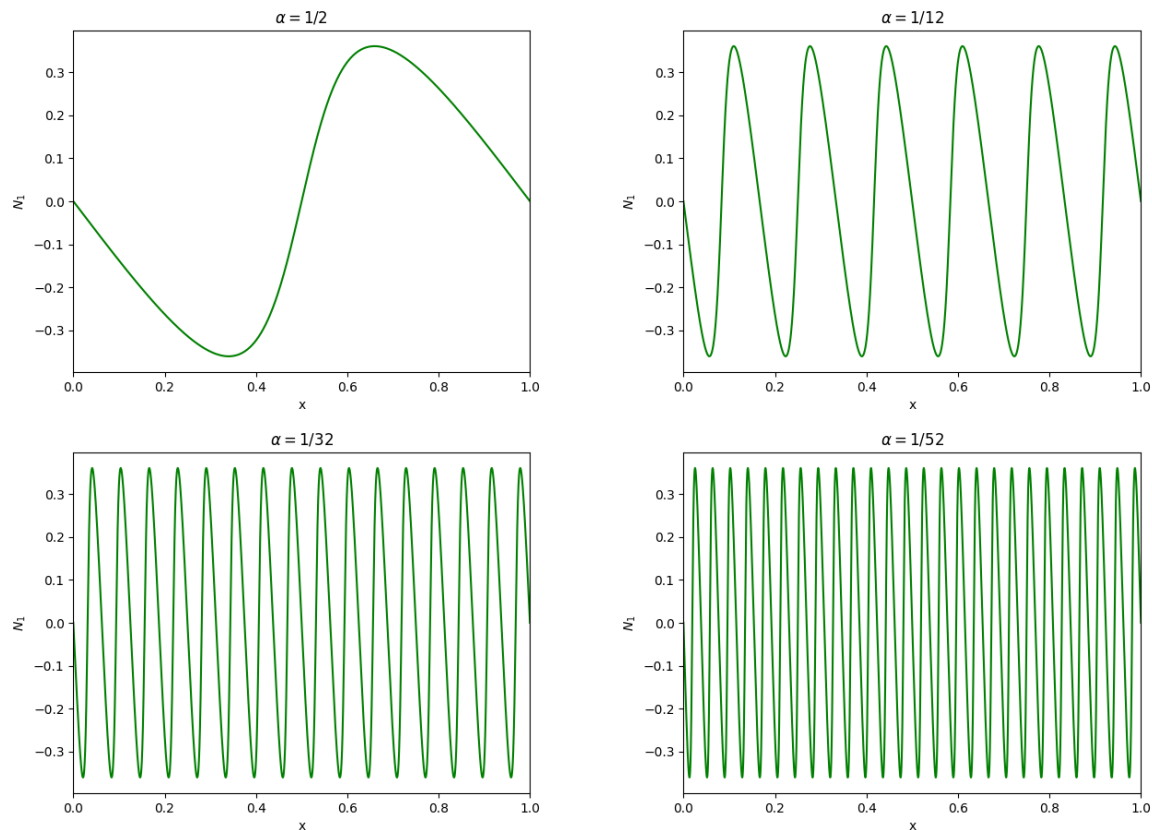
$$\begin{cases} \frac{d}{dx} \left[2 + 0.75 \cos(\pi y) \frac{du}{dx} \right] = -2 \cos(\pi y), \\ u^\alpha(0) = 0, \quad u^\alpha(1) = 1. \end{cases} \quad (12)$$

Primeiramente foi encontrado o coeficiente efetivo: $\hat{E} = \frac{\sqrt{55}}{4}$.

Depois, encontra-se o N_1 , através da equação (6), resultando em:

$$N_1(y) = \begin{cases} \frac{2}{\pi} \arctan \left(\frac{5 \tan(\pi y/2)}{\sqrt{55}} \right) - y, & \text{se } 0 < y < 1/2, \\ 0, & \text{se } y = 1/2, \\ \frac{2}{\pi} \arctan \left(\frac{5 \tan(\pi y/2)}{\sqrt{55}} \right) - y + 1, & \text{se } 1/2 < y < 1. \end{cases} \quad (13)$$

A imagem mostra a solução $N_1(x/\alpha)$ do problema local. Através dela, também é possível perceber, com os diferentes valores de α , a rápida oscilação conforme α se aproxima de zero.


 Figura 2: Gráficos N_1 para diferentes valores de α .

Ao tentar calcular o N_2 , encontra-se uma função que não há integral. E ainda, após o cálculo e o gráfico do N_1 finalizados, calcula-se o problema homogeneizado, que resulta em:

$$\begin{cases} \frac{d^2 u_0}{dx^2} = 0, & x \in \Omega, \\ u_0(0) = 0, & u_0(1) = 1. \end{cases} \quad (14)$$

E calcula-se então, a solução $u_0(x)$:

$$u_0(x) = x. \quad (15)$$

Como (15) resultou em x , tem-se que $\frac{du_0}{dx} = 1$ através de tais informações, calcula-se as SAFs $u^1(x, \alpha)$, representada em (10), que resultará:

$$u^{(1)}(x, \alpha) = x + \alpha N_1 \left(\frac{x}{\alpha} \right), \quad (16)$$

E como $\frac{d^2 u_0}{dx^2} = 0$, a SAF $u^{(2)}(x, \alpha)$, representada em (11) será:

$$u^{(2)}(x, \alpha) = x + \alpha N_1 \left(\frac{x}{\alpha} \right) + \alpha^2 N_2 \left(\frac{x}{\alpha} \right) 0. \quad (17)$$

Em relação às soluções $u_1(x, y)$ e $u_2(x, y)$, de (7) e de (15), temos que o primeiro deles é:

$$u_1(x, y) = N \quad (18)$$

Além disso, novamente, como $\frac{d^2 u_0}{dx^2} = 0$ e de (9) tem-se que $u_2(x, y)$ é:

$$u_2(x, y) = 0. \quad (19)$$

A solução analítica do problema original é:

$$u^{(\alpha)}(x) = \frac{8\alpha^2}{3\pi^2} \log\left(\frac{2 + 0.75 \cos(\pi x/\alpha)}{2.75}\right) + \frac{\sqrt{55}}{4} \left(1 + \frac{8\alpha^2}{3\pi^2} \log\left(\frac{2 + 0.75 \cos(\pi/\alpha)}{2.75}\right)\right) \int_0^x \frac{1}{2 + 0.75 \cos(\pi r/\alpha)} dr. \quad (20)$$

Em seguida, através da linguagem de programação *Python*, foi implementado um código para resolver o problema (12) através do Método de Diferenças Finitas Centradas, MDF, para que pudesse ser feita a comparação de ambos os resultados e validação do MHA.

Para aplicar o MDF, a equação (12) foi reorganizada em:

$$\frac{d^2 u}{dx^2} - \frac{3\pi}{4\alpha} \frac{\sin(\pi x/\alpha)}{2 + 0.75 \cos(\pi x/\alpha)} \frac{du}{dx} = \frac{-2 \cos(\pi x/\alpha)}{2 + 0.75 \cos(\pi x/\alpha)}. \quad (21)$$

A figura ilustra, para diferentes valores de α , as soluções: analítica, pelo Método de Diferenças Finitas, do problema homogeneizado e as SAFs $u^{(1)}$ e $u^{(2)}$.

Ao realizar as análises gráficas, nota-se que a solução analítica e a solução pelo MDF coincidem, em todos os gráficos. Isto indica que a solução analítica pelo MHA funciona, e com resultados satisfatórios.

Em relação as SAFs, tem-se que $\frac{d^2 u_0}{dx^2} = 0$, então, $\alpha^2 N_2(y) \frac{d^2 u_0}{dx^2} = 0$, tornando então, a SAF $u^{(2)}$ idêntica a SAF $u^{(1)}$. Assim, neste exemplo em específico não será possível identificar se a SAF $u^{(2)}$ será mais precisa que a SAF $u^{(1)}$, novos exemplos deverão ser testados em trabalhos futuros.

Outro ponto importante e que deve ser percebido, provém da observação de que, conforme α tende a zero, as soluções: analítica, pelo MDF, e as SAFs, ficam cada vez mais próximas da solução do problema homogeneizado.

Para que a legenda da figura abaixo fique compreensível: SDF é a solução pelo Método de Diferenças Finitas, SAF significa Solução Assintótica Forma, HOM é o problema homogeneizado e SA, é a solução analítica.

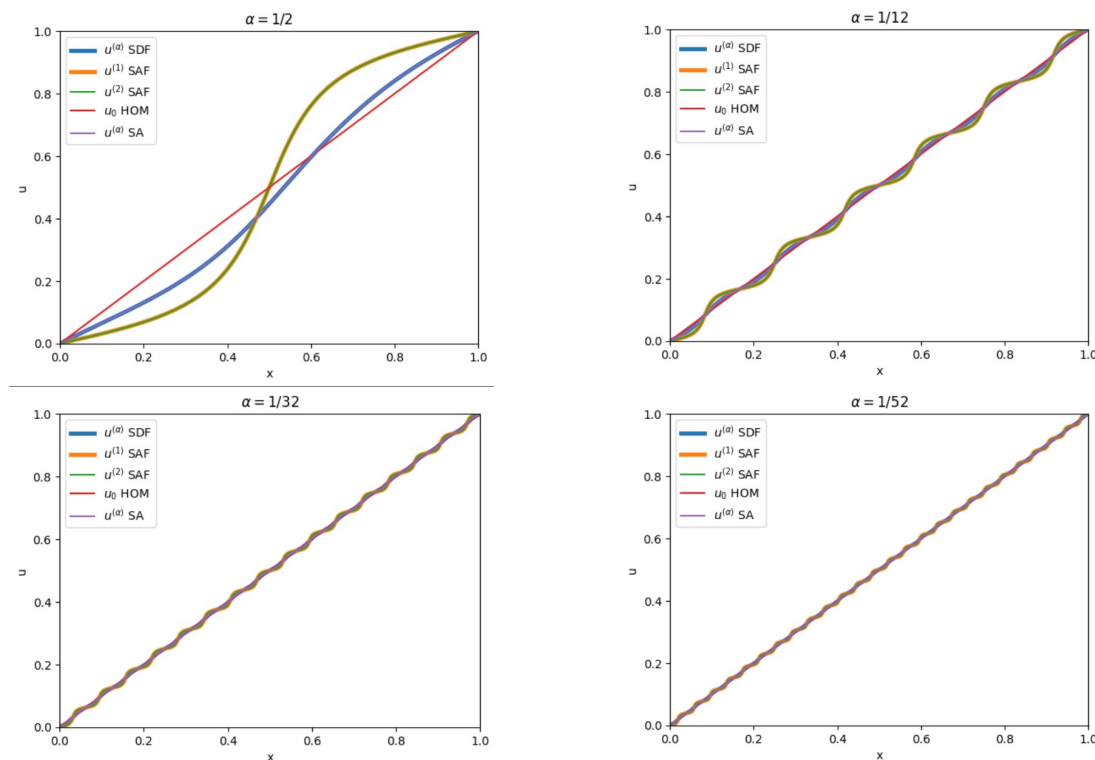


Figura 3: Gráficos u para diferentes valores de α , obtidos com diferentes métodos.

Conclusões

Este trabalho passou brevemente pela aplicação do Método da Homogeneização Assintótica em meios elásticos lineares e unidimensionais, e, para ilustrar, foi feito um exemplo que resultou em um problema homogeneizado com equação homogênea e condições de contorno não-homogêneas. Ambas as SAFs tiveram o mesmo resultado e ainda, a solução analítica do problema original e a solução pelo MDF resultaram no mesmo.

A partir dos cálculos e desenvolvimentos apresentados conclui-se que o MHA é uma boa escolha para resolver problemas não homogêneos. Ficou evidente que as soluções dos problemas homogeneizados encontradas foram uma boa aproximação para as soluções dos problemas originais, onde encontrou-se o comportamento efetivo do meio. E além disso, considerando as SAFs de primeira e segunda ordens pôde-se notar que para este exemplo, elas resultaram no mesmo, sendo então necessários, mais exemplos para analisar se mais termos na SAF torna os resultados ainda mais satisfatórios.

Referências

- BAKHVALOV, N.; PANASENKO, G. **Homogenisation: Averaging Processes in Periodic Media**. Holanda: Kluwer Academic Publishers, 1984.
- CAPDEVILLE, Yann; GUILLOT, Laurent; MARIGO, Jean-Jacques. 1-D non-periodic homogenization for the seismic wave equation. **Geophysical Journal International**, v. 181, p. 897–910, 2010.
- HE, Zhelong; PINDER, Marek-Jerzy. Locally exact asymptotic homogenization of periodic materials under anti-plane shear loading. **European Journal of Mechanics / A Solids**, v. 81, 2020.
- LEITZKE, Bruna S. **Soluções Assintóticas Formais de Segunda Ordem na Homogeneização Assintótica de Meios Microperiódicos**. 2017. Dissertação (Mestrado) – Universidade Federal de Pelotas, Pelotas.
- LIMA, Marcos Pinheiro de et al. Homogeneização Assintótica da Equação do Calor para Meios Unidimensionais Periódicos Continuamente Heterogêneos. **Periódicos Científicos da Universidade Federal do Rio Grande – FURG**, v. 26, p. 73–83, 2016.
- MILTON, Graeme W. **The Theory of Composites**. Cambridge: Cambridge University Press, 2004.
- PARTON, Vladimir Zalmanovich; KUDRYAVTSEV, Boris Aleksandrovich. **Engineering mechanics of composite structures**. USA: CRC Press, 1993.
- PERSSON, Lars Erik et al. **The Homogenization Method**. Suécia: Studentlitteratur, 1993.
- PETER, Malte A. Homogenisation of a chemical degradation mechanism inducing an evolving microstructure. **Comptes Rendus Mecanique**, v. 335, p. 679–684, 2007.
- PINHO-DA-CRUZ, Joaquim et al. Homogeneização de Propriedades de Materiais Compósitos em Termoelasticidade Linear. **Revista Iberoamericana de Ingeniería Mecánica**, v. 13, p. 3–23, 2009.
- SILVA, Uziel Paulo da. **Emprego do Método de Homogeneização Assintótica no Cálculo das Propriedades Efetivas de Estruturas Ósseas**. 2014. Dissertação (Doutorado) – Universidade de São Paulo, São Carlos.

Caracterização dos grafos com conectividade algébrica $n - 4$

Laura Maria Silva Miranda

*Instituto Federal do Sudeste de Minas Gerais, Juiz de Fora, MG
lauramariamiranda01@gmail.com*

Átila Arueira Jones

*Instituto Federal do Sudeste de Minas Gerais, Juiz de Fora, MG
atila.jones@ifsudestemg.edu.br*

Allana Sthel Santos de Oliveira

*Instituto Federal do Sudeste de Minas Gerais, Juiz de Fora, MG
allana.sthel@ifsudestemg.edu.br*

Resumo: A conectividade algébrica de um grafo, $a(G)$, é definida como o segundo menor autovalor laplaciano de G . Esse parâmetro tem uma boa relação com a conexidade do grafo, sendo alvo de inúmeros estudos em razão de sua alta aplicabilidade em novas tecnologias. Motivados por uma dessas aplicações e sabendo que o valor máximo de $a(G)$ é n , neste trabalho caracterizamos os grafos em que $a(G) = n - 4$, onde n é o número de vértices de G .

Palavras-chave: Conectividade algébrica. Caracterização. Grafos.

Introdução

Podemos definir um veículo autônomo como aquele que possui um conjunto de tecnologias cognitivas e de controle altamente automatizadas. A pesquisa pelo controle descentralizado desses veículos vem se desenvolvendo nos últimos anos devido à alta aplicabilidade que essa tecnologia oferece. A saber, o controle é descentralizado quando não existe uma central que tem acesso e controla todos os agentes, estes então interagem entre si. A comunicação entre veículos autônomos pode ser modelada por grafos orientados. Em (FAX; MURRAY, 2004) são utilizadas técnicas da teoria espectral de grafos para estudar a estabilização de veículos descentralizados, com comunicação bidirecional (sem orientação). No artigo é provado que a taxa de convergência para estabilidade está diretamente relacionada com o segundo menor autovalor da matriz laplaciana do grafo, conhecido como conectividade algébrica, de forma que o aumento deste valor implica no aumento da taxa de convergência.

Baseando-se na questão supracitada, é de interesse buscar grafos que maximizem a conectividade algébrica, visando uma eficiência na estabilidade do sistema. Na literatura, encontramos trabalhos que atuam neste interesse, como (OGIWARA, 2015), que busca maximizar tal parâmetro fixando o número de vértices e arestas; já em (LI, 2018) é estudado o aumento da conectividade algébrica através de inserção de arestas; e ainda vale citar (KOLOKOLNIKOV, 2015), que se restringe a específicas classes de grafos. Tendo vista do interesse nesta área, aliado à motivação da sua aplicação em controle de veículos autônomos, começamos a desenvolver resultados acerca da maximização da conectividade algébrica.

Neste trabalho, focamos na caracterização de grafos com n vértices cuja conectividade algébrica é igual a $n - 4$, estendendo os resultados obtidos em (DOS SANTOS, 2022).

Base teórica

Um grafo é uma estrutura $G = G(V, E)$, constituída por um conjunto finito e não vazio V , cujo elementos são denominados **vértices**, e um conjunto E de subconjuntos a dois elementos de V , denominados **arestas**. O **grau de um vértice** v , denotado por $d(v)$, é o número de arestas que incidem em v . Vértices u, v ligados por arestas são denominados **adjacentes** e denotamos por $uv \in E$. O **grau mínimo** de G é o

número $\delta(G) = \min\{d(v) : v \in V\}$, enquanto o **grau máximo** de G é dado por $\Delta(G) = \max\{d(v) : v \in V\}$. Diremos que um grafo $H(V', E')$ é **subgrafo** de $G(V, E)$, se $V' \subset V$ e $E' \subset E$, denotando assim $H \subset G$.

Um **caminho** é uma sequência finita $v_1, v_2, \dots, v_k$ de vértices distintos de um grafo $G(V, E)$ em que v_i e v_{i+1} são adjacentes para $1 \leq i \leq k-1$. Se v_1 é adjacente a v_k , temos um **ciclo**. O grafo formado por um caminho de n vértices é denotado por P_n , já C_n denota o grafo formado pelo ciclo de ordem n . O **grafo completo**, com n vértices, é aquele em que quaisquer dois vértices distintos são adjacentes, o denotaremos por K_n .

Um grafo G é dito **conexo** quando existe um caminho ligando cada par de vértices. Caso contrário, o grafo G é denominado **desconexo** e neste caso dizemos que $G' \subset G$ é uma **componente conexa** de G quando G' é um grafo conexo e não existe um grafo conexo $H \subset G$ tal que $G' \subset H$ e $G' \neq H$.

Sejam $G_1 = G_1(V_1, E_1)$ e $G_2 = G_2(V_2, E_2)$ dois grafos, com V_1 e V_2 disjuntos. O **grafo união**, $G = G_1 + G_2$, é aquele cujo conjunto de vértices é $V_1 \cup V_2$ e cujo conjunto de arestas é $E_1 \cup E_2$. Além disso, dado um $k \in \mathbb{N}$, podemos escrever kG para indicar o grafo que é união de k cópias disjuntas de G .

Diremos que G é combinação linear dos grafos H_i , $1 \leq i \leq A$, se G pode ser escrito como abaixo, onde $a_i \in \mathbb{N}$, $1 \leq i \leq A$ e $a_A \neq 0$,

$$G = a_1 H_1 + a_2 H_2 + \dots + a_A H_A = \sum_{i=1}^A a_i H_i \quad (1)$$

Dado o grafo $G = G(V, E)$, o seu **grafo complementar** $\bar{G} = \bar{G}(V, \bar{E})$ é aquele em que $\bar{V} = V$ e $v_i v_j \in \bar{E}$ se, e somente se, $v_i v_j \notin E$.

A Teoria Espectral de Grafos é uma linha de investigação onde se procura determinar características da estrutura de um grafo a partir de propriedades espectrais de matrizes associadas a ele. A **matriz de adjacência** de um grafo G , denotada por $A(G)$, é uma matriz quadrada de ordem n cujas entradas a_{ij} valem 1 se v_i e v_j são adjacentes e 0, caso contrário. Definimos $L(G) = D(G) - A(G)$ como a **matriz laplaciana** do grafo G , em que $D(G)$ é a **matriz diagonal dos graus** dos vértices de um grafo G , ou seja, $[D(G)]_{ii} = d(v_i)$.

O **autovalores laplacianos** de G , $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$, são os autovalores de $L(G)$. Se um autovalor μ tem multiplicidade $p > 1$, indicaremos por $\mu^{(p)}$. O multiconjunto constituído pelos autovalores laplacianos de G é denominado **espectro laplaciano** de G e é denotado por $\zeta(G)$.

O menor autovalor laplaciano de um grafo qualquer G é $\mu_n = 0$. Já o segundo menor autovalor laplaciano de G , μ_{n-1} , é denominado **conectividade algébrica** do grafo G e é denotado por $a(G)$. A conectividade algébrica é um parâmetro importante na teoria espectral de grafos e é alvo de muitos trabalhos, dos quais citamos o *survey* (DE ABREU, 2007). O maior autovalor do laplaciano de G é denominado **índice laplaciano** de G , denotado por $\mu_1(G)$.

A Tabela 1 mostra o espectro laplaciano dos principais grafos utilizados no presente texto, evidenciando ainda a conectividade algébrica e índice laplaciano. Estas e outras propriedades espectrais podem ser encontradas em (BROUWER; HAEMERS, 2011).

Tabela 1: Espectro Laplaciano de K_n, P_n, C_n

G	Espectro Laplaciano $\zeta(G)$	$a(G)$	$\mu_1(G)$
K_n	$\{n^{(n-1)}, 0^{(1)}\}$	n	n
P_n	$\left\{2 - 2 \cos\left(\frac{\pi j}{n}\right)\right\}, j \in \{0, \dots, n-1\}$	$2 - 2 \cos\left(\frac{\pi}{n}\right)$	$2 + 2 \cos\left(\frac{\pi}{n}\right)$
C_n	$\left\{2 - 2 \cos\left(\frac{2\pi j}{n}\right)\right\}, j \in \{0, \dots, n-1\}$	$2 - 2 \cos\left(\frac{2\pi}{n}\right)$	4, se n é par; $2 + 2 \cos\left(\frac{\pi}{n}\right)$, se n é ímpar.

O resultado a seguir pode ser visto como o ponto de partida dos estudos nesse trabalho, visto que associa o maior valor possível de $a(G)$ ao número de vértices n .

Proposição 1 (FIEDLER, 1973) *Para o grafo completo K_n vale $a(K_n) = n$. Se G , com n vértices, não é o grafo completo, então $a(G) \leq n - 2$.*

A **conectividade de vértices** de um grafo G , denotada por $k(G)$, é o menor número de vértices que, ao serem retirados, tornam o grafo desconexo. Já a **conectividade de arestas**, denotada por $k'(G)$, é o menor número de arestas que, ao serem retiradas, tornam o grafo desconexo.

Proposição 2 (FIEDLER, 1973) *Se G não é o grafo completo, então $a(G) \leq k(G) \leq k'(G) \leq \delta(G)$*

É possível ainda obter o comportamento do espectro laplaciano através das operações de união e complementar, descritos nas proposições abaixo.

Proposição 3 (MERRIS, 1998) *Se o espectro laplaciano de um grafo G é $\zeta(G) = \{\mu_1, \mu_2, \dots, \mu_{n-1}, 0\}$, então o espectro laplaciano de $\bar{G}$ é $\zeta(\bar{G}) = \{n - \mu_{n-1}, \dots, n - \mu_1, 0\}$. Em particular, $\mu_1(\bar{G}) = n - a(G)$.*

Proposição 4 (MERRIS, 1998) *Sejam G e H grafos, o índice laplaciano da união é obtido por*

$$\mu_1(G + H) = \max\{\mu_1(G), \mu_1(H)\} \quad (2)$$

O teorema a seguir relaciona o índice laplaciano e o grau máximo de um grafo G .

Proposição 5 (GRONE; MERRIS, 1994) *Seja G um grafo conexo simples com pelo menos uma aresta e grau máximo Δ . Então $\mu_1(G) \geq \Delta + 1$, com igualdade se, e somente se, existe um vértice adjacente a todos os outros vértices em G .*

Caracterização

Inspirados na desigualdade $a(G) \leq n - 2$, para $G \neq K_n$, posta na Proposição 1, o trabalho (DOS SANTOS, 2022) caracteriza todos os grafos em que $n - 3 \leq a(G) \leq n - 2$. Esses resultados são explicitados abaixo.

Proposição 6 (DOS SANTOS, 2022) *Seja G um grafo com n vértices. Temos que $a(G) = n - 2$ se, e somente se, $\delta(G) = n - 2$, ou seja, $\bar{G} = aK_2 + (n - 2a)K_1$, $a = 1, 2, 3, \dots, \lfloor \frac{n}{2} \rfloor$.*

Corolário 1 (DOS SANTOS, 2022) *Não existe grafo G com n vértices e $n - 3 < a(G) < n - 2$.*

Proposição 7 (DOS SANTOS, 2022) *Seja G um grafo com n vértices. Então $a(G) = n - 3$ se, e somente se, $\bar{G} = aK_1 + bK_2 + cK_3 + dP_3$, para a, b, c, d inteiros, em que $c + d > 0$*

Proposição 8 (DOS SANTOS, 2022) *Seja G um grafo com n vértices. Então $n - 4 < a(G) < n - 3$ se, e somente se,*

$$\bar{G} = \sum_{j=1}^A a_{2j+1} C_{2j+1} + \sum_{j=1}^B b_j P_j, \quad (3)$$

em que $A \geq 2$ ou $B \geq 4$. Além disso, em caso afirmativo, vale $a(G) = n - 2 - 2 \cos\left(\frac{\pi}{M}\right)$ para $M = \max\{2A + 1, B\}$.

No presente trabalho, estendemos esses resultados caracterizando grafos tais que $a(G) = n - 4$. Além disso, construímos uma família infinita de grafos que atingem a igualdade da Proposição 2. Vale destacar que as caracterizações foram obtidas com auxílio do software *Graph Filter* (JONES et al., 2022).

O lema enunciado a seguir garante que existe um número restrito de grafos conexos que atendem $\mu_1(G) = 4$ e $\Delta(G) = 3$.

Lema 1 *Seja G um grafo conexo com n vértices, $\mu_1(G) \leq 4$ e $\Delta(G) = 3$. Então $n = 4$, $\mu_1(G) = 4$ e as únicas possibilidades para G são:*

$$G = K_4, \quad G = S_3, \quad G = \overline{K_2 + 2K_1} \quad \text{ou} \quad G = \overline{K_1 + P_3}, \quad (4)$$

onde o grafo S_3 é a estrela de 3 pontas dado por $\overline{K_1 + K_3}$.

Demonstração: Se $\mu_1(G) = 4$, vale que $\mu_1(G) = \Delta + 1$, então da Proposição 5 concluímos que existe ao menos um vértice adjacente a todos os outros vértices de G . Mas $\Delta(G) = 3 = n - 1$, donde o grafo precisa ter 4 vértices, restringindo assim as possibilidades apenas para os grafos enunciados. Por outro lado se $\mu_1(G) < 4$, então $a(\overline{G}) > n - 4$, donde temos $\delta(\overline{G}) > n - 4$, pela Proposição 2. Ou seja, $\Delta(G) < 3$, o que é contraditório.

Podemos observar que se G é um dos grafos do enunciado, então $\mu(G) = 4$ e $\Delta(G) = 3$. ■

A seguir provamos o resultado principal deste trabalho.

Teorema 1 *Seja G um grafo com n vértices. Então $a(G) = n - 4$ se, e somente se,*

$$\overline{G} = aK_4 + bS_3 + c\overline{K_2 + 2K_1} + d\overline{K_1 + P_3} + \sum_{j=1}^{n-4} e_j P_j + \sum_{j=3}^n f_j C_j, \quad (5)$$

para a, b, c, d, e_j e f_j inteiros não negativos, em que $a + b + c + d + f_j > 0$ para algum $j \geq 4$ par.

Demonstração: Vamos supor o grafo G , com n vértices e $a(G) = n - 4$. Pela Proposição 2 temos $\delta(G) \geq n - 4$. A partir da Proposição 6 garantimos que $\delta(G) = n - 4$ ou $\delta(G) = n - 3$.

CASO 1: $\delta(G) = n - 3$.

Temos que $\Delta(\overline{G}) = 2$ e $\mu_1(\overline{G}) = 4$. Supondo que o grafo $\overline{G}$ seja conexo, a limitação do grau máximo garante que $\overline{G} = P_n$ ou $\overline{G} = C_n$. Mas como $\mu_1(\overline{G}) = 4$, temos que $\overline{G} = C_n$, para $n \geq 4$ par, segundo a Tabela 1. Por outro lado, se $\overline{G}$ é desconexo, podemos escrevê-lo como $\overline{G} = \sum_{i=1}^p H_i$, em que cada H_i é uma componente conexa que possui n_i vértices.

A partir da Proposição 4 temos $\mu_1(\overline{G}) = \max \{ \mu_1(H_i), 1 \leq i \leq p \}$, então para cada H_i vale $\Delta(H_i) \leq 2$ e $\mu_1(H_i) \leq 4$ sendo que ao menos uma componente deve atingir os valores máximos. Supondo uma componente conexa H_r que atinge $\mu_1(H_r) = 4$, obtemos $\Delta(H_r) = 2$, possibilidade já analisada no parágrafo acima, daí obtemos que $H_r = C_{n_r}$, para $n_r \geq 4$ par. Para as outras componentes temos $0 \leq \Delta(H_i) \leq 2$, então H_i é isomorfo a um caminho ou ciclo, com $n_i \geq 1$. Assim, concluímos que

$$\overline{G} = \sum_{j=1}^{n-4} e_j P_j + \sum_{j=3}^n f_j C_j, \quad (6)$$

sendo que existe $j \geq 4$ par tal que $f_j > 0$.

CASO 2: $\delta(G) = n - 4$.

Temos $\Delta(\overline{G}) = 3$ e $\mu_1(\overline{G}) = 4$. Vamos supor inicialmente que $\overline{G}$ seja conexo. Pelo Lema 1, $\overline{G}$ é isomorfo a um dos seguintes grafos: $K_4, S_3, \overline{K_2 + 2K_1}$ e $\overline{K_1 + P_3}$.

Por outro lado, supondo $\overline{G}$ desconexo, seguiremos raciocínio análogo ao CASO 1. Escrevendo $\overline{G}$ como união de várias componentes conexas, obtemos p componentes H_i , de n_i vértices, onde $\mu_1(\overline{G}) = \max \{ \mu_1(H_i), 1 \leq i \leq p \}$ e para cada H_i vale que $\Delta(H_i) \leq 3$ e $\mu_1(H_i) \leq 4$ sendo que ao menos uma componente deve atingir os valores máximos.

Considere que a componente H_r satisfaz $\Delta(H_r) = 3$, tal que $\mu_1(H_r) \leq 4$. Então pelo Lema 1, obtemos H_r isomorfo ao $K_4, S_3, \overline{K_2 + 2K_1}$ ou $\overline{K_1 + P_3}$, onde ainda vale $\mu_1(H_r) = 4$.

Para as outras componentes H_i , a Tabela 1 garante:

- se $\Delta(H_i) = 0$, então H_i é isomorfo a P_1 e $\mu_1(H_i) = 0$;
- se $\Delta(H_i) = 1$, então H_i é isomorfo a P_2 e $\mu_1(H_i) = 2$;
- se $\Delta(H_i) = 2$, então H_i é isomorfo a P_{n_1} ou a C_{n_1} , com $n_1 \geq 3$, donde $\mu_1(H_i) < 4$;
- se $\Delta(H_i) = 3$, retornamos ao caso máximo $\mu_1(H_i) = \mu_1(H_r) = 4$.

Logo, combinando ambos os CASOS 1 e 2, temos:

$$\overline{G} = aK_4 + bS_3 + c\overline{K_2 + 2K_1} + d\overline{K_1 + P_3} + \sum_{j=1}^{n-4} e_j P_j + \sum_{j=3}^n f_j C_j, \quad (7)$$

para a, b, c, d, e_j e f_j inteiros, em que $a + b + c + d > 0$ (referente ao CASO 2) ou exista ao menos um $f_j > 0$ tal que $j \geq 4$ par (referente ao CASO 1).

Para a recíproca, suponha $\overline{G}$ conforme na expressão (7), com $a + b + c + d + f_j > 0$ para algum $j \geq 4$ par. Então $\overline{G}$ possui pelo uma das componentes $K_4, S_3, \overline{K_2 + 2K_1}, \overline{K_1 + P_3}, C_{2k}$ para $k \geq 2$. Enquanto que a Tabela 1, garante $\mu_1(P_i) < 4$, e

$$\mu_1(K_4) = \mu_1(S_3) = \mu_1(\overline{K_2 + 2K_1}) = \mu_1(\overline{K_1 + P_3}) = \mu_1(C_{2k}) = 4, \text{ para } k \geq 2 \quad (8)$$

Assim, pela Proposição 4, temos que

$$\mu_1(\overline{G}) = \max \{ \mu_1(K_4), \mu_1(S_3), \mu_1(\overline{K_2 + 2K_1}), \mu_1(\overline{K_1 + P_3}), \mu_1(C_j), \mu_1(P_i) \} = 4. \quad (9)$$

Logo, da Proposição 3 segue que $a(G) = n - \mu_1(\overline{G}) = n - 4$.

■

A Tabela 2 resume todos os grafos que satisfazem a condição imposta para $a(G)$, reescrevendo sob o ponto de vista de caminhos e ciclos. As três primeiras linhas são resultados obtidos de (FIEDLER, 1973), as seguintes em (DOS SANTOS, 2022), enquanto que a última é contribuição deste trabalho.

Tabela 2: Caracterizações para $a(G) \geq n - 4$

$a(G)$	Grafos
$a(G) > n$	Não existe grafo
$a(G) = n$	$\overline{G} = nP_1$
$n - 2 < a(G) < n$	Não existe grafo
$a(G) = n - 2$	$\overline{G} = aP_2 + (n - 2a)P_1$
$n - 3 < a(G) < n - 2$	Não existe grafo
$a(G) = n - 3$	$\overline{G} = aP_1 + bP_2 + cC_3 + dP_3$
$n - 3 < a(G) < n - 4$	$\overline{G} = \sum_{j=1}^A a_{2j+1} C_{2j+1} + \sum_{j=1}^B b_j P_j$
$a(G) = n - 4$	$\overline{G} = aK_4 + bS_3 + c\overline{K_2 + 2P_1} + d\overline{P_1 + P_3} + \sum_{j=1}^{n-4} e_j P_j + \sum_{j=3}^n f_j C_j$

Observe que, pela Tabela 2, todos os grafos em que $n - 3 \leq a(G) \leq n$, podem ser vistos como complementares de combinações lineares de caminhos e ciclos. Já no caso $n - 4$, além desses, aparecem também os grafos obtidos no Lema 1.

Continuando a análise da Tabela 2 vale ressaltar ainda que todos os grafos que satisfazem $n - 3 \leq a(G) \leq n$ com $a(G)$ inteiro são cografos, o que não vale para $a(G) = n - 4$, conforme evidenciado nos exemplos abaixo. A Figura 1 representa o cografo $G_1 = \overline{K_4 + S_3 + P_1}$, enquanto que a Figura 2 esboça $G_2 = \overline{C_5 + K_4}$ não-cografo, e ambos satisfazem $a(G_1) = a(G_2) = n - 4 = 5$, como no Teorema 1.

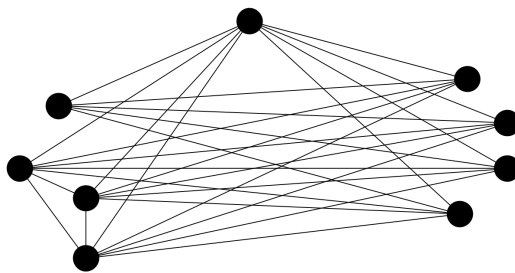


Figura 1: $G_1 = \overline{S_3 + P_1 + K_4}$.

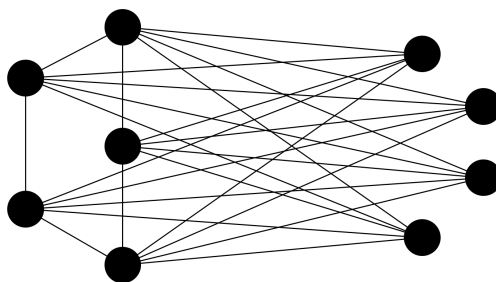


Figura 2: $G_2 = \overline{C_5 + K_4}$.

Conclusões

Este trabalho foi motivado pela importância da conectividade algébrica, não apenas para a Teoria Espectral de Grafos, mas também por suas aplicações. Apresentamos a caracterização dos grafos que atendem a maximização do invariante, mais especificamente aqueles em que $a(G) = n - 4$, permitindo observar o padrão da caracterização dos grafos com conectividade algébrica alta, sendo o $n - 4$ o primeiro caso em que o grafo não é formado apenas por combinação linear de ciclos e caminhos.

Agradecimentos

Os autores agradecem ao Instituto Federal do Sudeste de Minas Gerais (Juiz de Fora) pelo apoio financeiro a este trabalho.

Referências

- BROUWER, A. E.; HAEMERS, W. H. Spectra of graphs. **Springer Science & Business Media**, 2011.
- DE ABREU, N. M. M. Old and new results on algebraic connectivity of graphs. **Linear algebra and its applications**, v. 423, n. 1, p. 53-73, 2007.

DOS SANTOS, Y. R.; MIRANDA, L. M. S.; JONES A. A. Caracterização de grafos com conectividade algébrica máxima. In: Anais do Simpósio Brasileiro de Pesquisa Operacional. Campinas: Galoá, 2022. Disponível em: <https://proceedings.science/sbpo/sbpo-2022/trabalhos/caracterizacao-de-grafos-com-conectividade-algebrica-maxima?lang=pt-br>. Acessado em: 28 jul. 2023.

FAX, J. A.; MURRAY, R. M. Information flow and cooperative control of vehicle formations. **IEEE transactions on automatic control**, v. 49, n. 9, p. 1465-1476, 2004.

FIEDLER, M.. Algebraic connectivity of graphs. **Czechoslovak mathematical journal**, v. 23, n. 2, p. 298-305, 1973.

GRONE, R.; MERRIS, R.. The Laplacian spectrum of a graph II. **SIAM Journal on discrete mathematics**, v. 7, n. 2, p. 221-229, 1994.

JONES, A., et al. Graph Filter: software for manipulating and searching graphs. Versão 2.2.0, 17 de Maio, 2022. Disponível em <http://sistemas.jf.ifsudestemg.edu.br/graphfilter>. Acesso em: 20 jun. 2023.

KOLOKOLNIKOV, T.. Maximizing algebraic connectivity for certain families of graphs. **Linear Algebra and its Applications**, v. 471, p. 122-140, 2015.

LI, G. et al. Maximizing algebraic connectivity via minimum degree and maximum distance. **IEEE Access**, v. 6, p. 41249-41255, 2018.

MERRIS, R. Laplacian graph eigenvectors. **Linear algebra and its applications**, v. 278, n. 1-3, p. 221-236, 1998.

OGIWARA, K.; FUKAMI, T.; TAKAHASHI, N. Maximizing algebraic connectivity in the space of graphs with a fixed number of vertices and edges. **IEEE Transactions on Control of Network Systems**, v. 4, n. 2, p. 359-368, 2015.

CONSTRUÇÃO E CARACTERIZAÇÃO DE LOOPS DE CÓDIGO

Rosemary Miguel Pires
Universidade Federal Fluminense
rosemarypires@id.uff.br

Resumo: A construção de loops de código, como apresentamos neste trabalho, foi introduzida por Griess (1986). O desenvolvimento da teoria de loops de código tem bastante aplicabilidade no estudo das álgebras não associativas. Griess (1986) demonstrou que um loop de código é um loop de Moufang. Em (CHEIN; GOODAIRE, 1990) foi demonstrado que loops de código tem um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Neste trabalho, apresentamos como os loops de código são construídos. Para isto, desenvolvemos uma introdução inicial sobre códigos (lineares) pares, loops e loops de Moufang. Por fim, mostramos como determinar quando um certo loop finito pode ser caracterizado como loop de código.

Palavras-chave: Loops de Moufang. Códigos lineares pares. Loops de Código.

Introdução

Duas das quatro operações matemáticas básicas com números naturais não são associativas, o que por si só já justifica o estudo de estruturas não associativas. O surgimento de uma álgebra não associativa ocorreu no fim do século XIX, entretanto somente a partir da década de 30 do século XX, o estudo de álgebras não associativas, tais como as álgebras de Lie, as álgebras de Jordan e as álgebras alternativas, ganhou força e avançou em diversas direções.

A teoria de quasigrupos e loops também é jovem, tendo o seu início formal nos trabalhos pioneiros da matemática alemã Ruth Moufang, nos anos 30. Uma maneira intuitiva de “explicar” o que é um loop é dizer que ele é um grupo não associativo. Vale ressaltar que alguns resultados de investigações realizadas nas últimas décadas têm revelado que quasigrupos e loops possuem um importante papel em diversas especialidades da matemática, a exemplo de análise combinatória, teoria de códigos, geometria (projetiva, diferencial, hiperbólica, esférica, web), teoria de grupos (finita e de Lie) e teoria de nós. Esta teoria se enquadra no campo de estudo das Álgebras não Associativas, dentro da área de Álgebra.

A teoria de loops, em especial de loops de Moufang, é muito ampla e tem diversas aplicações. Dentre elas, uma das aplicações mais recentes é na Teoria de Loops de Código.

Griess (1986) deu início ao estudo dos Loops de Código para a construção de um grupo conhecido na literatura como “monstro”. Ele também demonstrou que um loop de código é um loop de Moufang. Os matemáticos Chein e Goodaire (1990) demonstraram que os loops de código tem uma propriedade especial: possuem um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Loops de Moufang, com uma ou mais destas propriedades de “unicidade”, têm um papel importante no estudo de loops que tem anéis de loop alternativos. Por exemplo, Chein e Goodaire provaram que um loop de Moufang, não-associativo com um único quadrado não trivial, deve ter um anel de loop alternativo sobre qualquer anel de característica dois.

Neste trabalho, apresentamos, primeiramente, alguns conceitos e resultados básicos da teoria de loops de Moufang. Também mostramos como construir loops de código e, provamos com detalhes que estes são loops de Moufang que satisfazem as propriedades de “unicidade” mencionadas acima. A seguir, mostramos que loops de código possuem um único elemento central não trivial, como consequência de uma caracterização dada em (CHEIN; GOODAIRE, 1990). Por fim, determinamos uma outra caracterização relacionada com E-loops. De modo geral, mostramos como determinar quando um certo loop finito pode ser caracterizado como loop de código.

Loops de Moufang

Nesta seção, apresentamos algumas definições e resultados preliminares sobre loops. Para mais detalhes e demonstrações veja (GOODAIRE; JASPERS; POLCINO, 1996) e (PFLUGFELDER, 1990).

Um **loop** é um conjunto L , munido com uma operação binária (denotada por justaposição ou por $\cdot$) tal que, dados quaisquer elementos x, y, z , a equação $x \cdot y = z$ determina o terceiro elemento de modo único e tal que existe um elemento identidade bilateral (denotado por 1).

Um loop L é um loop com a propriedade inversa (denotado por loop P.I.) se cada $x \in L$ tem um único inverso bilateral, que denotamos por x^{-1} , e se, para todo $x, y \in L$, o loop satisfaz $x^{-1}(xy) = y$ e $(yx)x^{-1} = y$, ou seja, as propriedades de inverso à esquerda e inverso à direita, respectivamente.

Teorema 1. *Em qualquer loop L , para quaisquer elementos x, y, z de L , as seguintes identidades (chamadas identidades de Moufang) são equivalentes:*

$$((xy)x)z = x(y(xz)), \quad (1)$$

$$((xy)z)y = x(y(zy)), \quad (2)$$

$$(xy)(zx) = (x(yz))x. \quad (3)$$

Se L é um loop satisfazendo qualquer uma dessas identidades, então L é um loop P.I. que também satisfaz

$$(yx)x = yx^2, \quad x(xy) = x^2y, \quad (xy)x = x(yx).$$

Um loop L é chamado um **loop de Moufang** se ele satisfaz qualquer uma das três identidades de Moufang.

Veremos a seguir um modo equivalente de determinar se um conjunto não vazio finito munido de uma operação binária é um loop. Para isto, precisamos definir a noção de quasigrupo. Seja G um conjunto não-vazio munido de uma operação binária “ $\cdot$ ” (denotado por $(G, \cdot)$, ou simplesmente, por G , e chamado de grupóide) e seja a qualquer elemento fixado em G . As aplicações de translação $L_a : G \rightarrow G$ e $R_a : G \rightarrow G$, são definidas respectivamente por $L_a(x) = a \cdot x$ e $R_a(x) = x \cdot a$, $\forall x \in G$.

Um grupóide $(G, \cdot)$ é chamado um quasigrupo se as aplicações $L_a : G \rightarrow G$ e $R_a : G \rightarrow G$ são bijeções para todo $a \in G$.

A definição de quasigrupo é equivalente a dizermos que dados quaisquer dois elementos de x, y, z em G o terceiro pode ser unicamente selecionado em G de modo que $x \cdot y = z$. Logo, temos que um loop $(L, \cdot)$ é um quasigrupo com elemento identidade (bilateral) 1.

Teorema 2. *Seja $(G, \cdot)$ um grupóide finito. São equivalentes:*

- (i) $(G, \cdot)$ é um quasigrupo.
- (ii) $L_a : G \rightarrow G$ e $R_a : G \rightarrow G$ são injetoras para todo $a \in G$.
- (iii) $L_a : G \rightarrow G$ e $R_a : G \rightarrow G$ são sobrejetoras para todo $a \in G$.
- (iv) As leis de cancelamento à esquerda e à direita valem para $(G, \cdot)$.
- (v) Cada elemento em G aparece somente uma vez em cada linha e em cada coluna da tabela de multiplicação para $(G, \cdot)$.

Definição 1. *Sejam x, y e z três elementos de um loop L . O comutador de x e y é o único elemento $[x, y]$ de L que satisfaz $xy = (yx)[x, y]$ e o associador de x, y e z é o único elemento (x, y, z) de L que satisfaz $(xy)z = (x(yz))(x, y, z)$.*

Seja L um loop. O núcleo à esquerda, o núcleo à direita e o núcleo do meio de L são definidos respectivamente por: $\mathcal{N}_l = \{a \in L \mid (a, x, y) = 1, \forall x, y \in L\}$, $\mathcal{N}_r = \{a \in L \mid (x, y, a) = 1, \forall x, y \in L\}$, e $\mathcal{N}_m = \{a \in L \mid (x, a, y) = 1, \forall x, y \in L\}$.

O núcleo de L é $\mathcal{N}(L) = \mathcal{N}_l \cap \mathcal{N}_r \cap \mathcal{N}_m$ e o centro de L é $\mathcal{Z}(L) = \{x \in \mathcal{N}(L) \mid [a, x] = 1, \forall a \in L\}$.

Proposição 1. *Todo núcleo de um loop é um subloop associativo e portanto um grupo. O centro de um loop é um grupo abeliano.*

As demonstrações dos próximos resultados podem ser encontradas em (CHEIN; GOODAIRE, 1990).

Lema 1. *Se L é um loop de Moufang, $z \in L$ é um elemento que comuta com todo elemento de L e $z^2 \in \mathcal{N}(L)$, então $z \in \mathcal{Z}(L)$.*

O próximo corolário será aplicado para dar uma caracterização dos loops de código (Ver Teorema 7).

Corolário 1. *Se L é um loop de Moufang, $z \in L$ é um elemento que comuta com todo elemento de L e $z^2 = 1$, então $z \in \mathcal{Z}(L)$.*

Teorema 3. *Se L é um loop de Moufang com um único quadrado não trivial, e, então $e^2 = 1$ e L é um grupo abeliano ou $[L, L] = (L, L, L) = L^2 = \{1, e\} \subseteq \mathcal{Z}(L)$.*

Teorema 4. *Se L é um loop de Moufang com $L^2 = \{1, e\}$, então, para todo $w, x, y, z \in L$,*

$$\begin{aligned} [xy, z] &= [x, z][y, z](x, y, z) \\ (wx, y, z) &= (w, y, z)(x, y, z). \end{aligned}$$

Teorema 5. (Moufang) *Se a, b, c são elementos de um loop de Moufang e $ab.c = a.bc$, o subloop gerado por $\{a, b, c\}$ é um grupo.*

Do Teorema 5, se $(x, y, z) \neq 1$, então $(x', y', z') \neq 1$ para qualquer permutação x', y', z' de x, y, z . Portanto se L tem um único quadrado não trivial e desta forma, um único associador não trivial, então $(x', y', z') = (x, y, z)$. Em outras palavras, o associador de três elementos é independente da ordem de seus elementos. Assim, a propriedade (2) do Teorema 4 também é válida para (x, wy, z) e (x, y, wz) .

Proposição 2. *Seja F um loop de Moufang.*

1. *Se $(x, y, z)^2 = 1$ e todos os comutadores e associadores de F são centrais, então*

$$[xy, z] = [x, z][y, z](x, y, z).$$

2. *Se os comutadores e associadores de F são centrais, então*

$$(wx, y, z) = (w, y, z)(x, y, z).$$

Usando a proposição anterior se prova que $(x, wy, z) = (x, w, z)(x, y, z)$ e $(x, y, wz) = (x, y, w)(x, y, z)$. De fato, observe que $[xy, z] = [z, xy]$. Daí, $(x, y, z) = (z, x, y)$. Logo, $(x, y, wz) = (wz, x, y) = (w, x, y)(z, x, y) = (y, w, x)(y, z, x) = (x, y, w)(x, y, z)$. Analogamente temos o outro caso.

Como consequência dos resultados acima, provamos o próximo resultado.

Proposição 3. *Se os quadrados e comutadores de um loop de Moufang F são centrais, então*

$$(xy)^2 = x^2 y^2 [x, y].$$

Loops de Código: construção e algumas propriedades

Nesta seção apresentamos uma introdução à Teoria de Loops de Código. Para isto, consideremos $\mathbb{F}_2^n$ como um espaço vetorial n -dimensional sobre o corpo de 2 elementos $\mathbb{F}_2 = \{0, 1\}$. Para vetores u e v em $\mathbb{F}_2^n$, $|v|$ (o peso de v) denota o número de coordenadas não-nulas de v e $|u \cap v|$ (peso de u intersecção v) denota o número de posições nas quais as coordenadas de u e v são ambas não-nulas, ou seja, se $u = (u_1, \dots, u_n)$ e $v = (v_1, \dots, v_n)$, com $u_i, v_i \in \mathbb{F}_2$ para todo $i = 1, \dots, n$, então $|v| = |\{i \mid v_i = 1\}|$ e $|u \cap v| = |\{i \mid u_i = v_i = 1\}|$.

Definição 2. Um código par é um subespaço $V \subseteq \mathbb{F}_2^n$ tal que $|v| \equiv 0 \pmod{4}$ e $|u \cap v| \equiv 0 \pmod{2}$ para quaisquer $u, v \in V$.

Como exemplo de código par, podemos considerar o subespaço $V \subseteq \mathbb{F}_2^7$ com base $B = \{v_1, v_2, v_3\}$ onde

$$v_1 = (1, 1, 1, 1, 0, 0, 0), \quad v_2 = (1, 1, 0, 0, 1, 1, 0), \quad v_3 = (1, 0, 1, 0, 1, 0, 1).$$

Definição 3. Seja V um código par. Uma função $\phi : V \times V \rightarrow \{1, -1\}$ definida para todo $u, v, w \in V$, da forma:

$$\phi(v, v) = (-1)^{\frac{|v|}{4}}; \quad (4)$$

$$\phi(v, w) = (-1)^{\frac{|v \cap w|}{2}} \phi(w, v); \quad (5)$$

$$\phi(0, v) = \phi(v, 0) = 1; \quad (6)$$

$$\phi(v + w, u) = \phi(v, w + u) \phi(v, w) \phi(w, u) (-1)^{|v \cap w \cap u|}. \quad (7)$$

é chamada “factor set”.

Agora denotamos por $L(V)$ o conjunto $V \cup (-V)$. Vamos definir uma operação binária “.” sobre $L(V)$, chamada de produto e denotada, quando conveniente, por justaposição. Sejam $v, w \in V$, definimos

$$\begin{aligned} v.w &= \phi(v, w)(v + w), \text{ onde } \phi(v, w) \in \{1, -1\}, \\ v.(-w) &= (-v).w = -(v.w), \\ (-v).(-w) &= v.w. \end{aligned} \quad (8)$$

Com o produto definido acima, se prova que $(L(V), .)$, ou simplesmente $L(V)$, tem uma estrutura de loop. Loops construídos desta forma, são chamados de **loops de código**. Como a dimensão do código par V é finita, o loop de código também é finito. Logo, definimos o posto de $L(V)$ como sendo a dimensão de V .

Como exemplo, considere V o código par do exemplo anterior e $L(V) = V \cup (-V)$, com a operação de multiplicação definida como acima. Claramente temos que $L(V)$ é um loop de código de posto 3 com 16 elementos. Existem, a menos de isomorfismo, apenas 5 loops de código de posto 3 (PIRES, 2011).

Pretendemos provar agora que $L(V)$ é de fato um loop. Definimos que dois elementos x, y de $L(V)$ da forma $x = au$ e $y = bv$, com $a, b \in \{1, -1\}$, são iguais se $a = b$ e $u = v$. Sempre que denotarmos um elemento de $L(V)$ da forma $x = av$ significa que $a \in \{1, -1\}$ e $v \in V$. Podemos representar o produto de dois elementos quaisquer $x = au$ e $y = bv$ de $L(V)$ da forma $xy = (ab\phi(u, v))(u + v)$.

Seja $L = L(V)$. Como L é um conjunto finito, podemos aplicar o Teorema 2 para provar que L é loop. Para isto, basta provarmos que para todo $a \in L$ que as aplicações de translação $L_a : L \rightarrow L$ e $R_a : L \rightarrow L$, definidas respectivamente por $L_a(x) = a.x$ e $R_a(x) = x.a$, para todo $x \in L$ são injetoras e que L tem elemento identidade.

Com efeito, sejam $x = au \in L$ e as aplicações de translação $L_x : L \rightarrow L$ e $R_x : L \rightarrow L$ definidas como antes. Sejam $y = bv$ e $z = cw$ elementos de L tais que $L_x(y) = L_x(z)$. Pela definição de L_x obtemos $xy = xz$, ou seja, $(au)(bv) = (au)(cw)$. Pela definição do produto em L temos que $(ab\phi(u, v))(u + v) =$

$(ac\phi(u, w))(u + w)$. Desta igualdade, vemos que $ab\phi(u, v) = ac\phi(u, w)$ e que $u + v = u + w$, o que nos dá $v = w$ e $\phi(u, v) = \phi(u, w)$, e, finalmente $b = c$. Logo, $y = bv = cw = z$, como queríamos. Portanto, L_x é injetora, para todo $x \in L$. Analogamente, prova-se que R_x é injetora, para todo $x \in L$. Agora, 0 é o elemento identidade de L , pois $0.v = \phi(0, v)(0 + v) = 1v = v$ e $0.(-v) = -0.v = -1v = -v$.

Consideramos através do lema abaixo algumas propriedades relativas ao código par V . Estas propriedades são obtidas com base num resultado mais geral apresentado no artigo (GRIESS, 1986, Lema 5, p. 225).

Lema 2. *Sejam x, y, z quaisquer elementos de um código par V . São válidas as seguintes propriedades:*

$$|x \cap y \cap (z + x)| \equiv |x \cap y \cap z| \pmod{2}; \quad (9)$$

$$\frac{1}{2}|y \cap (x + z)| \equiv \frac{1}{2}|y \cap x| + \frac{1}{2}|y \cap z| + |y \cap x \cap z| \pmod{2}; \quad (10)$$

$$\frac{1}{2}|x \cap (x + y + z)| \equiv \frac{1}{2}|x \cap y| + \frac{1}{2}|x \cap z| + |x \cap y \cap z| \pmod{2}. \quad (11)$$

Além de mostrar a existência de loops de código em seu trabalho, Griess (1986) também demonstrou que um loop de código é um loop de Moufang. Além disso, Chein e Goodaire (1990) demonstraram que um loop de código satisfaz uma propriedade bem especial relacionada aos quadrados, comutadores e associadores de seus elementos. Em outras palavras, em um loop de código existe um único quadrado não trivial, um único comutador e um único associador não trivial. Este resultado é apresentado no teorema seguinte. No trabalho destes autores foi introduzida apenas uma ideia da demonstração. Aqui, apresentamos uma demonstração incluindo todos os detalhes necessários para o entendimento.

Teorema 6. *O loop $L(V)$ é loop de Moufang e valem as seguintes propriedades, para quaisquer $u, v, w \in V$:*

$$v^2 = (-1)^{\frac{|v|}{4}} 0, \quad (12)$$

$$[u, v] = u^{-1}v^{-1}uv = (-1)^{\frac{|u \cap v|}{2}} 0, \quad (13)$$

$$(u, v, w) = ((uv)w)((u(vw))^{-1}) = (-1)^{|u \cap v \cap w|} 0. \quad (14)$$

Demonstração. Seja $L(V) = \{1, -1\} \times V$, onde V é um código par. Temos que provar que $L(V)$ satisfaz qualquer uma das identidades de Moufang. Vamos provar que vale a Identidade 3 do Teorema 1:

$$(ax)(by).(cz)(ax) = [(ax).(by)(cz)](ax), \quad \forall ax, by, cz \in L(V) \quad (15)$$

Pela definição do produto de $L(V)$, para verificar esta identidade de Moufang basta provarmos que

$$\phi(x, y)\phi(z, x)\phi(x + y, z + x) = \phi(y, z)\phi(x, y + z)\phi(x + y + z, x), \quad (16)$$

onde $\phi : V \times V \rightarrow \{1, -1\}$ é a função *factor set* (conforme Definição 3).

De fato, por um lado

$$(ax)(by).(cz)(ax) = (abac\phi(x, y)\phi(z, x)\phi(x + y, z + x))(y + z).$$

Por outro lado,

$$[(ax).(by)(cz)](ax) = (abac\phi(y, z)\phi(x, y + z)\phi(x + y + z, x))(y + z).$$

Agora, para demonstrar a Equação (16), primeiramente, observamos que pelas Equações (7) e (9),

$$\phi(x, y)\phi(x + y, z + x) = \phi(y, z + x)\phi(x, x + y + z)(-1)^{|x \cap y \cap z|}. \quad (17)$$

A partir da Equação (5), obtemos: $\phi(x, x+y+z) = \phi(x+y+z, x)(-1)^{|x \cap (x+y+z)|}$. Usando este resultado juntamente com a Equação (11), podemos reescrever a Equação (17) da forma:

$$\phi(x, y)\phi(x+y, z+x) = \phi(y, z+x)\phi(x+y+z, x)(-1)^{\frac{|x \cap y|}{2} + \frac{|x \cap z|}{2}}. \quad (18)$$

Multiplicando cada termo da última equação por $\phi(z, x)$, obtemos o seguinte resultado:

$$\phi(x, y)\phi(z, x)\phi(x+y, z+x) = \phi(y, z+x)\phi(z, x)\phi(x+y+z, x)\phi(z, x)(-1)^{\frac{|x \cap y|}{2} + \frac{|x \cap z|}{2}}. \quad (19)$$

Agora, observemos que a partir das Equações (5) e (7) obtemos:

$$\begin{aligned} \phi(x, z)\phi(x, z+y) &= \phi(x+z, y)\phi(z, y)(-1)^{|x \cap y \cap z|}; \\ \phi(y, x+z) &= \phi(x+z, y)(-1)^{\frac{x \cap y}{2} + \frac{y \cap z}{2} + |x \cap y \cap z|}. \end{aligned}$$

Assim encontramos a seguinte equação:

$$\phi(z, x)\phi(y, z)\phi(x, y+z)\phi(y, z+x)(-1)^{\frac{|x \cap y|}{2} + \frac{|x \cap z|}{2}} = 1. \quad (20)$$

Para concluirmos esta primeira parte da demonstração do teorema, basta ver que a Equação (20) é equivalente a equação seguinte:

$$\phi(y, z+x)\phi(z, x)\phi(x+y+z, x)\phi(z, x)(-1)^{\frac{|x \cap y|}{2} + \frac{|x \cap z|}{2}} = \phi(y, z)\phi(x, y+z)\phi(x+y+z, x). \quad (21)$$

Logo, a Equação (16) segue diretamente das Equações (19) e (21) e, portanto, $L(V)$ é um loop de Moufang. Agora iremos provar as propriedades de “unicidade” relacionadas ao quadrado (comutador, associador) enunciadas no Teorema.

Propriedade do quadrado: Como $v^2 = \phi(v, v)(v+v)$, e por definição $\phi(v, v) = (-1)^{\frac{|v|}{4}}$, então $v^2 = (-1)^{\frac{|v|}{4}}0$.

Propriedade do comutador: Por definição,

$$\begin{aligned} u.v &= \phi(u, v)(u+v) = (-1)^{\frac{|u \cap v|}{2}}\phi(v, u)(v+u) = \\ &= (-1)^{\frac{|u \cap v|}{2}}0.\phi(v, u)(v+u) = (-1)^{\frac{|u \cap v|}{2}}0.(v.u). \end{aligned}$$

Desta forma, pela unicidade do comutador, obtemos que $[u, v] = (-1)^{\frac{|u \cap v|}{2}}0$.

Propriedade do associador: Por definição, por um lado

$$(uv)w = \phi(u, v)\phi(u+v, w)(u+v+w) = (-1)^{|u \cap v \cap w|}\phi(v, w)\phi(u, v+w)(u+v+w).$$

Por outro lado,

$$(-1)^{|u \cap v \cap w|}0.(u(vw)) = (-1)^{|u \cap v \cap w|}0.(\phi(v, w)\phi(u, v+w)(u+v+w)),$$

ou seja,

$$(-1)^{|u \cap v \cap w|}0.(u(vw)) = (-1)^{|u \cap v \cap w|}\phi(v, w)\phi(u, v+w)\phi(0, u+v+w)(u+v+w).$$

Logo,

$$(-1)^{|u \cap v \cap w|}0.(u(vw)) = (uv)w.$$

Desta forma, pela unicidade do associador, concluímos que $(u, v, w) = (-1)^{|u \cap v \cap w|}0$.

□

Caracterizações de loops de código

Definição 4. Um loop de Moufang L é chamado *E-loop* se existe um subloop central Z de 2 elementos tal que $L/Z \in A$, onde A é a variedade de grupos com identidade $x^2 = 1$.

Teorema 7. Seja L um loop de Moufang. O loop L é um *E-loop* se, e somente se, $|L^2| \leq 2$.

Demonstração. Seja L um *E-loop*. Então existe um subloop central $Z = \{1, x\}$ tal que $\bar{y}^2 = \bar{1}$, para todo $\bar{y} \in L/Z$. Vamos mostrar que $L^2 \subseteq Z$. Se supormos que $e \in L^2$ seja tal que $e \neq 1$ e $e \neq x$, então existirá $u \in L$ tal que $u^2 = e$. Daí teremos $\bar{u}^2 \neq \bar{1}$, contradizendo a definição de L . Logo $u^2 = 1$ ou $u^2 = x$, ou seja $L^2 \subseteq Z$. Portanto, $|L^2| \leq 2$. Reciprocamente, suponhamos que $|L^2| \leq 2$. Caso $|L^2| = 1$, então L é um loop de Moufang comutativo. De fato, para quaisquer $x, y \in L$, $[x, y] = x^{-2}(xy^{-1})^2y^2 = 1$. Logo, pelo Corolário 1, qualquer elemento de L está no centro de L . Considere então um subloop central qualquer de dois elementos, digamos $Z = \{1, x\}$, onde $x \in L$ é qualquer. Assim, $L/Z \in A$, onde A é a variedade de grupos tais que $y^2 = 1$, para todo y . Agora suponhamos que $L^2 = \{1, e\}$ e que $L^2 = Z$. Pelo Teorema 3, L^2 é um subloop central e assim, podemos ver com simples cálculos que $\bar{y}^2 = \bar{1}$, para todo $\bar{y} \in L/Z$. $\square$

Teorema 8. (CHEIN; GOODAIRE, 1990) Um loop finito L é isomorfo a um loop de código se e somente se L é um loop de Moufang com $|L^2| \leq 2$.

Como consequência desta caracterização e do Teorema 3, observamos que se L é um loop de código não associativo então $|\mathcal{Z}(L)| = 2$.

Finalizamos este trabalho dando uma outra caracterização dos loops de código. A demonstração segue diretamente do Teorema 7 e do Teorema 8.

Teorema 9. Um loop de Moufang L é um loop de código se, e somente se, L é um *E-loop* finito.

Vários dos resultados deste trabalho são aplicados no estudo dos problemas de classificação e representação de loops de código. Para mais detalhes, veja os trabalhos de (PIRES, 2011) e (GRISHKOV; PIRES, 2018).

Conclusões

Neste trabalho, apresentamos uma introdução e construção de loops de código. Provamos com detalhes que loops de código são loops de Moufang com um único comutador não trivial, um único associador não trivial e um único quadrado não trivial. Por fim, exibimos uma caracterização destes loops dada em (CHEIN; GOODAIRE, 1990) e, provamos uma outra caracterização relacionada com os *E-loops*.

Referências

- CHEIN, O.; GOODAIRE, E. G. Moufang Loops with a Unique Nonidentity Commutator (Associator, Square). **J. Algebra** 130, 369-384. 1990.
- GOODAIRE, E. G.; JASPERS, E.; POLCINO, C. **Alternative Loop Rings**. North Holland Math, Studies N.184, Elsevier, Amsterdam, 1996.
- GRIESS, R. L. Code loops. **J. Algebra** 100, 224-234, 1986.
- GRISHKOV A.; PIRES, R. M. Variety of loops generated by code loops. **International Journal of Algebra and Computation**, Volume 28, No. 1, 163-177, 2018.
- PFLUGFELDER, H. O. **Quasigroups and Loops: An Introduction**. Berlin, 1990.
- PIRES, R. M. **Loops de código: automorfismos e representações**. 2011. Tese (Doutorado em Matemática) – IME/USP, São Paulo, 2011.

Número Cromático Antimágico Local dos Grafos *Fireflies*

Lara Rodrigues Ventura
Universidade Federal Fluminense
laraventura@id.uff.br

André Ebling Brondani
Universidade Federal Fluminense
andrebrondani@id.uff.br

Francisca Andrea Macedo Franca
Universidade Federal Fluminense
francisca_franca@id.uff.br

Resumo: Uma rotulação de arestas de um grafo conexo $G = (V, E)$, $f : E \rightarrow \{1, 2, \dots, |E|\}$, é dita antimágica local se é uma bijeção tal que para qualquer par de vértices adjacentes u e v , $w(u) \neq w(v)$, onde $w(u) = \sum f(e)$ com e variando sobre todas as arestas incidentes a u . O número cromático antimágico local de G é o número mínimo de rotulações de vértices distintas, induzidas sobre todas as rotulações antimágicas locais de G . Neste trabalho, determinamos o número cromático antimágico local da classe dos grafos *fireflies*.

Palavras-chave: Grafo *firefly*. Rotulação antimágica local. Número cromático antimágico local.

Introdução

A Teoria de Grafos é usada em vários campos, como modelagem de redes, *design* de banco de dados, agendamentos, problemas do caixeiro viajante, etc. (ELUMALAI, 2020). A rotulação de grafos é um dos tópicos de pesquisa em teoria de grafos que associa um elemento do grafo, como vértices ou arestas, a números inteiros denominados rótulos.

Sejam $G = (V, E)$ um grafo conexo simples e $f : E \rightarrow \{1, 2, \dots, |E|\}$ uma bijeção. Para cada $v \in V$, o peso de v é dado por $w(v) = \sum_{e \in E(v)} f(e)$, onde $E(v)$ denota o conjunto das arestas que incidem em v . Se $w(v) \neq w(u)$ para todo par de vértices distintos $v, u \in V$, então f é denominada rotulação antimágica de G . Se uma tal rotulação existir, G é dito antimágico. Hartsfield e Ringel introduziram o conceito de rotulação antimágica de um grafo em Hartsfield e Ringel (1990) e conjecturaram que todo grafo conexo, com exceção do grafo completo K_2 , é antimágico. Desde então, a conjectura tem recebido muita atenção, e foi provada para várias classes especiais de grafos. Entretanto, a conjectura ainda não foi resolvida, mesmo para algumas classes de grafos particularmente simples, como árvores. Para obter mais detalhes sobre classes conhecidas de grafos antimágicos, consulte Gallian (1997).

Em 2017, dois grupos de pesquisadores, independentemente, consideraram uma condição mais fraca de rótulos que distingue vizinhos, ao invés de rótulos que distinguem vértices. De acordo com Arumugan *et al.* (2017) e Bensmail *et al.* (2017), se $w(v) \neq w(u)$ para todo par de vértices adjacentes em G , então f é uma rotulação antimágica local de G . Ambos os grupos conjecturaram que todo grafo conexo, exceto K_2 , é antimágico local. Tal conjectura foi provada em (HASLEGRAVE, 2018) usando métodos probabilísticos. Além disso, Arumugan *et al.* (2017) definiram o número cromático antimágico local para um grafo. Tal invariante já foi estudado em algumas classes de grafos, veja (NAZULA; SLAMIN; DAFIK, 2018) e (SETHURAMAN; SHERMILY, 2021). Motivados por tal definição, neste trabalho, apresentamos uma rotulação antimágica local e o número cromático antimágico local para a classe dos grafos *fireflies*, que contém subclasses importantes de grafos tais como o grafo estrela, o grafo da amizade e o grafo borboleta. O texto está organizado com os resultados preliminares na próxima seção e os resultados obtidos na seção seguinte.

Preliminares

Nesta seção, são apresentados alguns conceitos e resultados básicos da Teoria de Grafos e algumas definições e teoremas relacionados ao número cromático antimágico local que são necessários para compreensão do texto. Para mais detalhes, consultar (MERRIS, 2001) e (HARTSFIELD; RINGEL, 2002).

Seja $G = (V, E)$ um grafo simples, com conjunto de vértices $V = \{v_1, v_2, \dots, v_n\}$ e conjunto de arestas $E = \{v_i v_j; i \neq j, v_i, v_j \in V\}$. Indicamos por $|V|$ e $|E|$, respectivamente, a *ordem* e o *tamanho* de G . Dois vértices distintos, v_i e v_j em V , são ditos *vértices adjacentes* se $v_i v_j \in E$ (e neste caso, diz-se que a aresta $v_i v_j$ *incide* em v_i), a vizinhança de v_i é o conjunto $N_G(v_i) = \{v_j \in V; v_i v_j \in E\}$ e o grau do vértice v_i é dado por $d(v_i) = |N_G(v_i)|$. Os vértices de grau 1 de G , quando existem, são ditos *vértices pendentes* e uma aresta que incide em um vértice pendente é dita *aresta pendente*. Uma sequência finita $(v_1, v_2, \dots, v_k)$ de vértices de G é dita uma *cadeia* de v_1 a v_k quando $v_i v_{i+1} \in E$, onde $1 \leq i \leq k-1$. A cadeia é dita *fechada* quando $v_1 = v_k$, caso contrário, a cadeia é *aberta*; e a cadeia é dita *simples* quando $v_i \neq v_j$, onde $2 \leq i, j \leq k$. Um *caminho* é uma cadeia simples e aberta e um *ciclo* é uma cadeia simples e fechada. Um *subgrafo* de G é um grafo $H = (W, F)$ tal que $W \subseteq V$ e $F \subseteq E$. Um grafo é dito *conexo* quando seu conjunto de vértices é unitário ou existe um caminho ligando cada par de vértices de G . Caso contrário, o grafo é denominado *desconexo*. Um grafo conexo e sem ciclos é denominado *árvore* e as suas *folhas* são os seus vértices pendentes.

Dados $G = (V, E)$ e $u \in V$, denotemos por $E(u)$ ao conjunto formado pelas arestas de G que incidem em u . Hartsfield e Ringel (2003) introduziram o conceito de rotulação antimágica de um grafo de acordo com a definição a seguir.

Definição 1. Sejam $G = (V, E)$ um grafo e $f : E \rightarrow \{1, 2, \dots, |E|\}$ uma bijeção. Para cada vértice $u \in V$, o peso de u é dado por $w(u) = \sum_{e \in E(u)} f(e)$. Quando $w(u) \neq w(v)$ para quaisquer dois vértices distintos $u, v \in V(G)$, diz-se que f é uma *rotulação antimágica* de G . Um grafo é chamado *antimágico* se tem uma rotulação antimágica.

Os conceitos de rotulação antimágica local e número cromático antimágico local de um grafo foram introduzidos por Arumugan *et al.* (2017), de acordo com as próximas definições.

Definição 2. Seja $G = (V, E)$ um grafo de ordem n e tamanho m sem vértices isolados. Uma *rotulação antimágica local* de um grafo G é uma bijeção $f : E \rightarrow \{1, 2, \dots, m\}$ de modo que os pesos de quaisquer dois vértices adjacentes u e v são distintos. Um grafo é chamado *antimágico local* se tem uma rotulação antimágica local.

Definição 3. Seja G um grafo de ordem n . O *número cromático antimágico local* de G , denotado por $\chi_{la}(G)$, é o número mínimo de rotulações de vértices distintas induzidas sobre todas as rotulações antimágicas locais de G .

Para todo grafo G , o número cromático de G , $\chi(G)$, se relaciona com o número cromático antimágico local de G pela desigualdade $\chi_{la}(G) \geq \chi(G)$ e a diferença $\chi_{la}(G) - \chi(G)$ pode ser arbitrariamente grande, como mostra o teorema a seguir, devido a Arumugan *et al.* (2017).

Teorema 1. *Seja T uma árvore de ordem $n \geq 3$. Se T possui ℓ folhas, então $\chi_{la}(T) \geq \ell + 1$*

Observação 1. *O Teorema 1 foi generalizado em (BACA et al., 2021), onde os autores mostraram que a mesma desigualdade se verifica para grafos de ordem $n \geq 3$ com ℓ vértices pendentes.*

Resultados obtidos

Nesta seção são apresentados os resultados obtidos sobre número cromático antimágico local dos grafos *fireflies*. De acordo com Aouchiche *et al.* (2011) um *grafo firefly*, denotado por $F_{r,s,t}$, é um grafo

de ordem $n = 2r + s + 2t + 1$, que consiste de r triângulos, s arestas pendentes e t caminhos pendentes de comprimento 2, todos compartilhando um vértice em comum. Denotemos por $\mathcal{F}_n$ o conjunto de todos os grafos *Fireflies* com n vértices. Note que $\mathcal{F}_n$ contém a estrela, $F_{0,s,0}$, o grafo *stretched stars*, $F_{0,s,t}$, o grafo da amizade, $F_{r,0,0}$, e o grafo borboleta, $F_{r,s,0}$. Além disso, para cada $n \geq 2$ temos

$$\mathcal{F}_n = \{F_{r,0,0}, F_{0,s,0}, F_{0,0,t}; r, s, t \geq 1\} \cup \{F_{r,s,0}, F_{r,0,t}, F_{0,s,t}; r, s, t \geq 1\} \cup \{F_{r,s,t}; r, s, t \geq 1\}.$$

Em (ARUMUGAN *et al.*, 2017) é determinado o número cromático antimágico local de várias classes de grafos tais como caminhos, ciclos, certos grafos bipartidos completos e os grafos roda. Em particular, os autores mostram que $\chi_{la}(F_{r,0,0}) = 3$, $\chi_{la}(F_{0,s,0}) = s + 1$ e $\chi_{la}(F_{0,1,t}) = t + 2$. A seguir, determinamos o número cromático antimágico local para os demais grafos em $\mathcal{F}_n$.

Teorema 2. *Se s e t são inteiros positivos, então o número cromático antimágico local do grafo $F_{0,s,t}$ é $\chi_{la}(F_{0,s,t}) = s + t + 1$.*

Demonstração. Sejam $V(F_{0,s,t}) = \{c\} \cup \{v_i | 1 \leq i \leq t\} \cup \{u_i | 1 \leq i \leq t\} \cup \{w_i | 1 \leq i \leq s\}$ o conjunto de vértices e $E(F_{0,s,t}) = \{cv_i | 1 \leq i \leq t\} \cup \{cw_i | 1 \leq i \leq s\} \cup \{v_i u_i | 1 \leq i \leq t\}$ o conjunto de arestas de $F_{0,s,t}$. Defina a rotulação das arestas $f : E(F_{0,s,t}) \rightarrow \{1, 2, \dots, |V| - 1\}$ por

$$\begin{aligned} f(v_i u_i) &= i, & \text{se } 1 \leq i \leq t; \\ f(cv_i) &= 2t + 1 - i, & \text{se } 1 \leq i \leq t; \\ f(cw_i) &= 2t + i, & \text{se } 1 \leq i \leq s. \end{aligned}$$

Note que f é uma rotulação antimágica local de $F_{0,s,t}$ e a rotulação de seus vértices, induzida por f , é dada por

$$\begin{aligned} w(u_i) &= i, & \text{se } 1 \leq i \leq t; \\ w(v_i) &= 2t + 1, & \text{se } 1 \leq i \leq t; \\ w(w_i) &= 2t + i, & \text{se } 1 \leq i \leq s; \\ w(c) &= \frac{(s+t)(s+3t+1)}{2}. \end{aligned}$$

Como $w(w_1) = 2t + 1$ obtemos $\chi_{la}(F_{0,s,t}) \leq s + t + 1$ e o resultado segue do Teorema 1. □

Na Figura 1, mostramos a rotulação antimágica local do grafo $F_{0,4,4}$ e seu respectivo número cromático antimágico local.

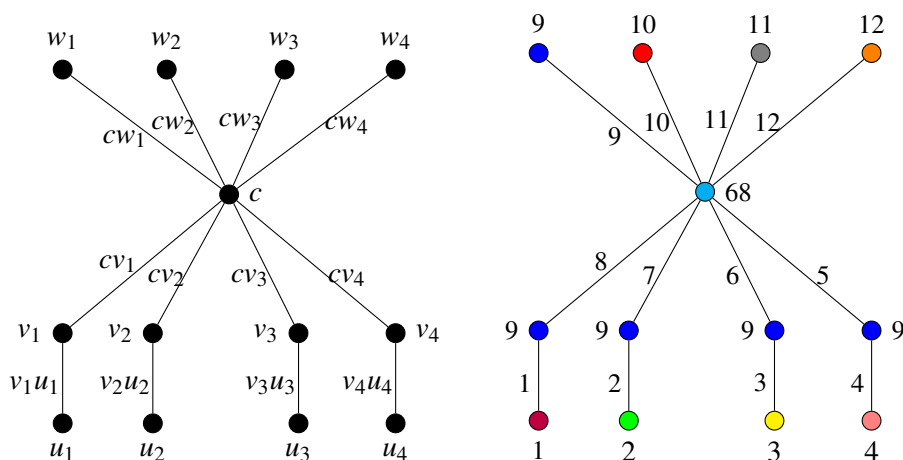


Figura 1: Grafo $F_{0,4,4}$ com $\chi_{la}(F_{0,4,4}) = 9$

Teorema 3. Dado o inteiro $t > 1$, o número cromático antimágico local do grafo $F_{0,0,t}$ é dado por

$$\chi_{la}(F_{0,0,t}) = \begin{cases} t+1, & \text{se } t \in \{2, 3\}; \\ t+2, & \text{se } t \geq 4. \end{cases}$$

Demonstração. Sejam $V(F_{0,0,t}) = \{c\} \cup \{v_i | 1 \leq i \leq t\} \cup \{u_i | 1 \leq i \leq t\}$ o conjunto de vértices e $E(F_{0,0,t}) = \{cv_i | 1 \leq i \leq t\} \cup \{v_i u_i | 1 \leq i \leq t\}$ o conjunto de arestas de $F_{0,0,t}$. Defina a rotulação de arestas $f : E(F_{0,0,t}) \rightarrow \{1, 2, \dots, 2t\}$ por

$$\begin{aligned} f(cv_i) &= i & \text{se } 1 \leq i \leq t; \\ f(v_i u_i) &= 2t+1-i & \text{se } 1 \leq i \leq t. \end{aligned}$$

Note que f é uma rotulação antimágica local de $F_{0,0,t}$ e a rotulação de seus vértices, induzida por f , é dada por

$$\begin{aligned} w(v_i) &= 2t+1, & \text{se } 1 \leq i \leq t; \\ w(u_i) &= 2t+1-i, & \text{se } 1 \leq i \leq t; \\ w(c) &= \frac{t(t+1)}{2}. \end{aligned}$$

Observe que, para $t \in \{2, 3\}$, $\chi_{la}(F_{0,0,t}) \leq t+1$. Do Teorema 1, $\chi_{la}(F_{0,0,t}) = t+1$ para $t \in \{2, 3\}$.

Para $t \geq 4$ obtemos $\chi_{la}(F_{0,0,t}) \leq t+2$, pois $w(c) > 2t+1$. Agora, como $F_{0,0,t}$ possui $2t$ arestas e $t \geq 4$, para qualquer outra rotulação antimágica local de $F_{0,0,t}$ tem-se

$$w(c) = \sum_{k=1}^t f(v_k c) \geq \frac{t(t+1)}{2} > 2t \geq f(v_i u_i) = w(u_i), \quad 1 \leq i \leq t. \quad (1)$$

Como $E(F_{0,0,t}) = \{v_i u_i, v_i c; 1 \leq i \leq t\}$, da desigualdade (1) concluímos que $\chi_{la}(F_{0,0,t}) \geq t+2$, o que conclui a demonstração do teorema. $\square$

Na Figura 2 exibimos a rotulação antimágica local do grafo $F_{0,0,5}$ e seu respectivo número cromático antimágico local.

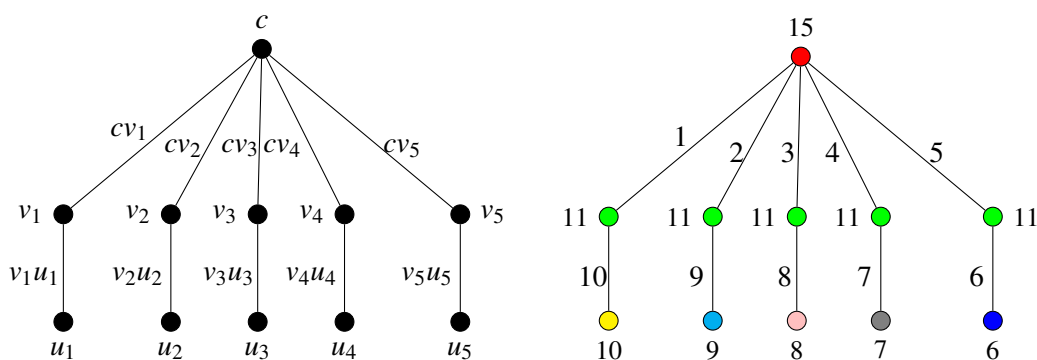


Figura 2: Grafo $F_{0,0,5}$ com $\chi_{la}(F_{0,0,5}) = 7$.

Teorema 4. Se r e s são inteiros positivos, então o número cromático antimágico local do grafo $F_{r,s,0}$ é dado por

$$\chi_{la}(F_{r,s,0}) = \begin{cases} 3, & \text{se } s = 1; \\ s+1, & \text{se } s \geq 2. \end{cases}$$

Demonstração. Sejam $V(F_{r,s,0}) = \{u_i : 1 \leq i \leq r\} \cup \{v_i : 1 \leq i \leq r\} \cup \{w_i : 1 \leq i \leq s\} \cup \{c\}$ o conjunto de vértices e $E(F_{r,s,0}) = \{u_i v_i : 1 \leq i \leq r\} \cup \{cu_i : 1 \leq i \leq r\} \cup \{cv_i : 1 \leq i \leq r\} \cup \{cw_i : 1 \leq i \leq s\}$ o conjunto de arestas de $F_{r,s,0}$. Como $\chi(F_{r,s,0}) = 3$, segue que $\chi_{la}(F_{r,s,0}) \geq 3$.

Considere a rotulação de arestas $f : E(F_{r,s,0}) \rightarrow \{1, 2, \dots, 3r + s\}$ dada por

$$\begin{aligned} f(w_1 c) &= 2r + 1, \\ f(w_i c) &= 3r + i, & \text{onde } 2 \leq i \leq s; \\ f(u_i v_i) &= i, & \text{onde } 1 \leq i \leq r; \\ f(cu_i) &= 2r + 1 - i, & \text{onde } 1 \leq i \leq r; \\ f(cv_i) &= 3r + 2 - i, & \text{onde } 1 \leq i \leq r, \end{aligned}$$

Note que f é uma rotulação antimágica local de $F_{r,s,0}$ e a rotulação de seus vértices, induzida por f , é dada por

$$\begin{aligned} w(w_1) &= 2r + 1, \\ w(w_i) &= 3r + i, & \text{onde } 2 \leq i \leq s; \\ w(u_i) &= 2r + 1, \\ w(v_i) &= 3r + 2, \\ w(c) &= (2r + 1)^2 + \frac{(s-1)}{2}(6r + s + 2). \end{aligned}$$

Logo, $\chi_{la}(F_{r,1,0}) = 3$. Para o caso em que $s \geq 2$, obtemos $\chi_{la}(F_{r,s,0}) \leq s + 1$ e o resultado segue da Observação 1. □

Na Figura 3, mostramos a rotulação antimágica local do grafo $F_{2,4,0}$ e seu respectivo número cromático antimágico local.

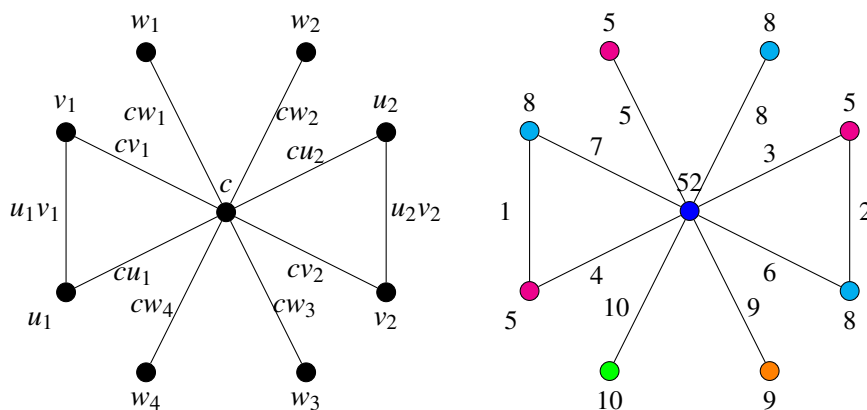


Figura 3: Grafo $F_{2,4,0}$ com $\chi_{la}(F_{2,4,0}) = 5$

Teorema 5. Se r e t são inteiros positivos, então o número cromático antimágico local do grafo $F_{r,0,t}$ é dado por

$$\chi_{la}(F_{r,0,t}) = \begin{cases} t + 3, & \text{se } t < r; \\ t + 2, & \text{se } t \geq r. \end{cases}$$

Demonstração. Considere os conjuntos de vértices e arestas de $F_{r,0,t}$ dados, respectivamente, por

$$\begin{aligned} V(F_{r,0,t}) &= \{u_i : 1 \leq i \leq t\} \cup \{v_i : 1 \leq i \leq t\} \cup \{w_i : 1 \leq i \leq r\} \cup \{z_i : 1 \leq i \leq r\} \cup \{c\} \\ E(F_{r,0,t}) &= \{u_i v_i : 1 \leq i \leq t\} \cup \{w_i z_i : 1 \leq i \leq r\} \cup \{c v_i : 1 \leq i \leq t\} \cup \{c w_i : 1 \leq i \leq r\} \cup \{c z_i : 1 \leq i \leq r\}. \end{aligned}$$

Seja g uma rotulação antimágica local qualquer de $F_{r,0,t}$. Como g é uma bijeção e $F_{r,0,t}$ tem $3r + 2t$ arestas temos $w(c) \geq \sum_{k=1}^{2r+t} k > 3r + 2t$. Assim, $w(u_i) < w(v_i) < w(c)$ para todo i , $1 \leq i \leq t$, e portanto

$$\chi_{la}(F_{r,0,t}) \geq t + 2. \quad (2)$$

Definamos a rotulação de arestas $f : E(F_{r,0,t}) \rightarrow \{1, 2, \dots, 3r + 2t\}$ por

$$\begin{aligned} f(cv_i) &= i, & 1 \leq i \leq t; \\ f(w_i z_i) &= t + i, & 1 \leq i \leq r; \\ f(cw_i) &= 2r + t + 1 - i, & 1 \leq i \leq r; \\ f(cz_i) &= 3r + t + 1 - i, & 1 \leq i \leq r; \\ f(u_i v_i) &= 3r + 2t + 1 - i, & 1 \leq i \leq t. \end{aligned}$$

Note que f é uma rotulação antimágica local de $F_{r,0,t}$ e a rotulação de seus vértices, induzida por f , é dada por

$$\begin{aligned} w(v_i) &= 3r + 2t + 1, & 1 \leq i \leq t; \\ w(u_i) &= 3r + 2t + 1 - i, & 1 \leq i \leq t; \\ w(w_i) &= 2r + 2t + 1, & 1 \leq i \leq r; \\ w(z_i) &= 3r + 2t + 1, & 1 \leq i \leq r; \\ w(c) &= r(4r + 2t + 1) + \frac{t}{2}(t + 1). \end{aligned}$$

Note que $w(u_i) < w(v_i) = w(z_i) < w(c)$ e que $w(u_i) - w(w_i) = r - i$ para todo i , $1 \leq i \leq t$. Assim, quando $r \leq t$, obtemos $w(u_r) = w(w_r)$ e, neste caso, $\chi_{la}(F_{r,0,t}) \leq t + 2$. Logo, da equação (2), temos $\chi_{la}(F_{r,0,t}) = t + 2$ para $r \leq t$. Porém, se $r > t$, então $w(w_i) < w(u_i) < w(v_i) = w(z_i) < w(c)$, donde concluímos que $\chi_{la}(F_{r,0,t}) \leq t + 3$. Por contradição, prova-se a desigualdade contrária e concluímos a demonstração. $\square$

Na Figura 4, mostramos a rotulação antimágica local do grafo $F_{3,0,1}$ e seu respectivo número cromático antimágico local.

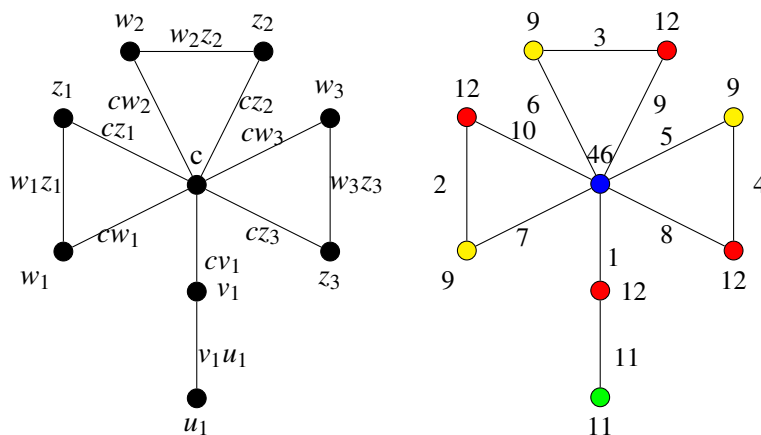


Figura 4: Grafo $F_{3,0,1}$ com $\chi_{la}(F_{3,0,1}) = 4$

O teorema a seguir, cuja demonstração é semelhante aos resultados anteriores, completa a descrição do número cromático antimágico local dos grafos *fireflies*.

Teorema 6. *Sejam r, s e t inteiros positivos. O número cromático antimágico local do grafo $F_{r,s,t}$ é dado por*

$$\chi_{la}(F_{r,s,t}) = \begin{cases} s + t + 2, & \text{se } t < r; \\ s + t + 1, & \text{se } t \geq r. \end{cases}$$

Conclusões

Neste trabalho concluímos a pesquisa sobre o número cromático antimágico local dos grafos da classe *firefly*, cujo início se deu em (ARUMUGAN *et al.*, 2017). O número cromático antimágico local de todas as árvores nesta classe varia entre $\ell + 1$ e $\ell + 2$, sendo ℓ o número de suas folhas. Tais números correspondem aos obtidos em vários artigos sobre tal assunto e nossos resultados reforçam a observação de vários autores de que o número cromático antimágico local de uma árvore deve ser um desses dois valores.

Agradecimentos

Este trabalho é parcialmente financiado pelo MATH-AMSUD no âmbito do projeto GSA, equipe brasileira financiada pela CAPES no projeto 88881.694479/2022-01. A primeira autora agradece o apoio da FAPERJ, processo E-26/202.105/2021. Os demais autores agradecem o apoio da FAPERJ, processo E-20/2022-284573.

Referências

- AOUCHICHE, M.; HANSEN, P.; LUCAS, C. On the extremal values of the second largest Q -eigenvalue. **Linear Algebra and its Applications**, v. 435, p. 2591-2606, 2011.
- ARUMUGAN, S. *et al.* Local antimagic vertex coloring of a graph. **Graphs and Combinatorics**, v. 33, p. 275-285, 2017.
- BACA, M.; SEMANICOVÁ-FENOVCÍKOVÁ, A.; WANG, T.-M. Local antimagic chromatic number for copies of graphs. **Mathematics**, v. 9, n. 11, 1230, 2021.
- BENSMAIL, J.; SENHAJI, M.; LYNGSIE, K.S. On a combination of the 1-2-3 conjecture and the antimagic labeling conjecture. **Discrete Mathematics and Theoretical Computer Science**, v.19, n. 1, 2017.
- ELUMALAI, A. Graph coloring and labelling applications in computer sciences. **Malaya Journal of Matematik**, v. S, n. 2, p. 4039-4041, 2020.
- GALLIAN, J. A. A dynamic survey of graph labeling. **Electronic Journal of Combinatorics**, DS6, 1997. ed. 20, 2017.
- HAGEDORN, T. R. Magic rectangles revisited. **Discrete Mathematics**, v. 207, p. 65-72, 1999.
- HARTSFIELD, N.; RINGEL, G. **Pearls in Graph Theory: A Comprehensive Introduction**. San Diego: Academic Press Inc., 1990.
- HASLEGRAVE, J. Proof of a local antimagic conjecture. **Discrete Mathematics and Theoretical Computer Science**, v. 20, n. 1, 2018.
- MERRIS, R. **Graph Theory**. Canada: JOHN WILEY SONS, INC., 2001.
- NAZULA, N. H.; SLAMIN; DAFIK. Local antimagic vertex coloring of unicyclic graphs. **Indonesian Journal of Combinatorics**, v. 2, n. 1, p. 30-34, 2018.
- SETHURAMAN, G.; SHERMILY, K.M. Antimagic labeling of new classes of trees. **AKCE International Journal of Graphs and Combinatorics**, v. 18, n. 2, p. 110–116, 2021.

O A_α —ESPECTRO DO GRAFO *BLOW-UP*

Uilton Cesar Peres Junior

Centro Federal de Educação Celso Suckow da Fonseca - CEFET/RJ

uilton.junior@aluno.cefet-rj.br

Carla Silva Oliveira

Escola Nacional de Ciências Estatísticas - ENCE/IBGE

carla.oliveira@ibge.gov.br

André Ebling Brondani

Universidade Federal Fluminense - UFF

andrebrondani@id.uff.br

Resumo: Seja G um grafo conexo de ordem n , $A(G)$ sua matriz de adjacência e $D(G)$ sua matriz diagonal dos graus. Em 2017, Nikiforov propôs uma combinação linear convexa que generaliza as matrizes $A(G)$ e $D(G)$ definida da seguinte maneira:

$$A_\alpha(G) = \alpha D(G) + (1 - \alpha)A(G), 0 \leq \alpha \leq 1.$$

Neste artigo, exploramos o espectro da matriz A_α em relação a operação *blow-up* de um grafo, exibimos seu polinômio característico e o seu espectro.

Palavras-chave: Matriz A_α . A_α -espectro. Operação *blow-up*.

Introdução

Seja $G = (V(G), E(G))$ um grafo simples de ordem $n = |V(G)|$. Quando $\{v_i, v_j\} \in E$, v_i e v_j são ditos adjacentes. O grafo complementar de G é o grafo $\overline{G} = (\overline{V}, \overline{E})$ em que $\overline{V} = V(G)$ e $\{v_i, v_j\} \in \overline{E}$ se e somente se $\{v_i, v_j\} \notin E$. O conjunto de vizinhos de um vértice v é denotado por $N_G(v)$. O grau de um vértice $v \in V$, é $d(v) = |N_G(v)|$. Um subgrafo H de G é um grafo $H = (W, F)$ em que $W \subseteq V$ e $F \subseteq E$. Um grafo G é denominado r -regular se cada um de seus vértices tem grau r . O grafo kG é um grafo desconexo com k componentes isomorfas ao grafo G . Os grafos completo, caminho e ciclo de ordem n são denotados por K_n , P_n e C_n , respectivamente. Um grafo G é dito bipartido quando $V = X \cup Y$ tal que $X \cap Y = \emptyset$ e dois vértices adjacentes não pertencem a mesma partição. Além disso, quando todos os vértices de X são adjacentes a todos os vértices de Y , tal grafo é dito bipartido completo e é denotado por $K_{p,q}$, onde $|X| = p$ e $|Y| = q$.

Denotemos por $\mathbb{M}_n(\mathbb{R})$ o conjunto de todas as matrizes quadradas de ordem n com coeficientes em $\mathbb{R}$. Denotamos por I_n a matriz identidade de ordem n e J_n a matriz de ordem n cujos elementos são todos iguais a 1. Dada uma matriz simétrica $M(G)$ associada a um grafo G de ordem n , definimos o M -polinômio característico de G como $p_M(G, x) = \det(xI_n - M(G))$, e suas raízes são denominadas autovalores de $M(G)$. Como usual, indexamos os autovalores de $M(G)$ de forma decrescente, e os denotamos por $\lambda_1(M(G)) \geq \lambda_2(M(G)) \geq \dots \geq \lambda_n(M(G))$. O M -espectro é o multiconjunto dos autovalores de $M(G)$, denotado por $\text{Spec}(M(G)) = \{\lambda_1^{(s_1)}, \lambda_2^{(s_2)}, \dots, \lambda_k^{(s_k)}\}$, onde $s_i = m(\lambda_i)$ é a multiplicidade do autovalor $\lambda_i(M(G))$. Quando todos os autovalores de M são não negativos, M é dita positiva semidefinida.

A matriz de adjacência de G é denotada por $A(G) = [a_{ij}]$, onde $a_{ij} = 1$ se $\{v_i, v_j\} \in E$ e $a_{ij} = 0$, caso contrário. A matriz diagonal dos graus, $D(G) = [d_{ij}]$, é definida por $d_{ii} = d(v_i)$, e $d_{ij} = 0$, $\forall i \neq j$ e a matriz laplaciana sem sinal, denotada por $Q(G)$, é definida como $Q(G) = D(G) + A(G)$ e o menor autovalor desta matriz é denotado por $q_n(G)$. Em 2017, Nikiforov (NIKIFOROV, 2017) definiu a matriz $A_\alpha(G)$ como uma combinação linear convexa de $A(G)$ e $D(G)$ da seguinte forma

$$A_\alpha(G) = \alpha D(G) + (1 - \alpha)A(G), \text{ para } \alpha \in [0, 1].$$

É fácil ver que $A_0(G) = A(G)$, $A_1(G) = D(G)$ e $2A_{1/2}(G) = Q(G)$. Os autovalores de $A_\alpha(G)$ são denotados por $\lambda_i(A_\alpha(G))$ onde $1 \leq i \leq n$, em ordem não crescente, ou seja, $\lambda_1(A_\alpha(G)) \geq \dots \geq \lambda_n(A_\alpha(G))$. Alguns artigos explicitam os autovalores da matriz A_α de algumas classes de grafos, como podemos ver em (NIKIFOROV, 2017), (BRONDANI; OLIVEIRA; FRANÇA, 2019), (BRONDANI; FRANÇA; OLIVEIRA, 2022), (LIN; XUE; SHU, 2018), (ROJO, 2017). Nikiforov e Rojo (2017) definiram α_0 como o menor valor no intervalo $[0, 1]$ tal que $\lambda_n(A_{\alpha_0}(G)) \geq 0$. Desta forma, A_α é positiva semidefinida se e somente se $\alpha_0 \leq \alpha \leq 1$. Neste mesmo artigo, foi levantado o seguinte problema “dado um grafo G , encontrar $\alpha_0(G)$ ”. Este estudo foi feito em artigos como (NIKIFOROV; ROJO, 2017) e (BRONDANI; FRANÇA; OLIVEIRA, 2022).

Os autovalores do grafo *blow-up*, definido na última seção do artigo, e de seu grafo complementar para as matrizes de adjacência e laplaciana sem sinal já foram determinados na literatura como pode ser visto em (NIKIFOROV, 2006) e (LIMA; NIKIFOROV; OLIVEIRA, 2016). Os autores em (LIMA; NIKIFOROV; OLIVEIRA, 2016) usaram a operação *blow-up* para avaliar o comportamento assintótico do valor máximo do autovalor $q_n(G)$ para grafos livres de subgrafos completos de ordem $r + 1$.

Neste trabalho, explicitamos os autovalores e o polinômio característico do grafo *blow-up* e de seu grafo complementar em termos dos autovalores da matriz A_α do grafo.

O texto está organizado com os resultados preliminares na próxima seção e os resultados obtidos na seção seguinte.

Preliminares

Nesta seção, são apresentados os resultados preliminares necessários que são utilizados ao longo do texto, que podem ser vistos em (SILVESTER, 2000) e (NIKIFOROV, 2017). Conceitos gerais envolvendo as teorias de grafos e matrizes que são utilizados neste trabalho podem ser encontrados em (MERRIS, 2001), (BALAKRISHNAN; RANGANATHAN, 2012), (HARARY, 1969) e (HORN; JOHNSON, 1969).

A proposição a seguir relaciona o k -ésimo autovalor de $A_\alpha(G)$ com o grau do k -ésimo vértice de G .

Proposição 1 (NIKIFOROV, 2017). *Se G é um grafo de ordem n com graus $d(v_1) \geq \dots \geq d(v_n)$ e com $A_\alpha(G) = A_\alpha$, então $\lambda_k(A_\alpha) \leq d(v_k)$. Em particular, $\lambda_1(A_\alpha) \leq \Delta(G)$ e $\lambda_n(A_\alpha) < \delta\alpha$.*

As Proposições 2 e 3 explicitam completamente o espectro do grafo completo e do grafo bipartido completo.

Proposição 2 (NIKIFOROV, 2017). *Os autovalores de $A_\alpha(K_n)$ são*

$$\lambda_1(A_\alpha(K_n)) = n - 1 \text{ e } \lambda_k(A_\alpha(K_n)) = \alpha n - 1, \text{ onde } 2 \leq k \leq n.$$

Proposição 3 (NIKIFOROV, 2017). *Seja $p \geq q \geq 1$. Se $\alpha \in [0, 1]$, os autovalores de $A_\alpha(K_{p,q})$ são*

$$\begin{aligned} \lambda_1(A_\alpha(K_{p,q})) &= \frac{1}{2} \left[\alpha(p+q) + \sqrt{\alpha^2(p+q)^2 + 4pq(1-2\alpha)} \right]; \\ \lambda_n(A_\alpha(K_{p,q})) &= \frac{1}{2} \left[(\alpha(p+q) - \sqrt{\alpha^2(p+q)^2 + 4pq(1-2\alpha)}) \right]; \\ \lambda_k(A_\alpha(K_{p,q})) &= \alpha p, \text{ onde } 1 < k \leq q; \\ \lambda_k(A_\alpha(K_{p,q})) &= \alpha q, \text{ onde } q < k < p+q. \end{aligned}$$

A seguir, apresentamos o produto de Kronecker de matrizes e algumas de suas propriedades.

Definição 1. Sejam $A = [a_{ij}] \in \mathbb{M}_{m,n}(\mathbb{R})$ e $B \in \mathbb{M}_{p,q}(\mathbb{R})$. O produto de Kronecker de A e B , denotado por $A \otimes B$, é uma matriz em $\mathbb{M}_{mp,nq}(\mathbb{R})$ definida por

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

Por definição, o produto de Kronecker é escrito como uma matriz em blocos, isto é, $A \otimes B = [a_{ij}B]$, onde $a_{ij}B$ é o (i, j) -ésimo bloco de $A \otimes B$. Vale ressaltar que, como no produto usual de matrizes, o produto de Kronecker não é comutativo. A seguir citamos algumas propriedades que são válidas para o produto de Kronecker de matrizes.

Teorema 1 (BROXSON, 2006). Se $A, B, C \in \mathbb{M}_n(\mathbb{R})$, então

- (i) $(\lambda A) \otimes B = \lambda(A \otimes B) = A \otimes (\lambda B)$;
- (ii) $(A \otimes B)^T = A^T \otimes B^T$;
- (iii) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$;
- (iv) $(A + B) \otimes C = A \otimes C + B \otimes C$;
- (v) $A \otimes (B + C) = A \otimes B + A \otimes C$;
- (vi) $A \otimes 0 = 0 \otimes A = 0$;
- (vii) $I_n \otimes I_m = I_{nm}$

Para concluir esta seção, exibimos a definição da operação *join* e de sua generalização, a operação *H-join*.

Definição 2. Sejam G e H grafos tais que $V(G) \cap V(H) = \emptyset$. O *join* de G e H , denotado por $G \vee H$, é o grafo com conjunto de vértices $V(G \vee H) = V(G) \cup V(H)$, e conjunto de arestas $E(G \vee H) = E(G) \cup E(H) \cup \{\{u, v\} : u \in V(G) \text{ e } v \in V(H)\}$.

Definição 3. Dado um grafo H com $V(H) = \{1, 2, \dots, p\}$ e uma família de grafos $\mathcal{F} = \{G_1, G_2, \dots, G_p\}$ tal que $|V(G_j)| = n_j$, $1 \leq j \leq p$. O *H-join* de $G_1, G_2, \dots, G_p$ é o grafo $G = \bigvee_H \{G_j; 1 \leq j \leq p\}$ tal que

$$V(G) = \bigcup_{j=1}^p V(G_j) \text{ e } E(G) = \left(\bigcup_{j=1}^p E(G_j) \right) \cup \left(\bigcup_{\{r,s\} \in E(H)} E(G_r \vee G_s) \right).$$

Observação 1. Note que quando $H = K_2$ o *H-join* é a operação *join*, descrita na Definição 2.

Exemplo 1. Na Figura 2, ilustramos o grafo *H-join* de $\mathcal{F} = \{C_4, K_2, 2K_1\}$, com $H = P_3$.

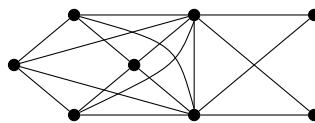


Figura 1: Grafo $\bigvee_{P_3} \{C_4, K_2, 2K_1\}$.

Resultados

Nesta seção, apresentamos o grafo *blow-up* de um grafo G e obtemos resultados acerca de seu A_α -espectro.

Definição 4. Dado um grafo G com $V(G) = \{v_1, \dots, v_n\}$ e um vetor $m = [m_1, \dots, m_n]$ de inteiros positivos, denotamos por $G \circ m$ o grafo obtido de G trocando cada vértice $v_i \in V(G)$ por um conjunto independente de m_i vértices $v_i^1, v_i^2, \dots, v_i^{m_i}$ e $\{v_i^s, v_j^p\} \in E(G \circ m)$ se e somente se $\{v_i, v_j\} \in E(G)$. O grafo $G \circ m$ é chamado de *blow-up* de G . Denotamos o grafo $G \circ m$ por $G^{(t)}$ quando $m_i = m_j$, onde $1 \leq i, j \leq n$.

A Figura 2 mostra os blow-ups $P_3 \circ \mathbf{m}$, onde $\mathbf{m} = (3, 4, 2)$ e $P_3^{(2)}$.

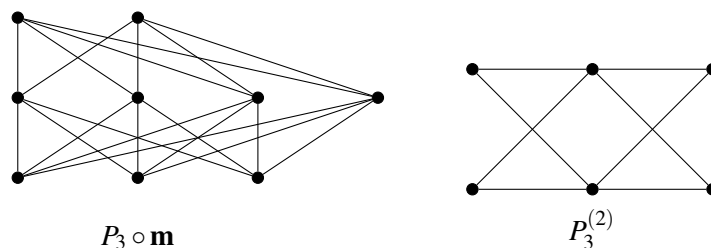


Figura 2: Grafos blow-up

Podemos observar que o *blow-up* de um grafo G de ordem n , $G \circ \mathbf{m}$, onde $\mathbf{m} = [m_1, \dots, m_n]$, é um caso particular do *H-join* quando tomamos $H = G$ e $\mathcal{F} = \{m_1 K_1, \dots, m_n K_1\}$. Em (CARDOSO *et al.*, 2022), os autores expressaram os autovalores da operação *H-join* da matriz A_α utilizando a matriz cadeia e autovalores principais do grafo. Neste trabalho, apresentamos o A_α -espectro do *blow-up* do tipo $G^{(t)}$ utilizando uma outra técnica de demonstração através do produto de Kronecker, além de explicitar o A_α -espectro do grafo complementar da operação *blow-up*, o que ainda não havia sido feito. A matriz de adjacência de $G^{(t)}$ é dada por

$$A(G^{(t)}) = \begin{bmatrix} A(G) & A(G) & \cdots & A(G) \\ A(G) & A(G) & \cdots & A(G) \\ \vdots & \vdots & \ddots & \vdots \\ A(G) & A(G) & \cdots & A(G) \end{bmatrix} = J_t \otimes A(G),$$

e a matriz dos graus

$$D(G^{(t)}) = \begin{bmatrix} tD(G) & 0 & \cdots & 0 \\ 0 & tD(G) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & tD(G) \end{bmatrix} = tI_t \otimes D(G).$$

Desta forma, encontramos a matriz A_α de $G^{(t)}$, dada por

$$A_\alpha(G^{(t)}) = \alpha t I_t \otimes D(G) + (1 - \alpha) J_t \otimes A(G).$$

Agora, para o grafo complementar de $G^{(t)}$, temos

$$A(\overline{G^{(t)}}) = \begin{bmatrix} J_n - I_n - A(G) & J_n - A(G) & \cdots & J_n - A(G) \\ J_n - A(G) & J_n - I_n - A(G) & \cdots & J_n - A(G) \\ \vdots & \vdots & \ddots & \vdots \\ J_n - A(G) & J_n - A(G) & \cdots & J_n - A(G) \end{bmatrix} = J_t \otimes J_n - I_t \otimes I_n - J_t \otimes A(G),$$

que podemos reescrever como $A(\overline{G^{(t)}}) = J_t \otimes (J_n - A(G)) - I_t \otimes I_n$. A matriz dos graus de $\overline{G^{(t)}}$ é dada por

$$D(\overline{G^{(t)}}) = \begin{bmatrix} (tn-1)I_n - tD(G) & 0 & \cdots & 0 \\ 0 & (tn-1)I_n - tD(G) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (tn-1)I_n - tD(G) \end{bmatrix} = I_t \otimes ((tn-1)I_n - tD(G)).$$

Assim,

$$A_\alpha(\overline{G^{(t)}}) = \alpha D(\overline{G^{(t)}}) + (1-\alpha)A(\overline{G^{(t)}}) = I_t \otimes [\alpha t(nI_n - D(G)) - I_n] + (1-\alpha)J_t \otimes (A(\overline{G})).$$

O próximo teorema descreve o espectro das matrizes $A_\alpha(G^{(t)})$ e $A_\alpha(\overline{G^{(t)}})$.

Teorema 2. *Seja G um grafo de ordem n . Para cada i , $1 \leq i \leq n$, as seguintes afirmações são válidas.*

- (i) *Os autovalores de $A_\alpha(G^{(t)})$ são $t\lambda_i(A_\alpha(G))$ e $\alpha td(v_i)$, onde $\alpha td(v_i)$ tem multiplicidade igual a pelo menos $t-1$.*
- (ii) *Os autovalores de $A_\alpha(\overline{G^{(t)}})$ são $t\lambda_i(A_\alpha(\overline{G})) + t-1$ e $\alpha t(n-d(v_i)) - 1$, onde $\alpha t(n-d(v_i)) - 1$ tem multiplicidade igual a pelo menos $t-1$.*

Demonstração. Seja G um grafo de ordem n . Dado um inteiro $t \geq 2$, temos

$$A_\alpha(G^{(t)}) = \alpha t I_t \otimes D(G) + (1-\alpha)J_t \otimes A(G).$$

Como $A_\alpha(G)$ é uma matriz simétrica, existe uma base ortonormal $\{x_1, x_2, \dots, x_n\}$ formada por autovetores de $A_\alpha(G)$. Para cada i , $1 \leq i \leq n$, consideremos o vetor $j_t \otimes x_i$, onde j_t é o vetor com t entradas iguais a 1. Desta forma,

$$\begin{aligned} A_\alpha(G^{(t)})(j_t \otimes x_i) &= \alpha t j_t \otimes D(G)x_i + (1-\alpha)[J_t j_t \otimes A(G)x_i] \\ &= t j_t \otimes \alpha D(G)x_i + t j_t \otimes (1-\alpha)A(G)x_i \\ &= t j_t \otimes A_\alpha(G)x_i \\ &= t\lambda_i(A_\alpha(G))(j_t \otimes x_i). \end{aligned}$$

Assim, para cada i , $t\lambda_i(A_\alpha(G))$ é um autovalor de $A_\alpha(G^{(t)})$. Agora, seja $p_t^{(i,j)}$ um vetor de ordem t onde a i -ésima entrada é igual a 1, a j -ésima entrada igual a -1, com $i \neq j$ e todas as outras entradas iguais a 0, e, seja $\{y_1, \dots, y_n\}$ uma base ortonormal de autovetores associados à $D(G)$. Desta forma,

$$\begin{aligned} A_\alpha(G^{(t)})(p_t^{(i,j)} \otimes y_i) &= \alpha t p_t^{(i,j)} \otimes D(G)y_i + (1-\alpha)[J_t p_t^{(i,j)} \otimes A(G)y_i] \\ &= \alpha t p_t^{(i,j)} \otimes \lambda_i(D(G))y_i \\ &= \alpha td(v_i)(p_t^{(i,j)} \otimes y_i). \end{aligned}$$

Portanto, para cada i , $1 \leq i \leq n$, $\alpha td(v_i)$ é autovalor de $A_\alpha(G^{(t)})$ com multiplicidade ao menos $t-1$, pois podemos escolher $p_t^{(i,j)}$ de $t-1$ maneiras diferentes. Para obter o resultado para item (ii), basta aplicar a mesma técnica de prova para $\overline{G^{(t)}}$. $\square$

Exemplo 2. Na Figura 3 ilustramos o grafo blow-up $P_3^{(3)}$ e o seu grafo complementar $\overline{P_3^{(3)}}$.

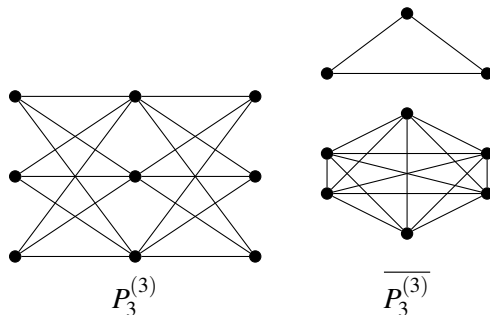


Figura 3: Grafos $P_3^{(3)}$ e $\overline{P_3^{(3)}}$.

Como $\text{Spec}(A_\alpha(P_3)) = \left\{ \alpha, \frac{1}{2} \left(3\alpha \pm \sqrt{9\alpha^2 - 16\alpha + 8} \right) \right\}$ e $\text{Spec}(A_\alpha(\overline{P_3})) = \{0, 1, 2\alpha - 1\}$, do Teorema 2 obtemos

$$\begin{aligned} \text{Spec} \left(A_\alpha \left(P_3^{(3)} \right) \right) &= \left\{ (3\alpha)^{[5]}, (6\alpha)^{[2]}, \frac{3}{2} \left(3\alpha \pm \sqrt{9\alpha^2 - 16\alpha + 8} \right) \right\}; \\ \text{Spec} \left(A_\alpha \left(\overline{P_3^{(3)}} \right) \right) &= \left\{ 2, 5, (6\alpha - 1)^{[5]}, (3\alpha - 1)^{[2]} \right\}. \end{aligned}$$

Corolário 3. Se G é um grafo de ordem n e sequência de graus $d(G) = (d(v_1), \dots, d(v_n))$ então

$$p_{A_\alpha}(G^{(t)}, x) = p_{A_\alpha}(G, x/t) \prod_{i=1}^n (x - t\alpha d(v_i)).$$

Em particular, $\lambda_n(A_\alpha(G^{(t)})) = t\lambda_n(A_\alpha(G))$.

Demonstração. Como consequência do Teorema 2 temos o espectro de $A_\alpha(G^{(t)})$ e consequentemente o seu polinômio característico. Pela Proposição 1 temos que $\lambda_n(A_\alpha(D_2(G))) < \alpha\delta(G^{(t)}) = \alpha t\delta(G) = \alpha t d(v_n)$. Portanto, $\lambda_n(A_\alpha(G^{(t)})) = t\lambda_n(A_\alpha(G))$. $\square$

Exemplo 3. Como consequência do Corolário 3, podemos explicitar o polinômio característico de $A_\alpha(K_n^{(t)})$ e $A_\alpha(K_{p,q}^{(t)})$, utilizando as Proposições 2 e 3, da seguinte maneira:

$$p_{A_\alpha}(K_n^{(t)}, x) = \left(x - tn + t \right) \left(x - \alpha tn + t \right)^{n-1} \left(x - \alpha tn - \alpha t \right)^n$$

e

$$p_{A_\alpha}(K_{p,q}^{(t)}, x) = \left(x - \alpha tp \right)^{2q-1} \left(x - \alpha tq \right)^{2p-1} \left(x^2 - \alpha t(p+q)x - \frac{t^2}{4} \alpha^2 (p+q)^2 + 4pq(1-2\alpha) \right).$$

Do Corolário 3 obtemos a observação a seguir.

Observação 2. $A_\alpha(G^{(t)})$ é positiva semidefinida se e somente se $A_\alpha(G)$ também for.

Para $l \geq 2$, obtemos $G_{l+1}^{(t)}$ a partir de $G_l^{(t)}$ aplicando a operação $G^{(t)}$ ao grafo $G_l^{(t)}$, ou seja $G_{l+1}^{(t)} = (G_l^{(t)})^{(t)}$ e consideramos $G_1^{(t)} = G^{(t)}$. Desta forma, a partir da Proposição 1 chegamos à Observação 2.

Observação 3. Se G é um grafo então $\alpha_0(G) = \alpha_0(G^{(t)}) = \alpha_0(G_2^{(t)}) = \dots = \alpha_0(G_l^{(t)})$, para $l \geq 1$.

Conclusões

Neste trabalho, exibimos o A_α -espectro do *blow-up* de um grafo G em termos do A_α -espectro de G e concluímos que o menor valor de α para o qual $A_\alpha(G)$ é semidefinida positiva não se altera quando aplicamos a operação *blow-up*.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ processo E-20/2022-284573. A pesquisadora Carla Oliveira é apoiada pelo CNPq Grant 304548/2020-0.

Referências

- BALAKRISHNAN, R.; RANGANATHAN, K. **A Textbook of Graph Theory**. New York: Springer, 2012.
- BRONDANI, A. E.; OLIVEIRA, C. S.; FRANÇA, F. A. M.; de LIMA, L. A_α -spectrum of a firefly graph. **Electronic Notes in Theoretical Computer Science**, v. 346, p. 209–219, 2019.
- BRONDANI, A.; FRANÇA, F.; e OLIVEIRA, C. Positive semidefiniteness of $A_\alpha(G)$ on some families of graphs. **Discrete Applied Mathematics**, v. 323, p. 113–123, 2022.
- BROXSON, B. J. **The kronecker product**. Dissertação. – University of North Florida, Florida, 2006.
- CARDOSO, D. M.; GOMES, H.; PINHEIRO, S. J. The H -join of arbitrary families of graphs. **Linear Algebra and its Applications**, v. 648, p. 160-180, 2022.
- HARARY, F. **Graph Theory**. Philippines: Addison-Wesley Plusbling Company, Inc., 1969.
- HORN, R. A.; JOHNSON, C. R. **Matrix analysis**. New York: Cambridge University Press Cambridge [Cambridgeshire], 1985.
- LIN, H.; XUE, J.; e SHU, J. On the A_α -spectra of graphs. **Linear Algebra and its Applications**, v. 556, p. 210–219, 2018.
- LIMA, L.; NIKIFOROV, V.; OLIVEIRA, C. The clique number and the smallest Q -eigenvalue of graphs. **Discrete Mathematics**, v. 339, p. 1744-1752, 2016.
- MERRIS, R. **Graph Theory**. Canada: JOHN WILEY SONS, INC., 2001.
- NIKIFOROV, V. Linear combinations of graph eigenvalues. **Electronic Journal of Linear Algebra**, v. 15, p. 329-336, 2006.
- NIKIFOROV, V. Merging the A - and Q -spectral theories. **Applicable Analysis and Discrete Mathematics**, v.11, p. 81–107, 2017.
- NIKIFOROV, V.; ROJO, O. A note on the positive semidefiniteness of $A_\alpha(G)$. **Linear Algebra and its Applications**, v. 519, p. 156–163, 2017.

Propriedades espectrais de grafos benzenóides

Allan José Amaral Ribeiro
Universidade Federal Fluminense
allanjose@id.uff.br

Francisca Andrea Macedo França
Universidade Federal Fluminense
francisca_franca@id.uff.br

André Ebling Brondani
Universidade Federal Fluminense
andrebrondani@id.uff.br

Resumo: A teoria química dos grafos é uma área moderna e importante da matemática discreta e da química. Grafos benzenóides são redes obtidas pela disposição de hexágonos regulares congruentes no plano, de modo que dois hexágonos sejam disjuntos ou possuam uma aresta comum. Neste trabalho investigamos uma classe de grafos benzenóides, exibimos o seu espectro e expressamos a quantidade de seus autovalores principais em termos de sua estrutura.

Palavras-chave: Matriz de adjacência. Espectro. Espectro principal. Grafos benzenóides.

Introdução

A química matemática é um ramo da química teórica que discute a estrutura molecular por métodos matemáticos sem necessariamente se referir à mecânica quântica. A teoria química dos grafos é um ramo da química matemática que combina matemática, química e teoria dos grafos. Sendo que a teoria dos grafos é utilizada para modelar matematicamente moléculas, a fim de obter informações sobre as propriedades físicas desses compostos químicos. Para mais detalhes, sugerimos (DAS; GUTMAN, 2004) e (RANJINI; LOKESHA; CANGUL, 2011).

Um grafo molecular é um grafo em que os vértices correspondem aos átomos e as arestas às ligações químicas. Os índices topológicos e espectrais dos grafos são definidos e usados em muitas áreas para estudar propriedades de diferentes objetos, como átomos e moléculas. Tais índices são definidos como invariantes de grafos que correspondem e refletem diversas propriedades físicas, químicas, farmacológicas, farmacêuticas e biológicas das espécies químicas subjacentes, veja (GUTMAN, 2017).

A contagem de cadeias em grafos moleculares tem uma longa tradição na química teórica. O fato de que o número de cadeias fechadas é igual ao k -ésimo momento espectral do respectivo grafo molecular é importante na teoria orbital molecular de Hückel (GUTMAN; FURTULA, 2017). Contagens de cadeias fechadas iniciando em um determinado átomo são descritas em vários textos químicos; elas são idênticas às conectividades obtidas pelo Algoritmo de Morgan (MORGAN, 1965) e usadas para discriminar e enumerar os átomos em uma molécula orgânica. Além disso, têm um papel significativo em algoritmos para a percepção de simetrias topológicas (GUTMAN, 2017), que podem estar relacionadas a certos índices topológicos ou espectrais.

Neste trabalho, motivados pela aplicabilidade da contagem de cadeias em grafos moleculares, utilizamos técnicas espectrais para determinar o espectro e a quantidade de autovalores que são necessários para a contagem do número de cadeias em determinados hidrocarbonetos benzenóides. Tais autovalores foram denominados autovalores principais em (CVETKOVIĆ, 1978). Além disso, neste mesmo artigo o autor prova que um grafo possui exatamente 1 autovalor principal se e somente se o grafo é regular e propõe a investigação dos grafos com exatamente $s \geq 2$ autovalores principais.

O trabalho está estruturado da seguinte forma: na próxima seção apresentamos algumas definições e resultados necessários para a compreensão do texto. Nas demais seções apresentamos as contribuições deste trabalho que envolvem a determinação do espectro de uma classe de grafos benzenóides e a solução do problema proposto por Cvetković (1978) para esta mesma classe de grafos. Por fim, apresentamos uma breve conclusão do trabalho.

Preliminares

Nesta seção, além de fixarmos algumas notações, são revistos alguns resultados e conceitos básicos sobre teoria de grafos e teoria de matrizes que são utilizados ao longo do texto. Para mais detalhes sugerimos (DIESTEL, 2000) e (HORN; JOHNSON, 2013).

O espaço vetorial formado pelas matrizes de ordem $m \times n$ e com coeficientes reais é denotado por $\mathbb{M}_{m \times n}(\mathbb{R})$ ou, simplesmente, $\mathbb{M}_n(\mathbb{R})$ quando $m = n$. Dado uma matriz A denotamos a matriz transposta de A por A^T , sua matriz inversa, quando esta existir, por A^{-1} , e por $A[\{i\}, \{j\}]$ a submatriz obtida de A excluindo-se a linha i e a coluna j . Além disso, a matriz identidade e a matriz nula em $\mathbb{M}_n(\mathbb{R})$ são denotadas por I_n e 0_n , respectivamente.

Seja $G = (V, E)$ um grafo simples com $V = \{v_1, v_2, \dots, v_n\}$. A cardinalidade de V é denominada ordem de G , a cardinalidade de E é o tamanho de G e dois vértices $u, v \in V$ são adjacentes quando $\{u, v\} \in E$. Uma cadeia de comprimento k em G é uma sequência de vértices, digamos $(v_0, v_1, \dots, v_k)$, em que $\{v_i, v_{i+1}\} \in E$, onde $0 \leq i \leq k-1$. A cadeia é dita fechada quando $v_0 = v_k$. A matriz de adjacência de G , $A(G) = [a_{ij}] \in \mathbb{M}_n(\mathbb{R})$, tem as entradas dadas por

$$a_{ij} = \begin{cases} 1, & \text{se } \{v_i, v_j\} \in E; \\ 0, & \text{caso contrário.} \end{cases}$$

O polinômio característico de G é definido por $\varphi(G, x) = \det(xI - A)$ e suas raízes são denominadas autovalores de G . O multiconjunto constituído pelos autovalores de G é denominado espectro de G e é denotado por $\sigma(G)$.

A seguir listamos alguns resultados importantes para as próximas seções que envolvem o cálculo de determinante para matrizes em blocos e a determinação de autovalores de certas matrizes tridiagonais. Finalizamos esta seção descrevendo a classe dos grafos benzenóides e definindo uma subclasse desta, que é o objeto de estudo desta pesquisa.

Proposição 1 (HORN; JOHNSON, 2013). *Se A e D são matrizes quadradas, então*

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{cases} \det(A) \det(D - CA^{-1}B), & \text{quando } A^{-1} \text{ existe;} \\ \det(D) \det(A - BD^{-1}C), & \text{quando } D^{-1} \text{ existe.} \end{cases}$$

Proposição 2 (NATH; PAUL, 2014). *Se A e B são matrizes quadradas e $P = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$, então $\sigma(P) = \sigma(A+B) \cup \sigma(A-B)$.*

A Proposição 3, devido a Losonczi (1992), exhibe o espectro de um tipo especial de matriz tridiagonal.

Proposição 3. *Considere a matriz tridiagonal de ordem n , $T_n =$*

$$\begin{bmatrix} b - \alpha & c & 0 & \cdots & 0 \\ a & b & c & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a & b & c \\ 0 & \cdots & 0 & a & b - \beta \end{bmatrix}.$$

(i) *Se $\alpha = \beta = \sqrt{ac} \neq 0$, então $\sigma(T_n) = \left\{ b + 2\sqrt{ac} \cos\left(\frac{j\pi}{n}\right); 1 \leq j \leq n \right\}$.*

$$(ii) \text{ Se } \alpha = \beta = 0, \text{ então } \sigma(T_n) = \left\{ b + 2c \sqrt{\frac{a}{c}} \cos \left(\frac{j\pi}{n+1} \right); 1 \leq j \leq n \right\}.$$

De acordo com Gutman (2017), *grafos benzenóides* são redes obtidas pelo arranjo de hexágonos regulares congruentes no plano de modo que dois hexágonos são disjuntos ou possuem uma aresta em comum.

Considere a subclasse de grafos benzenóides, $\mathcal{F} = \{F_h | (h = \text{número de hexágonos}) \wedge (h \geq 1)\}$, construída da seguinte forma:

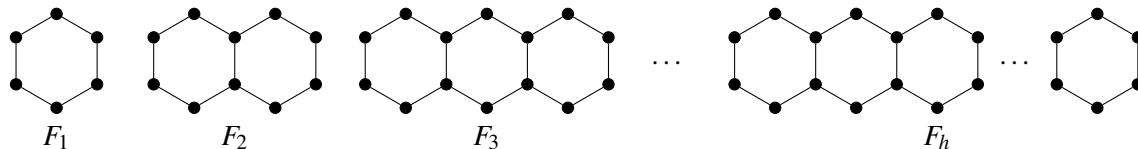


Figura 1: Faixas benzenóides.

Os elementos de $\mathcal{F}$ são denominados *faixas de benzenóides*. Além disso, dado $h > 0$, o grafo $F_h \in \mathcal{F}$ é bipartido, tem ordem $n = 4h + 2$ e tamanho $m = 5h + 1$.

O espectro dos grafos da classe $\mathcal{F}$

Nesta seção apresentamos o espectro dos grafos benzenóides pertencentes à classe $\mathcal{F}$. Dado $h > 0$, existe uma rotulação conveniente dos vértices de $F_h \in \mathcal{F}$ de modo que a sua matriz de adjacência seja dada por

$$A(F_h) = \begin{bmatrix} 0_{h+1} & C_{(h+1) \times h} & I_{h+1} & 0_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h & 0_{h \times (h+1)} & 0_h \\ I_{h+1} & 0_{(h+1) \times h} & 0_{h+1} & C_{(h+1) \times h} \\ 0_{h \times (h+1)} & 0_h & C_{h \times (h+1)}^T & 0_h \end{bmatrix}, \quad (1)$$

onde $C = [c_{ij}] \in \mathbb{M}_{(h+1) \times h}(\mathbb{R})$ é definida por $c_{ij} = \begin{cases} 1, & \text{se } i \in \{j, j+1\}; \\ 0, & \text{caso contrário.} \end{cases}$

A Figura 2 exhibe o grafo F_3 rotulado de modo que a sua matriz de adjacência seja da forma dada em (1).

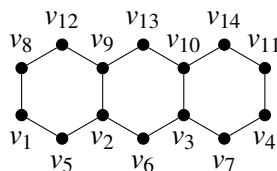


Figura 2: Grafo F_3 com a rotulação adotada.

Lema 4. *Seja C a submatriz de $A(F_h)$, definida em (1). Se $E_{h+1} = CC^T \in \mathbb{M}_{h+1}(\mathbb{R})$ e $Z_h = C^T C \in \mathbb{M}_h(\mathbb{R})$, então*

$$\sigma(E_{h+1}) = \left\{ 2 + 2 \cos \left(\frac{j\pi}{h+1} \right); 1 \leq j \leq h+1 \right\} \text{ e } \sigma(Z_h) = \left\{ 2 + 2 \cos \left(\frac{j\pi}{h+2} \right); 1 \leq j \leq h \right\}.$$

Demonstração. De fato, basta notar que as matrizes E_{h+1} e Z_h têm a forma da matriz dada na Proposição 3, com $a = c = 1$ e $b = 2$. Para E_{h+1} tem-se $\alpha = \beta = 1$ e para Z_h , $\alpha = \beta = 0$. $\square$

Teorema 5. Para cada inteiro positivo h , o espectro do grafo $F_h \in \mathcal{F}$ é dado por

$$\sigma(F_h) = \left\{ -1, 1, \frac{1}{2} \left(1 \pm \sqrt{9 + 8 \cos \left(\frac{j\pi}{h+1} \right)} \right), \frac{1}{2} \left(-1 \pm \sqrt{9 + 8 \cos \left(\frac{j\pi}{h+1} \right)} \right); 1 \leq j \leq h \right\}.$$

Demonstração. Inicialmente, note que a matriz $A(F_h)$, dada em (1), tem a forma

$$A(F_h) = \begin{bmatrix} M & N \\ N & M \end{bmatrix}, \text{ onde } M = \begin{bmatrix} 0_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix} \text{ e } N = \begin{bmatrix} I_{h+1} & 0_{(h+1) \times h} \\ 0_{h \times (h+1)} & 0_h \end{bmatrix}.$$

Da Proposição 2, o espectro de $A(F_h)$ é a união dos espectros das matrizes

$$R_1 = \begin{bmatrix} I_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix} \text{ e } R_2 = \begin{bmatrix} -I_{h+1} & C_{(h+1) \times h} \\ C_{h \times (h+1)}^T & 0_h \end{bmatrix}.$$

Sejam $E_{h+1} = CC^T$ e $Z_h = C^T C$. Do Lema 4, $\det(Z_h) \neq 0$ e da Proposição 1, $\det(R_1) = (-1)^h \det(Z_h)$, $\det(R_2) = -\det(Z_h)$. Portanto, $x = 0$ não é autovalor para R_1 e nem para R_2 . Aplicando a Proposição 1 e visto que o determinante de uma matriz simétrica corresponde ao produto de seus autovalores, para $x \neq 0$, obtemos:

$$\begin{aligned} \varphi(R_1, x) &= \det \begin{bmatrix} (x-1)I_{h+1} & -C_{(h+1) \times h} \\ -C_{h \times (h+1)}^T & xI_h \end{bmatrix} \\ &= x^h \det \left[(x-1)I_{h+1} - \frac{1}{x} E_{h+1} \right] \\ &= \frac{1}{x} \det [x(x-1)I_{h+1} - E_{h+1}] \\ &= \frac{1}{x} \prod_{j=1}^{h+1} (x^2 - x - \mu_j), \end{aligned} \quad (2)$$

onde, do Lema 4, $\mu_j = 2 + 2 \cos \left(\frac{j\pi}{h+1} \right)$, $1 \leq j \leq h+1$, é um autovalor de E_{h+1} . Como $x \neq 0$, da equação (2), obtemos

$$\sigma(R_1) = \left\{ 1, \frac{1}{2} \left(1 \pm \sqrt{9 + 8 \cos \left(\frac{j\pi}{h+1} \right)} \right); 1 \leq j \leq h \right\}.$$

Analogamente, mostra-se que

$$\sigma(R_2) = \left\{ -1, \frac{1}{2} \left(-1 \pm \sqrt{9 + 8 \cos \left(\frac{j\pi}{h+1} \right)} \right); 1 \leq j \leq h \right\}$$

e o resultado segue da Proposição 2. □

A cardinalidade do espectro principal dos grafos da classe $\mathcal{F}$

Denotemos por $\mathbf{1}_n$ o vetor cujas n coordenadas são iguais a 1 e seja $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ uma base ortonormal de autovetores da matriz de adjacência, A , de um grafo G . O Teorema da Decomposição Espectral estabelece que $A = \sum_{i=1}^n \lambda_i \mathbf{x}_i \mathbf{x}_i^T$, onde λ_i é um autovalor de A correspondente ao autovetor $\mathbf{x}_i$. Como a (i, j) -ésima entrada da matriz A^k corresponde ao número de cadeias de comprimento k que iniciam no vértice v_i e terminam no vértice v_j , concluímos que

$$\mathbf{1}_n^T A^k \mathbf{1}_n = \sum_{i=1}^n \lambda_i^k (\mathbf{x}_i^T \mathbf{1}_n)^T (\mathbf{x}_i^T \mathbf{1}_n)$$

equivale ao número de cadeias de comprimento k em G . Os distintos autovalores que contribuem para a contagem do número de cadeias de comprimento k em um grafo são aqueles cujo autoespaço associado não é ortogonal ao vetor $\mathbf{1}_n$.

Seja $G = (V, E)$ um grafo. Uma *partição equilibrada* $\pi = \{V_1, V_2, \dots, V_k\}$ de G é uma partição de V com a seguinte propriedade: dadas duas células, V_i e V_j de π , existe uma constante q_{ij} , tal que cada vértice $v \in V_i$ tem exatamente q_{ij} vizinhos em V_j . A matriz $Q = [q_{ij}] \in \mathbb{M}_k(\mathbb{R})$ é denominada *matriz quociente* de G com respeito à partição π .

Dado $h \geq 1$, considere a rotulação do grafo F_h dada na Figura 3.

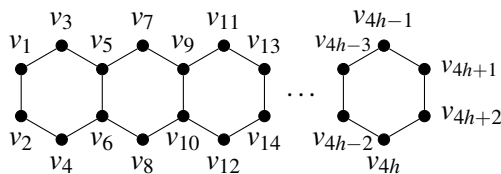


Figura 3: Rotulação de F_h

Sejam $F_h = (V, E) \in \mathcal{F}$, $V = \{v_1, v_2, \dots, v_{4h+2}\}$ e considere os subconjuntos de V , definidos por $U_i = \{v_{2i-1}, v_{2i}\}$ e $W_j = \{v_{4h-2j-1}, v_{4h-2j}\}$, onde $1 \leq i \leq h+1$ e $0 \leq j \leq h-2$. Definimos a partição $\pi(F_h) = \{V_1, V_2, \dots, V_{h+1}\}$, de acordo com a rotulação dada na Figura 3, em que

$$V_i = \begin{cases} U_1 \cup \{v_{4h+1}, v_{4h+2}\}, & \text{se } i = 1; \\ U_2 \cup W_0, & \text{se } i = 2; \\ U_i \cup W_{i-2}, & \text{se } 3 \leq i \leq h; \\ U_{h+1}, & \text{se } i = h+1. \end{cases}$$

Para posterior referência, enunciamos o lema a seguir que decorre diretamente da definição e cuja justificativa é imediata.

Lema 6. Para $h \geq 2$, a partição $\pi(F_h)$ é uma partição equilibrada de F_h .

A matriz quociente de $\pi(F_h)$, $Q_h \in \mathbb{M}_{h+1}(\mathbb{R})$, é definida por $Q_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$ e para $h \geq 3$ é dada,

recursivamente, por:

$$Q_3 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \text{ e } Q_h = \begin{bmatrix} Q_{h-1}[\{h\}, \{h\}] & P \\ P^T & S \end{bmatrix},$$

onde $Q_{h-1}[\{h\}, \{h\}] \in \mathbb{M}_{h-1}(\mathbb{R})$ é obtida de Q_{h-1} excluindo-se a linha h e a coluna h ; $S = \begin{bmatrix} a & 1 \\ 2 & 1-a \end{bmatrix}$

com $a = \frac{1+(-1)^{h-1}}{2}$; $P = [p_{ij}] \in \mathbb{M}_{(h-1) \times 2}(\mathbb{R})$ tal que $p_{ij} = \begin{cases} 1, & \text{se } (i, j) = (h-1, 1); \\ 0, & \text{caso contrário.} \end{cases}$

Seja C uma matriz quadrada de ordem n . Denotamos por $W(C) = [\mathbf{1} \ C\mathbf{1} \ C^2\mathbf{1} \ \dots \ C^{n-1}\mathbf{1}]$ a *matriz cadeia* de C . Em particular, se $C = A(G)$ é a matriz de adjacência de um grafo G , então $W(G) = W(A(G)) = [w_{ij}]$ é denominada *matriz cadeia de G* , onde w_{ij} é o número de cadeias em G de comprimento j que começam no vértice v_i .

Lema 7 (HAGOS, 2002). O posto da matriz cadeia de um grafo G é igual ao número dos seus autovalores principais.

Lema 8 (HUANG et al., 2015). Sejam $\pi(G)$ e Q uma partição equilibrada de um grafo G e a matriz quociente de G com respeito a $\pi(G)$, respectivamente. Então o número de autovalores principais de G é igual ao posto de $W(Q)$.

Para $h \geq 2$, seja $W_h = W(Q_h) = [\mathbf{1} \quad Q_h \mathbf{1} \quad Q_h^2 \mathbf{1} \quad \dots \quad Q_h^h \mathbf{1}] \in \mathbb{M}_{h+1}(\mathbb{R})$ a matriz cadeia da matriz quociente Q_h associada a partição equilibrada $\pi(F_h)$. Por exemplo,

$$W_2 = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 & 5 \\ 1 & 3 & 7 \end{bmatrix} \quad \text{e} \quad W_3 = \begin{bmatrix} 1 & 2 & 4 & 9 \\ 1 & 2 & 5 & 11 \\ 1 & 3 & 7 & 18 \\ 1 & 2 & 6 & 14 \end{bmatrix}.$$

Para $h \geq 4$, o termo geral da matriz cadeia $W_h = [w_{ij}]$ são dados por:

$$w_{ij} = \begin{cases} 1, & \text{se } j = 1; \\ w_{1(j-1)} + w_{2(j-1)}, & \text{se } (i = 1) \wedge (2 \leq j \leq h+1); \\ w_{(i-1)(j-1)} + w_{i(j-1)} + w_{(i+1)(j-1)}, & \text{se } (i \text{ é ímpar}) \wedge (2 \leq j \leq h+1); \\ w_{(i-1)(j-1)} + w_{(i+1)(j-1)}, & \text{se } (i \text{ é par}) \wedge (2 \leq j \leq h+1); \\ 2w_{(h+1)(j-1)} + w_{(h+1)(j-1)}, & \text{se } (i = h+1) \wedge (i \text{ é ímpar}) \wedge (2 \leq j \leq h+1); \\ 2w_{(h+1)(j-1)}, & \text{se } (i = h+1) \wedge (i \text{ é par}) \wedge (2 \leq j \leq h+1). \end{cases}$$

Os dois lemas a seguir podem ser demonstrados através de indução sobre h e por esta razão omitiremos suas demonstrações.

Lema 9. Para $h \geq 3$, considere a matriz cadeia $W_h = [w_{ij}]$. Dado o inteiro j tal que $1 \leq j \leq h-1$, temos:

$$w_{ij} = w_{(i+2)j}, \text{ onde } j \leq i \leq h-1.$$

Lema 10. Para $h \geq 3$, considere a matriz cadeia $W_h = [w_{ij}]$. Temos:

- (i) $w_{(j+1)j} - w_{(j-1)j} = 1, \quad 2 \leq j \leq h;$
- (ii) $w_{(h+1)(h+1)} - w_{(h-1)(h+1)} = \begin{cases} \frac{h+4}{2}, & \text{se } h \text{ é par;} \\ \frac{h+3}{2}, & \text{se } h \text{ é ímpar.} \end{cases}$

Por meio de teste computacionais, programado na linguagem *python*, examinamos o determinante e o posto da matriz cadeia a partir da quantidade de hexágonos dos grafos em $\mathcal{F}$. O grafo F_1 possui exatamente um autovalor principal, pois é um grafo regular. Como $\det(W_2) = -1$, o grafo F_2 possui três autovalores principais, que são os autovalores da matriz Q_2 . Para $h \geq 3$, considere a matriz cadeia $W_h = W(Q_h)$. Substituindo a linha $h+1$ de W_h pela subtração da mesma pela linha $h-1$ e aplicando o Lema 9 e o Lema 10, obtemos

$$\det(W_h) = \begin{cases} \frac{(h+4)}{2} \det(W_{h-1}) - \det(W_{h-1}[\{h\}, \{h\}]), & \text{se } h \text{ é par;} \\ \frac{(h+3)}{2} \det(W_{h-1}) - \det(W_{h-1}[\{h\}, \{h\}]), & \text{se } h \text{ é ímpar.} \end{cases}$$

Aplicando indução sobre h , verifica-se que

$$\det(W_h) = \begin{cases} -1, & \text{se } 8q-6 \leq h \leq 8q-4; \\ 1, & \text{se } 8q-2 \leq h \leq 8q; \\ 0, & \text{se } h = 4q+1, \end{cases}$$

onde $q \in \mathbb{N}$. Note que para $h = 4q+2$,

$$\det(W_{4q+2}) = (2q+3) \det(W_{4q+1}) - \det(W_{4q+1}[\{4q+2\}, \{4q+2\}]).$$

Como $\det(W_{4q+2}) \neq 0$ e $\det(W_{4q+1}) = 0$, obtemos que $\det(W_{4q+1}[\{4q+2\}, \{4q+2\}]) \neq 0$ e portanto, $\lambda = 0$ é um autovalor simples da matriz W_{4q+1} , para todo $q \geq 1$. Os fatos citados anteriormente justificam o próximo teorema.

Teorema 11. O número de autovalores principais de $F_h \in \mathcal{F}$ é igual h , quando $h \equiv 1 \pmod{4}$ e é igual a $h+1$ nos demais casos.

Conclusões

Neste artigo estudamos propriedades espectrais da classe de grafos benzenóides $\mathcal{F}$. Para cada $h \geq 2$, determinamos expressões explícitas para os autovalores de F_h , no Teorema 5. Por fim, no Teorema 11, mostramos que a quantidade de autovalores principais de tais grafos é igual a h quando $h \equiv 1 \pmod{4}$ e é igual a $h + 1$ nos outros casos. Este resultado resolve a questão, proposta por Cvetković, sobre a classificação dos grafos com exatamente $s \geq 2$ autovalores principais para os grafos na classe $\mathcal{F}$.

Agradecimentos

O primeiro autor agradece o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES). Os demais autores agradecem a Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ processo E-20/2022-284573.

Referências

- CVETKOVIĆ, D. M. The main part of the spectrum, divisors and switching of graphs. **Publications de l'Institut Mathématique (Beograd)**, v. 37, p. 31-38, 1978.
- DAS, K.C.; GUTMAN, I. Some properties of the second Zagreb index. **MATCH Communications in Mathematical and in Computer Chemistry**, v. 52, p. 103-112, 2004.
- DIESTEL, R. **Graph theory, in: Graduate Texts in Mathematics 173**. 2. ed. Springer-Verlag, Berlin Heidelberg, 2000.
- GUTMAN, I. **Selected Theorems in Chemical Graph Theory**. 1.ed. Univ. Kragujevac, Kragujevac, 2017.
- GUTMAN, I.; FURTULA, B. The Total π -Electron Energy Saga. **Croatica Chemica Acta**, v. 90, n. 3, p. 359-368, 2017.
- HAGOS, E.M., Some results on graph spectra, **Linear Algebra and its Applications**, v. 356, p. 103-111, 2002.
- HORN, R. A.; JOHNSON, C. R. **Matrix Analysis**. 2. ed. Cambridge, 2013.
- HUANG, X.; HUANG, Q.; LU, L. Construction of graphs with exactly k main eigenvalues. **Linear Algebra and its Applications**, v. 486, p. 204-218, 2015.
- LOSONCZI, L. Eigenvalues and eigenvectors of some tridiagonal matrices. **Acta Mathematica Hungarica**, v. 60, p. 309-332, 1992.
- MORGAN, H. L. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. **Journal of Chemical Documentation**, v. 5, n. 2, p. 107-113, 1965.
- NATH, M.; PAUL, S. On the distance Laplacian spectra of graphs. **Linear Algebra and its Applications**, v. 460, p. 97-110, 2014.
- RANJINI, P. S.; LOKESHA, V.; CANGUL, I. N., On the Zagreb indices of the line graphs of the subdivision graphs. **Applied Mathematics and Computation**, v. 218, p. 699-702, 2011.

Boa Colocação e Analiticidade para um problemas Vigas Termoelásticas do tipo III

Leandro Tomaz de Araujo

Universidade Federal Rural do Rio de Janeiro

leandroaraujo@ufrj.br

Andrea Luiza Gonçalves Martinho

Universidade Federal Rural do Rio de Janeiro

andrea@ufrj.br

Resumo: Neste trabalho estudaremos um problema envolvendo vigas termoelásticas do tipo III onde a viga esta engastada nas extremidades e de material completamente termoelástico. O objetivo deste trabalho é apresentar a boa colocação e analiticidade para este problema.

Palavras-chave: Semigrupo, Analiticidade, Sistemas Termoelásticos

1 Introdução

Um problema é dito do tipo III quando o fluxo de calor ocorre através de pulsos de calor em sólidos rígidos ou materiais elásticos como um especie de *ondas térmicas* em velocidade finita de propagação, por exemplo, é como os efeitos térmicos em uma viga em balanço. Note que experimentalmente em baixas temperaturas, o calor se propaga como ondas térmicas. Esse fenômeno é conhecido *second sound*, sendo semelhante a propagação de ondas no ar. A condição do Tipo III é portanto a resposta deste material a este fenômeno térmico é baseado em funções constitutivas como visto anteriormente. Este tipo de condição foi apresentada em 1991 quando Green & Naghdi [3, 4] apresentaram as três teorias termoelásticas que a denominaram do tipo I, II e III. Neste trabalho mostraremos a boa colocação do problema e a analiticidade para um problema associando ao sistema de equações:

$$\rho u_{tt} + \alpha u_{xxxx} + m\theta_{txx} = 0 \text{ para } x \in (0, \ell), t > 0, \quad (1)$$

$$c\theta_{tt} - \kappa\theta_{xx} - \kappa_0\theta_{xxt} - mu_{xxt} = 0 \text{ para } x \in (0, \ell), t > 0, \quad (2)$$

com as condições iniciais da forma

$$u(0, x) = u_0(x), u_t(0, x) = u_1(x), \text{ em } (0, \ell), \quad (3)$$

$$\theta(0, x) = \theta_0(x), \theta_t(0, x) = \theta_1(x) \text{ em } (0, \ell), \quad (4)$$

e munida as condições de contorno:

$$u(t, 0) = u(t, \ell) = u_x(t, 0) = u_x(t, \ell) = \theta_x(t, 0) = \theta_x(t, \ell) = 0 \text{ em } (0, +\infty). \quad (5)$$

As hipóteses que foram assumidas foram $\rho, \alpha, c, \kappa, m$ e κ_0 são constantes positivas.

Este trabalho esta organizado da seguinte maneira. Na seção 2 apresentamos o espaço fase que é o maior espaço de Hilbert onde o funcional de energia faz sentido e definiremos operador linear e dissipativo $\mathcal{A}$ e o seu domínio que permitem reescrever o problema acima em um problema de Cauchy. Além disso, provaremos que operador $\mathcal{A}$ é dissipativo através da seguinte adaptação da proposição 4.6 de [7]. Na seção 3 estudaremos a existência e unicidade de solução para o sistema (1)-(5), e para tal é suficiente demonstrar que $0 \in \rho(\mathcal{A})$ uma vez que $\mathcal{H}$ é um espaço de Hilbert e o operador $\mathcal{A}$ é linear, dissipativo e com domínio denso. Finalmente, na seção 4 provaremos uma propriedade muito importante: a analiticidade do semigrupo associando ao sistema (1)-(5). Para tal utilizaremos o seguinte resultado

Teorema 1. *Seja $\rho(\mathcal{A})$ o conjunto resolvente do operador $\mathcal{A}$. Suponhamos que $i\mathbb{R} \subset \rho(\mathcal{A})$. Então, um C_0 -semigrupo de contrações $T(t)$ em um espaço de Hilbert $\mathcal{H}$ com a norma $\|\cdot\|_{\mathcal{H}}$ é analítico se, e somente se,*

$$\limsup_{|\lambda| \rightarrow +\infty} \|\lambda(i\lambda I - \mathcal{A})^{-1}\|_{\mathcal{L}(\mathcal{H})} < \infty.$$

Demonstração. Veja [5], p.5. □

Os dois resultado mais importante deste trabalho pode ser resumido na forma dos Teorema 3 e 4. Vale ressaltar que este problema já foi estudado antes em [6], mas a regularidade e abordagem feita aqui é diferente e por Avalos & Rivera [2] com outra configuração.

2 Espaço Fase e Operador $\mathcal{A}$

A energia associada ao modelo (1)-(5) é dada por

$$E(t) = \frac{1}{2} \int_0^\ell \left(\rho |u_t|^2 + \alpha |u_{xx}|^2 + c |\theta_t|^2 + \kappa |\theta_x|^2 \right) dx.$$

De onde temos

$$\frac{d}{dt} E(t) = - \int_0^\ell \kappa_0 |\theta_{xt}|^2 dx. \quad (6)$$

Denotando por $u_t = v$, $\theta_t = \Theta$, e $U = (u, v, \theta, \Theta)^T$. Definimos o espaço de fase $\mathcal{H}$ da seguinte forma

$$\mathcal{H} = H_0^2(0, \ell) \times L^2(0, \ell) \times H_*^1(0, \ell) \times L_*^2(0, \ell),$$

onde

$$L_*^2(0, \ell) = \left\{ g \in L^2(0, \ell); \int_0^\ell g(s) ds = 0 \right\}, \quad H_*^m(0, \ell) = H^m(0, \ell) \cap L_*^2(0, \ell), \quad m = 1, 2.$$

Note que $\mathcal{H}$ é um espaço de Hilbert equipado com o produto interno

$$(U, U^*)_{\mathcal{H}} = \int_0^\ell \left(\rho v \overline{v^*} + \alpha u_{xx} \overline{u_{xx}^*} + c \Theta \overline{\Theta^*} + \kappa \theta_x \overline{\theta_x^*} \right) dx,$$

onde $U = (u, v, \theta, \Theta)^T$ e $U^* = (u^*, v^*, \theta^*, \Theta^*)^T$. O operador $\mathcal{A}$, para $\rho = c = 1$ é dado pela seguinte expressão

$$U = \begin{pmatrix} u \\ v \\ \theta \\ \Theta \end{pmatrix}, \quad \mathcal{A}U = \begin{pmatrix} v \\ -\alpha u_{xxx} - m \Theta_{xx} \\ \Theta \\ \kappa \theta_{xx} + \kappa_0 \Theta_{xx} + m v_{xx} \end{pmatrix}.$$

Portanto, o sistema (1)-(5) pode ser reescrito como

$$U_t - \mathcal{A}U = 0, \quad U(0) = U_0. \quad (7)$$

com domínio $D(\mathcal{A}) = \{U \in \mathcal{H}; \mathcal{A}U \in \mathcal{H}\}$ é dado por

$$D(\mathcal{A}) = \left\{ U \in \mathcal{H} \left| \begin{array}{l} v \in H_0^2(0, \ell), \boxed{u \in H^3(0, \ell)} \\ -\alpha u_{xx} - m \Theta \in H^2(0, \ell), \\ \Theta \in H_*^1(0, \ell) \\ \kappa \theta + \kappa_0 \Theta + m v \in H_*^2(0, \ell). \end{array} \right. \right\}. \quad (8)$$

Uma das principais dificuldades deste problema é que o domínio não pode ser representado na forma de um espaço de Sobolev, mas na forma de um espaço vetorial dependendo dos operadores diferenciais da equação. Além disso, podemos verificar que o operador $\mathcal{A}$ definido acima é dissipativo pois depois de um cálculo rápido obtemos

$$\operatorname{Re}(\mathcal{A}U, U)_{\mathcal{H}} = - \int_0^\ell \kappa_0 |\Theta_x|^2 dx \leq 0. \quad (9)$$

Para mostrar a boa colocação do problema utilizaremos, além do domínio do operador, o sistema resolvente associado ao modelo

$$(i\lambda I - \mathcal{A})U = F. \quad (10)$$

Em termos de suas componentes, o sistema resolvente pode ser escrito como

$$i\lambda u - v = f_1, \quad (11)$$

$$i\lambda v + \alpha u_{xxxx} + m\Theta_{xx} = f_2, \quad (12)$$

$$i\lambda \theta - \Theta = f_3, \quad (13)$$

$$i\lambda \Theta - \kappa \theta_{xx} - \kappa_0 \Theta_{xx} - mv_{xx} = f_4, \quad (14)$$

verificando a condição de contorno

$$u(0) = u(\ell) = u_x(0) = u_x(\ell) = \theta_x(0) = \theta_x(\ell) = 0. \quad (15)$$

Tomando o produto interno da equação resolvente (10) com U e depois tomando a parte real obtemos

$$\int_0^\ell \kappa_0 |\Theta_x|^2 dx = \operatorname{Re} (U, F)_{\mathcal{H}} \quad (16)$$

Agora, usando a desigualdade de Cauchy-Schwarz obtemos

$$\int_0^\ell |\Theta_x|^2 dx \leq C \|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}},$$

onde C é uma constante positiva.

Uma das dificuldades para mostrar a existência é mostrar a densidade do domínio do operador $\mathcal{A}$. Para este fim, usaremos uma adaptação da Proposição 4.6 de Pazy [7] que fornece condições suficientes para $\overline{D(\mathcal{A})} = \mathcal{H}$.

Teorema 2. *Seja $\mathcal{A}$ um operador dissipativo verificando $R(\mathcal{A}) = \mathcal{H}$. Se o espaço $\mathcal{H}$ é reflexivo então $\overline{D(\mathcal{A})} = \mathcal{H}$.*

3 Existência e Unicidade de soluções

Finalmente, estamos em condições de provar o resultado principal desta seção que pode ser resumido na forma do Teorema a seguir.

Teorema 3. *O operador $\mathcal{A}$ é o gerador infinitesimal de um semigrupo- C_0 de contrações $(S(t))_{t \geq 0}$ sobre o espaço $\mathcal{H}$. Além disso, para todo dado inicial $U_0 \in D(\mathcal{A})$ existe uma única solução do problema (7) verificando*

$$U \in C^1([0, +\infty[, \mathcal{H}) \cap C^0([0, +\infty[, D(\mathcal{A})).$$

Demonstração. Como o operador $\mathcal{A}$ é dissipativo. É suficiente mostrar que o domínio $D(\mathcal{A})$ é denso em $\mathcal{H}$ e que $0 \in \varrho(\mathcal{A})$. Para este propósito, provaremos que para todo $F = (f_1, f_2, f_3, f_4)^T \in \mathcal{H}$ existe um único $U = (u, v, \theta, \Theta)^T \in D(\mathcal{A})$ tal que $\mathcal{A}U = F$, ou equivalentemente em termos de suas componentes,

$$-v = f_1, \quad (17)$$

$$\alpha u_{xxxx} + m\Theta_{xx} = f_2, \quad (18)$$

$$-\Theta = f_3, \quad (19)$$

$$-\kappa\theta_{xx} - \kappa_0\Theta_{xx} - mv_{xx} = f_4. \quad (20)$$

De fato, substituindo a equação (19) em (18), e depois as equações (17) e (19) em (20) obtemos

$$\alpha u_{xxxx} = -mf_3'' + f_2, \quad (21)$$

$$-\kappa\theta_{xx} = -\kappa_0 f_3'' + mf_1'' + f_4. \quad (22)$$

Note que o sistema anterior é desacoplado. Então, usando as condições de contorno obtemos

$$\alpha u_{xxxx} = -mf_3'' + f_2 \in H^{-1}(0, \ell), \quad (23)$$

$$u(0) = u(\ell) = u_x(0) = u_x(\ell) = 0, \quad (24)$$

$$-\kappa\theta_{xx} = -\kappa_0 f_3'' + mf_1'' + f_4 \in H^{-1}(0, \ell), \quad (25)$$

$$\theta_x(0) = \theta_x(\ell) = 0. \quad (26)$$

Usando o Lema de Lax-Milgran, não é difícil de verificar que o sistema (25)-(26) possui uma única solução $\theta \in H_*^1(0, \ell)$ verificando

$$-\kappa\theta_{xx} = mf_1'' - \kappa_0 f_3'' + f_4 \in L^2(0, \ell). \quad (27)$$

A seguir resolvemos (21)-(24). Para isto introduzimos a seguinte forma sesquilinear

$$a(u, v) = \int_0^\ell \alpha u_{xx} \overline{v_{xx}} dx.$$

Note que $a(\cdot, \cdot)$ é simétrica, continua e coerciva sobre o espaço $H_0^2(0, \ell) \times H_0^2(0, \ell)$. Então, a forma linear

$$T(v) = \int_0^\ell (mf_3' \overline{v_x} + f_2 \overline{v}) dx,$$

verifica $T \in H^{-2}$. Portanto, pelo Lema de Lax-Milgran, existe uma única solução $u \in H^3(0, \ell)$ tal que

$$\int_0^\ell \alpha u_{xx} \overline{v_{xx}} dx = \int_0^\ell (mf_3' \overline{v_x} + f_2 \overline{v}) dx, \quad \forall v \in H_0^2(0, \ell). \quad (28)$$

Para cada $v \in C_0^\infty(0, \ell)$ temos que u verifica (21) no sentido das distribuições e

$$\alpha u_{xxxx} + mf_3'' = f_2 \in L^2(0, \ell). \quad (29)$$

Usando (27)-(29) concluímos que $U = (u, v, \theta, \Theta)^T \in D(\mathcal{A})$. Logo, $R(\mathcal{A}) = \mathcal{H}$; e como $\mathcal{H}$ é reflexivo e $\mathcal{A}$ é dissipativo, então $D(\mathcal{A})$ é denso, por causa do Teorema 2. Portanto, $\mathcal{A}$ é o gerador infinitesimal de um semigrupo- C_0 de contrações. $\square$

4 Analiticidade do Semigrupo

Nosso ponto de partida é provar a estabilidade forte que resumimos na forma da próxima Proposição e é similar a prova do Lema 3.2 de Avalos & Rivera [2], por essa razão nos restringimos a dar um esquema da prova.

Proposição 1 (Estabilidade Forte). *O operador $\mathcal{A}$ verifica $i\mathbb{R} \subset \rho(\mathcal{A})$.*

Demonstração. Denotemos por

$$\mathcal{N} = \{s \in \mathbb{R}^+ :] - is, is[\subset \varrho(\mathcal{A})\}.$$

Como $0 \in \varrho(\mathcal{A})$, $\mathcal{N} \neq \emptyset$. Agora, denotando $\sigma = \sup \mathcal{N}$ temos as seguintes possibilidades: $\sigma = +\infty$ ou $0 < \sigma < \infty$. É suficiente demonstrar que $0 < \sigma < \infty$ não pode acontecer. A demonstração é feita por contradição aonde se constrói uma sequência $U_n \in D(\mathcal{A})$ tal que $\|U_n\|_{\mathcal{H}} = 1$ e $U_n \rightarrow 0$. $\square$

Antes de provarmos a analiticidade do semigrupo associado ao sistema (1)-(5) vamos precisar de alguns resultados técnicos. Além disso, assuma que $U = (u, v, \theta, \Theta)^T \in D(\mathcal{A})$ solução da equação resolvente $i\lambda U - \mathcal{A}U = F$, onde $F = (f_1, f_2, f_3, f_4)^T \in \mathcal{H}$.

Lema 1. *Sob as condições anteriores,*

$$\int_0^\ell |v_x|^2 dx \leq C|\lambda| \|U\|_{\mathcal{H}}^2 + C\|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}}. \quad (30)$$

onde C é uma constante positiva.

Demonstração. Como $v \in H_0^2(0, \ell)$, segue pela Desigualdade de Gagliardo-Nirenberg (veja teorema 1.4.5 de [5]) e usando a equação (11). $\square$

Proposição 2 (1ª Estimativa). *Suponha as condições anteriores, então para todo $\epsilon > 0$,*

$$\int_0^\ell |\lambda \theta_x|^2 dx \leq C\|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}} + C\|F\|_{\mathcal{H}}^2, \quad (31)$$

$$\int_0^\ell |\lambda \Theta|^2 dx \leq 2\epsilon |\lambda|^2 \|U\|_{\mathcal{H}}^2 + C_\epsilon \|F\|_{\mathcal{H}}^2. \quad (32)$$

onde C e C_ϵ são constantes positivas.

Demonstração. Usando (13) na equação (16) segue a primeira desigualdade. Agora, multiplicando a equação (14) por $i\overline{\lambda\Theta}$ e tomando a parte real e usando as equações (13), (16) e desigualdade de Young obtemos a segunda desigualdade segue para λ grande. $\square$

Para encontrar as limitações que implica na analiticidade do semigrupo, introduzimos o funcional

$$\mathcal{J}(x) = \frac{1}{2} \left(|v(x)|^2 + \alpha |u_{xx}(x)|^2 \right).$$

Sob estas notações temos

Lema 2. *Sob as condições anteriores, então para todo $\epsilon > 0$,*

$$\left| \frac{\ell}{2} \mathcal{J}(0) + \frac{\ell}{2} \mathcal{J}(\ell) - \int_0^\ell \left(\mathcal{J}(x) + \alpha |u_{xx}|^2 \right) dx \right| \leq C_\epsilon \|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}} + \epsilon \int_0^\ell \mathcal{J}(x) dx, \quad (33)$$

onde C_ϵ é uma constante positiva.

Demonstração. O resultado é obtido multiplicando (12) por $q\overline{u_x}$ onde $q(x) = x - \frac{\ell}{2}$ e integrando por partes. $\square$

A proposição a seguir a qual omitiremos a prova pois é similar a prova do Lema 3.5 de Avalos & Rivera [2].

Lema 3. *Suponha as condições anteriores, então para todo $\epsilon > 0$,*

$$\left| u_{xx}(0) - \frac{u_{xxx}(0)}{i\sigma} \right| \leq \epsilon |u_{xx}(0)| + \epsilon \left| \frac{u_{xxx}(0)}{i\sigma} \right| + C \|U\|_{\mathcal{H}}^{1/2} \|F\|_{\mathcal{H}}^{1/2} + \frac{1}{\sigma^{1/2}} \|F\|_{\mathcal{H}}. \quad (34)$$

$$\left| u_{xx}(\ell) - \frac{u_{xxx}(\ell)}{i\sigma} \right| \leq \epsilon |u_{xx}(\ell)| + \epsilon \left| \frac{u_{xxx}(\ell)}{i\sigma} \right| + C \|U\|_{\mathcal{H}}^{1/2} \|F\|_{\mathcal{H}}^{1/2} + \frac{1}{\sigma^{1/2}} \|F\|_{\mathcal{H}}. \quad (35)$$

onde C é uma contante positiva e σ suficientemente grande.

O seguinte Lema ocupa um papel importante na demonstração do nosso resultado.

Lema 4. *Sob as mesmas condições anteriores temos que a solução do sistema resolvente verifica*

$$\int_0^\ell |u_{xxx}|^2 dx \leq C|\lambda| \|U\|_{\mathcal{H}}^2 + C \|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}},$$

onde C é uma constante positiva.

Demonstração. Multiplicando a equação (12) por $\overline{u_{xx}}$ e integrando por partes e na sequencia usando (11) obtemos

$$\int_0^\ell \alpha |u_{xxx}|^2 dx = \int_0^\ell |v_x|^2 dx - \int_0^\ell (v_x \overline{f_1'} + f_2 \overline{u_{xx}}) dx + [\alpha u_{xxx} \overline{u_{xx}}]_0^\ell - \int_0^\ell m \Theta_{xx} \overline{u_{xx}} dx.$$

Agora, tomando o valor absoluto e depois usando a Proposição 2, a equação (16) e os Lemas 2 e 3 obtemos o resultado para λ suficientemente grande. $\square$

Proposição 3 (2ª Estimativa). *Suponha as condições anteriores, então para todo $\epsilon > 0$,*

$$\int_0^\ell |\lambda u_{xx}|^2 dx \leq 2\epsilon |\lambda|^2 \|U\|_{\mathcal{H}}^2 + C_\epsilon \|F\|_{\mathcal{H}}^2,$$

onde C_ϵ é uma constante positiva.

Demonstração. Substituindo a equação (11) na equação (14), e depois multiplicando a equação por $i\lambda \overline{u_{xx}}$ e depois tomando a parte real. O resultado segue usando a desigualdade de Young e a equação (16) obtemos

$$\left| \int_0^\ell (\kappa \theta_{xx} + \kappa \Theta_{xx}) i \lambda \overline{u_{xx}} dx \right| \leq C_\epsilon |\lambda| \|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}} + \epsilon |\lambda| \|u_{xx}\|_{L^2}^2.$$

Agora, usando a Proposição 2 e o Lema 4 e depois a Proposição 2 segue o resultado. $\square$

Proposição 4 (3ª Estimativa). *Sob as mesmas condições anteriores temos que a solução do sistema resolvente verifica para todo $\epsilon > 0$,*

$$\int_0^\ell |\lambda v|^2 dx \leq 2\epsilon |\lambda|^2 \|U\|_{\mathcal{H}}^2 + C_\epsilon \|F\|_{\mathcal{H}}^2.$$

onde C_ϵ é uma contante positiva.

Demonstração. Multiplicando a equação (12) por $\overline{i\lambda v}$ e integrando sobre o intervalo $(0, \ell)$ e depois usando as Proposições 3, 2 e 3 segue o resultado. $\square$

Finalmente, estamos em condições de mostrar a analiticidade do semigrupo- C_0 de contrações.

Teorema 4. *O semigrupo associado ao sistema (1)-(5) é analítico.*

Demonstração. Note que da Proposição 1, $i\mathbb{R} \subset \rho(\mathcal{A})$. Agora pelas Proposição 2, Proposição 3 e Proposição 4 obtemos

$$|\lambda|^2 \|U\|_{\mathcal{H}}^2 = \int_0^\ell (|\lambda v|^2 + \alpha |\lambda u_{xx}|^2 + |\lambda \Theta|^2 + \kappa |\lambda \theta|^2) dx \leq 2\epsilon |\lambda|^2 \|U\|_{\mathcal{H}}^2 + C_\epsilon \|F\|_{\mathcal{H}}^2.$$

assim temos que $|\lambda|^2 \|U\|_{\mathcal{H}}^2 \leq C_\epsilon \|F\|_{\mathcal{H}}^2$. De onde segue o resultado. $\square$

5 Conclusões

Neste trabalho mostramos a boa colocação dos modelos de vigas termoelásticas do tipo III global. Para tal demonstração, utilizamos a teoria clássica de Semigrupos e Análise Funcional que pode ser encontrada em [1, 7]. Além disso, mostramos que a dissipação produzida é suficiente para que o modelo seja analítico. Esta propriedade é muito importante porque em particular implica nas seguintes propriedades: estabilidade exponencial do modelo, o semigrupo possui a propriedade de crescimento definido pelo espectro, a solução possui efeito regularizante sobre os dados iniciais e a solução admite uma expansão em séries de potências convergentes no tempo.

6 Agradecimentos

Este trabalho foi desenvolvido sobre orientação do Prof. Jaime M. Rivera como parte da tese de doutorado dos autores pelo Instituto de Matemática da UFRJ.

Referências

- [1] ADAMS, R. A.; FOURNIER, J.J. F, *Sobolev Spaces*. Second Edition. Academic Press, 2003.
- [2] ÁVALOS, G. G., Muñoz RIVERA, J., & OCHOA OCHOA, E. (2022). *Analiticity of the Type III Thermoelastic Euler Bernoulli Model with Tip*. Acta Applicandae Mathematicae, 182(1). <https://doi.org/10.1007/s10440-022-00543-5>
- [3] GREEN, A. E., & NAGHDI, P. M. (1991). *A re-examination of the basic postulates of thermomechanics*. Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences, 432(1885), 171-194. <https://doi.org/10.1098/rspa.1991.0012>
- [4] GREEN, A. E. & NAGHDI, P. M. (1993). *Thermoelasticity without energy dissipation*. Journal of Elasticity, 31(3), 189-208. <https://doi.org/10.1007/BF00044969>
- [5] LIU, Z. & ZHENG, S. *Semigroups Associated with Dissipative Systems*, CRC Research Notes in Mathematics, vol. 398, Chapman and Hall, 1999.
- [6] LIU, Z. & QUINTANILLA, R. (2010). *Analyticity of solutions in type III thermoelastic plates*. IMA Journal of Applied Mathematics, 75(3), 356-365. <https://doi.org/10.1093/imamat/hxq016>
- [7] Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Vol. 44). Springer New York . <https://doi.org/10.1007/978-1-4612-5561-1>

Boa Colocação e Estabilidade Exponencial para dois problemas de Viga Viscoelástica

Andrea Luiza Gonçalves Martinho
Universidade Federal Rural do Rio de Janeiro
andrea@ufrj.br

Leandro Tomaz de Araujo
Universidade Federal Rural do Rio de Janeiro
leandroaraujo@ufrj.br

Resumo: Neste trabalho estudaremos a existência e a unicidade de soluções para dois sistemas de vigas de material viscoelásticos bem como a estabilidade exponencial. Além disso, mostraremos que a perda de analiticidade para um dos sistemas.

Palavras-chave: Semigrupos, Estabilidade Exponencial, Perda de Analiticidade.

Introdução

Neste trabalho estudaremos o comportamento assintótico e parte do efeito regularizante de uma viga totalmente de material viscoelástico de comprimento ℓ . Este estudo foi motivado pela grande importância que os materiais viscoelásticos exercem na engenharia, na física e em muitas outras áreas como podemos ver, por exemplo, em [3, 4]. Mais precisamente, consideraremos uma viga feita de material viscoelástico estruturada no intervalo $[0, \ell]$ com especificamente com duas configurações para as extremidades. A primeira configuração envolve uma viga engastada em suas extremidades, e a segunda configuração envolve uma viga apoiada em suas extremidades. Em qualquer um dos casos consideraremos a deformação da viga em um ponto x no tempo t é u , onde $u = u(x, t)$ com $x \in]0, \ell[$. O que direciona a investigação da viga unidimensional considerada através de um modelo envolvendo um problema com valor de contorno e valor inicial. Os modelos que consideraremos para a viga viscoelástica são dados abaixo:

$$\rho u_{tt} + \alpha u_{xxxx} - \alpha_0 u_{txx} = 0, \text{ para } (x, t) \in]0, \ell[\times]0, +\infty[\quad (1)$$

com condições iniciais

$$u(x, 0) = u_0(x) \text{ em }]0, \ell[; \quad (2)$$

$$u_t(x, 0) = u_1(x) \text{ em }]0, \ell[; \quad (3)$$

e associado com uma das condições de fronteira do tipo

$$u(0, t) = u(\ell, t) = u_x(0, t) = u_x(\ell, t) = 0, \text{ em }]0, +\infty[\quad (4)$$

ou

$$u(0, t) = u(\ell, t) = u_{xx}(0, t) = u_{xx}(\ell, t) = 0, \text{ em }]0, +\infty[\quad (5)$$

onde ρ, α, α_0 são constantes positivas.

O comportamento assintótico é a parte mais relevante deste trabalho o qual consiste no decaimento exponencial para os problemas envolvendo a viga viscoelástica para um problema determinado por (1)-(3) e associado com condições de fronteira do tipo (4) ou (5). Este resultado foi obtido usando o Teorema de Pruss que pode ser encontrado em Zheng & Liu [5].

Em relação ao efeito regularizante o resultado obtido foi que para um problema envolvendo uma viga viscoelástica para um problema determinado por (1)-(3) e associado com condições de

fronteira do tipo (5) temos a perda de analiticidade. Este resultado foi obtido usando o teorema 4 que pode ser encontrado em [5].

Agora, observe que em ambos modelos obteremos a existência e unicidade de soluções para estes problemas com técnicas de Semigrupos e Espaços de Sobolev que podem ser encontrados em [1, 2, 6].

Este trabalho esta organizado como segue. Na seção 1 apresentaremos o espaço fase e o domínio do operador. Na seção 2 apresentaremos a boa colocação para ambos os problemas usando a teoria bem conhecida. Na seção 3 apresentamos a estabilidade exponencial. Finalmente, na seção 4 apresentamos a perda de analiticidade para o problema determinado por (1)-(3) e associado com condições de fronteira do tipo (5).

1 Espaço fase e operador $\mathcal{A}$

Nesta seção obteremos um resultado de existência e unicidade de soluções do problema determinado por (1)-(3) e associado com condições de fronteira do tipo (4) ou (5). Note que a energia associada ao modelo (1)-(3) é dada por

$$E(t) = \frac{1}{2} \int_0^\ell \left(\rho |u_t|^2 + \alpha |u_{xx}|^2 \right) dx.$$

Um cálculo rapido mostra que

$$\frac{d}{dt} E(t) = - \int_0^\ell \alpha_0 |u_{xt}|^2 dx.$$

Assim, para definir o espaço fase considere os seguintes espaços de Hilbert

$$V_1 = H_0^2(0, \ell), \quad V_2 = H_*^2(0, \ell) \cap H_0^1(0, \ell). \\ H_*^2(0, \ell) := \left\{ \omega \in H^2(0, \ell) : \int_0^\ell \omega'(x) dx = 0 \right\}.$$

Denotando por $u_t = v$. Definimos o espaço de fase:

1. Problema (1)-(3) e associado com condições de fronteira do tipo (4):

$$\mathcal{H}_1 = V_1 \times L^2(0, \ell).$$

2. Problema (1)-(3) e associado com condições de fronteira do tipo (5):

$$\mathcal{H}_2 = V_2 \times L^2(0, \ell).$$

Note que $\mathcal{H}_1$ e $\mathcal{H}_2$ são espaços de Hilbert equipado com o produto interno da forma

$$(U, U^*)_{\mathcal{H}} = \int_0^\ell (\rho v \overline{v^*} + \alpha u_{xx} \overline{u_{xx}^*}) dx,$$

onde $U = (u, v)^T$ e $U^* = (u^*, v^*)^T$. Para $j = 1, 2$, o operador $\mathcal{A}_j$, para $\rho = 1$ é dado pela seguinte expressão

$$\mathcal{A} = \mathcal{A}_j = \begin{pmatrix} 0 & I \\ -\alpha \partial_{xxxx} & \alpha_0 \partial_{xx} \end{pmatrix} \text{ para } j = 1, 2.$$

Portanto, o sistema (1)-(3) pode ser reescrito como

$$\begin{aligned} U_t - \mathcal{A}U &= 0, \\ U(0) &= U_0. \end{aligned} \tag{6}$$

com domínio é dado por

$$D(\mathcal{A}_j) = H^4(0, \ell) \cap V_j \times V_j \text{ para } j = 1, 2$$

Por simplicidade de notação, escreveremos $\mathcal{A}$ e $\mathcal{H}$ ao invés de $\mathcal{A}_j$ e $\mathcal{H}_j$ para $j = 1, 2$. Tomemos $F = (f_1, f_2)^t \in \mathcal{H}$. Agora, observando que a equação do resolvente $(i\lambda I - \mathcal{A})^{-1}F = U$ é equivalente a $i\lambda U - \mathcal{A}U = F$, $\lambda \in \mathbb{R}$, temos que em termos das componentes pode ser escrito como

$$i\lambda u - v = f_1 \tag{7}$$

$$i\lambda \rho v + \alpha u_{xxxx} - \alpha_0 v_{xx} = f_2, \tag{8}$$

satisfazendo (4) e (3). Agora, tomando o produto interno da equação do resolvente $i\lambda U - \mathcal{A}U = F$, $\lambda \in \mathbb{R}$ com U , e depois tomando a parte real e usando (11) temos que

$$\int_0^\ell \alpha_0 |v_x|^2 dx = \operatorname{Re} \langle F, U \rangle_{\mathcal{H}}, \tag{9}$$

e usando a desigualdade de Cauchy-Schwarz temos que

$$\int_0^\ell |v_x|^2 dx \leq C \|U\|_{\mathcal{H}} \|F\|_{\mathcal{H}}, \tag{10}$$

onde $C > 0$ é uma constante. Assim temos estimada a norma de v_x .

2 Boa Colocação

A seguir vamos provar a boa colocação do problemas determinado por (1) e (3) e associado com condições de fronteira do tipo (4) ou (5).

Lema 1. *Para os casos 1 e 2,*

1. $\mathcal{A}$ é dissipativo em $\mathcal{H}$;
2. $0 \in \varrho(\mathcal{A})$. Além disso, $\mathcal{A}$ gera um semigrupo C_0 de contrações sobre $\mathcal{H}$.

Demonstração. Para o caso 1 e 2, observe que

$$\operatorname{Re} \langle \mathcal{A}U, U \rangle_{\mathcal{H}} = - \int_0^\ell \alpha_0 |v_x|^2 dx \leq 0. \tag{11}$$

e assim temos que $\mathcal{A}$ é um operador dissipativo. Agora, provaremos que $0 \in \varrho(\mathcal{A})$. Provaremos que para todo $F = (f_1, f_2)^T \in \mathcal{H}$ existe uma única solução $U \in D(\mathcal{A})$ tal que

$$\mathcal{A}U = F.$$

Em termos das componentes temos que

$$-v = f_1 \in H^2(0, \ell) \tag{12}$$

$$\alpha u_{xxxx} - \alpha_0 v_{xx} = \rho f_2 \in L^2(0, \ell). \tag{13}$$

Da primeira equação temos $v = -f_1 \in H^2(0, \ell)$. Portanto de (13) segue que

$$\alpha u_{xxx} + \alpha_0 f_{1,xx} = \rho f_2 \in L^2(0, \ell).$$

Usando o Lema de Lax Milgran para a forma bilinear

$$a(u, w) = \int_0^\ell \alpha u_{xx} w_{xx} dx,$$

encontramos que existe uma única função $u \in H_0^2(0, \ell)$ verificando as equações acima. Note que pela regularidade elíptica temos que $u \in H^4(0, \ell)$. Portanto existe $U = (u, -f_1)^T \in D(\mathcal{A})$ verificando a equação resolvente. Como $\mathcal{A}$ é fechado, dissipativo e $0 \in \varrho(\mathcal{A})$, temos que para algum $\epsilon > 0$, $\text{Im}(\epsilon I - \mathcal{A}) = \mathcal{H}$. Usando o Teorema 4.6 Capítulo 1 pág. 16 of [6] concluímos que $D(\mathcal{A})$ é denso sobre $\mathcal{H}$. Então $\mathcal{A}$ é o gerador infinitesimal de um semigrupo C_0 de contrações. Portanto, segue o resultado. $\square$

Finalmente,

Teorema 1. *Para qualquer $U_0 \in \mathcal{H}$ a solução U satisfaz $U \in C([0, T]; \mathcal{H})$. Além disso, para cada $U_0 \in D(\mathcal{A})$,*

$$U \in C^1([0, T]; \mathcal{H}) \cap C([0, T]; D(\mathcal{A})).$$

3 Estabilidade Exponencial

A nossa principal ferramenta será a caracterização de Pruss que pode ser encontrado em [5], que estabelecemos a seguir

Teorema 2. *Seja $S(t) = e^{At}$ um semigrupo C_0 de contrações sobre um espaço de Hilbert $\mathcal{H}$. Então $S(t)$ é exponencialmente estável, se e somente se,*

$$\varrho(\mathcal{A}) \supset \{i\beta; \beta \in \mathbb{R}\} \quad e \quad \overline{\lim_{|\beta| \rightarrow \infty}} \|(i\beta I - \mathcal{A})^{-1}\|_{\mathcal{H}} < \infty.$$

Lema 2. *O domínio $D(\mathcal{A})$ tem imersão compacta no espaço de fase $\mathcal{H}$.*

O Lema 2 implica que o espectro de $\mathcal{A}$ é formado exclusivamente por autovalores. Isto porque resolvente é um operador compacto.

Lema 3. *Com as notações anteriores temos que $i\mathbb{R} \subset \varrho(\mathcal{A})$.*

Demonstração. Pelo Lema 2 bastará mostrar que não existem autovalores imaginários. De fato, suponhamos que exista um autovalor imaginário, isto é, existe $i\lambda \in \sigma(\mathcal{A})$, com $\lambda \in \mathbb{R}$ tal que

$$\mathcal{A}U = i\lambda U \quad \text{com } U \neq 0.$$

Tomando produto interno na equação anterior com U concluímos que

$$\text{Re } \langle \mathcal{A}U, U \rangle_{\mathcal{H}} = 0.$$

Usando (9) temos

$$\int_0^\ell \alpha_0 |v_x|^2 dx = 0.$$

Logo $v = 0$ o que implica que $u = 0$ sobre o intervalo $]0, \ell[$, ou seja, $U = (u, v) = (0, 0)$ em $]0, \ell[$ o que implica que o autovetor é identicamente nulo o que é uma contradição. Portanto não pode existir autovalores imaginários. De onde segue o resultado. $\square$

Agora, considere

$$\mathcal{I}(x) = \frac{1}{2}(\rho|v|^2 + \alpha|u_{xx}|^2)(x) \quad \text{e} \quad \mathfrak{R} = \|U\|_{\mathcal{H}}\|F\|_{\mathcal{H}} + \|F\|_{\mathcal{H}}^2.$$

Finalmente,

Teorema 3. *O semigrupo definido pela equação (1)-(3) é exponencialmente estável.*

Demonstração. Usando a desigualdade (10) e a desigualdade de Poincaré obtemos

$$\int_0^\ell |v|^2 dx \leq \int_0^\ell |v_x|^2 dx \leq C\|U\|_{\mathcal{H}}\|F\|_{\mathcal{H}}. \quad (14)$$

Pelo Teorema das derivadas intermediárias temos

$$\begin{aligned} \|u_{xx}\|_{L^2} &\leq \frac{c}{|\lambda|}\|v\|_{L^2} + c\|v + f_1\|_{L^2}^{1/2} \left[c\|u_{xx}\|_{L^2}^{1/2} + c\|v\|_{L^2}^{1/2} \right] + \\ &\quad + \frac{c}{|\lambda|^{1/2}}\|v + f_1\|_{L^2}^{1/2} \left[c\|f_2\|_{L^2}^{1/2} + c\|f_{1,xx}\|_{L^2}^{1/2} \right] \end{aligned}$$

Usando a desigualdade triangular e a desigualdade de Young temos

$$\|u_{xx}\|_{L^2} \leq \epsilon\|u_{xx}\|_{L^2} + \epsilon\|v\|_{L^2} + c_\epsilon\mathfrak{R}^{1/2}.$$

E assim temos que

$$\int_0^\ell \alpha|u_{xx}|^2 dx \leq c_\epsilon \mathfrak{R} + \epsilon \int_0^\ell \mathcal{I}(x) dx. \quad (15)$$

Usando a desigualdade acima junto com (14) segue nossa conclusão. $\square$

4 Perda de Analiticidade

A seguir vamos mostrar a perda de analiticidade de uma viga viscoelástico para o problema (1) e (3) com condição de fronteira (5) usando um resultado que enunciaremos abaixo e pode ser encontrado em [5]

Teorema 4. *Seja $\varrho(\mathcal{A})$ o conjunto resolvente do operador linear $\mathcal{A}$. Então, um semigrupo de contrações $S(t) = e^{\mathcal{A}t}$ em um espaço de Hilbert $\mathcal{H}$ é analítico se e somente se $i\mathbb{R} \subset \varrho(\mathcal{A})$ e*

$$\limsup_{|\lambda| \rightarrow +\infty} \|\lambda(i\lambda_n I - \mathcal{A})^{-1}\|_{\mathcal{H}} < \infty.$$

O que permite demonstrar o seguinte resultado:

Teorema 5. *O semigrupo $e^{\mathcal{A}_2 t}$ definido sobre o espaço de Hilbert $\mathcal{H}_2$ é não analítico.*

Demonstração. É suficiente mostrar que existe uma sequência $\{\lambda_n\}$ de números reais e uma sequência limitada U_n em $\mathcal{H}_2$ tal que $\lambda_n \rightarrow +\infty$ e

$$\limsup_{n \rightarrow +\infty} \|\lambda_n(i\lambda_n I - \mathcal{A}_2)^{-1}\|_{\mathcal{H}_2} = +\infty$$

Com efeito, para cada $n \in \mathbb{N}$ tome

$$F_n = \left(0, \cos\left(\frac{n\pi x}{\ell}\right) \right)^T.$$

Seja $U_n = (u_n, v_n)^T \in D(\mathcal{A}_2)$ a solução única da equação resolvente $i\lambda_n - \mathcal{A}_2 U_n = F_n$ com $\lambda_n \in \mathbb{R}$, ou equivalentemente,

$$i\lambda_n u - v = 0, \text{ em } (0, \ell) \times \mathbb{R}^+ \quad (16)$$

$$i\lambda_n \rho v - \alpha u_{xxxx} + \alpha_0 v_{xx} = \text{sen}\left(\frac{n\pi x}{\ell}\right), \text{ em } (0, \ell) \times \mathbb{R}^+. \quad (17)$$

Pelas condições de contorno (5) as soluções do sistema são da forma

$$u_n = A_n \text{sen}\left(\frac{n\pi x}{\ell}\right).$$

Por (16),

$$v_n = i\lambda_n A_n \text{sen}\left(\frac{n\pi x}{\ell}\right).$$

Por (17),

$$\left(\alpha \left(\frac{n\pi}{\ell}\right)^4 - \lambda_n^2\right) A_n \text{sen}\left(\frac{n\pi x}{\ell}\right) + \alpha_0 i \left(\frac{n\pi}{\ell}\right)^2 A_n \text{sen}\left(\frac{n\pi x}{\ell}\right) = \text{sen}\left(\frac{n\pi x}{\ell}\right).$$

Tome $\lambda_n = \sqrt{\alpha} \left(\frac{n\pi}{\ell}\right)^2$ de modo que obtemos

$$A_n = \frac{-i}{\alpha_0} \left(\frac{\ell}{n\pi}\right)^2.$$

Então

$$\|U_n\|_{\mathcal{H}_2}^2 \geq \alpha \int_0^\ell |u_{n,xx}|^2 dx = \frac{\alpha}{\alpha_0^2}.$$

De modo que

$$\|\lambda_n U_n\|_{\mathcal{H}_2}^2 = |\lambda_n|^2 \|U_n\|_{\mathcal{H}_2}^2 \geq \frac{\sqrt{\alpha^3}}{\alpha_0^2} \left(\frac{n\pi}{\ell}\right)^2$$

Fazendo $n \rightarrow +\infty$ o resultado segue. □

5 Conclusões

Neste trabalho mostramos existência e unicidade de soluções para dois modelos de vigas viscoelásticas bem como a estabilidade exponencial para ambos problemas. Note que mostramos a perda de analiticidade apenas para o problema determinado por (1)-(3) e associado com condições de fronteira do tipo (5). Acreditamos que o primeiro problema não seja analítico, mas ainda é uma questão em aberto.

6 Agradecimentos

Este trabalho foi desenvolvido sobre orientação do Prof. Jaime M. Rivera como parte da tese de doutorado dos autores pelo Instituto de Matemática da UFRJ.

Referências

- [1] ADAMS, R. A. *Sobolev Spaces*, Academic Press, New York 1975.
- [2] BREZIS, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer-Verlag, Berlin (2011).
- [3] CHRISTENSEN, R. M. (1982). *Theory of Viscoelasticity*. Elsevier. <https://doi.org/10.1016/B978-0-12-174252-2.X5001-7>
- [4] LAKES, R. S. (2009). *Viscoelastic Materials*. Cambridge University Press.
- [5] LIU, Z. & ZHENG, S. *Semigroups associated to dissipative systems*. Chapman e Hall/CRC Research Notes in Mathematics, 398, (1999).
- [6] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations* (Vol. 44). Springer New York. <https://doi.org/10.1007/978-1-4612-5561-1>

Desenvolvimento de Software para Dimensionamento e Seleção de Estruturas de Edifício em Concreto Armado com Redes Neurais Artificiais Otimizadas e AHP

Adriano Ricardo Borges Costa

UFF -Universidade Federal Fluminense - EEIMVR

adborgvr@gmail.com

Eliane da Silva Christo

Universidade Federal Fluminense

elianechristo@id.uff.br

Tiago de Araujo Neves

Universidade Federal Fluminense

tneves@id.uff.br

Resumo: A Engenharia evoluiu com a Computação e agora, algoritmos de aprendizado de máquina (ML) como as Redes Neurais Artificiais (RNAs) também vêm sendo utilizados. As RNAs vêm ajudando na resolução de problemas com grande massa de dados e de média à alta complexidade. Os engenheiros projetam edifícios e estes estão ficando cada vez mais complexos no design e em exigências estruturais. Dimensionar e orçar as opções de projetos que, mesmo usando métodos computacionais convencionais, podem ser tarefas complexas envolvendo muitos recálculos e análises, logo, a Engenharia é uma das áreas ideais na aplicação de RNAs. Depois da geração das opções dos edifícios focando "resistência x suficiência", selecionam-se os melhores projetos para o cliente focando na "qualidade x menor custo". Muitas revisões jogam o engenheiro contra os prazos de aprovação, contratação e a construção da obra. Este trabalho visa gerar um protótipo de software para engenharia com o combinado RNA + AHP e confirmar seu potencial. O protótipo dimensiona as armaduras das lajes de projetos via RNAs e os orça via AHP. Os resultados foram bastante promissores e atenderam aos objetivos, mostrando que RNAs e AHP podem ser mais uma relevante combinação na área de projetos de obra.

Palavras-chave: RNA. AHP. Projeto de Edifício. Concreto Armado. Perceptron Multicamada.

1. Introdução

A Engenharia utiliza a Computação em projetos há décadas e, como é sempre necessário criar n propostas de projeto (Fig.2), dimensioná-las demanda numerosos recálculos e análises, tornando-se tarefa complexa e repetitiva (Botelho, 2013). Edifícios têm muitos elementos a serem dimensionados e orçados como as lajes (Fig.1), que têm armaduras principais positivas e negativas e suas secundárias nos pavimentos tipo, sem falar nas vigas e pilares (Uberaba, 2009). Surge então, a proposta de um software para dimensionar / orçamento através de RNA e AHP. A arquitetura de RNA adotada foi a perceptron multicamada (PMC). Com a topologia e quantidade certa de neurônios, uma parametrização e funções apropriadas (Gráfs. 3 e 4), e valores corretos dos aceleradores de aprendizado a RNA gera uma função-solução através da sua estrutura (Kovács, 1996). Vários pré-treinamentos são feitos para se chegar à melhor topologia e parâmetros, (Silva, 2010) e este trabalho, além de produzir o software, também testou na prática a teoria da liga RNA-AHP, no cálculo e orçamento das armaduras.

2. Dimensionamento e Orçamento de Estrutura de Edifícios

Os engenheiros projetam edifícios para clientes baseados no dimensionamento e orçamento da estrutura e acabamento. Basicamente os requisitos são: Orçamento do cliente; Levantamento topográfico; Sondagem do Terreno; Cartas ambientais locais. Após os requisitos, são geradas plantas

de pré-projeto, com várias opções de estrutura e design e seus orçamentos (Fig.2) para posterior aprova do cliente. A complexidade, volume de dados e revisões dependerá dos projetos.

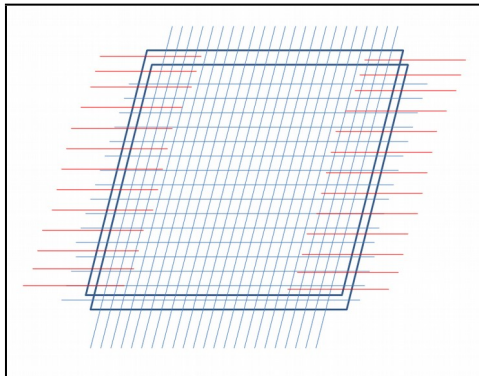


Fig.1: Ex. Distrib. (simplif.) Armad. Laje

Planta Baixa do Projeto 6														
	4.0		10.0		13.0		18.0		21.0		27.0			
4.0	6.0x4.0 Quarto		3.0x4.0 Copa		3.0x4.0 Sala		5.0x4.5 A.Serv		Swite		3.0x4.0 Sala	6.0x4.0 Copa	Quarto	
	Coz.				Coz.		WC				Coz			
8.0	6.0x4.5 Swite		Coz		3.0x4.5 Escada / Hall		5.0x4.5 Sala		Copa		3.0x4.5 Escada / Hall		6.0x4.5 Coz.	Swite
	Wc		A.Serv		Vent						Vent		A.Serv	Wc
12.5	6.0x4.5 Copa		Sala		3.0x4.5 Hall /Elevador		5.0x4.5 Copa		Sala		3.0x4.5 Hall /Elevador		6.0x4.5 Sala	Copa
	Wc				WC						WC			Wc
17.0	6.0x4.0 Quarto		Coz.		3.0x4.0 A.Serv		5.0x4.0 Swite		Coz.		3.0x4.0 A.Serv		6.0x4.0 Coz.	Quarto
							WC		A.Serv.					
21.0														

Fig.2: Ex. Planta Baixa (simplif.) de um Projeto

3. Redes Neurais Artificiais - Resumo da Estrutura e Funcionamento

RNAs são tipos de algoritmo para tratar problemas complexos, sendo uma malha de entradas/processadores/saídas (Figs. 3 e 4) formando uma rede matricial de camadas interconectando dados e equações. Os processadores abstraem neurônios biológicos. Problemas podem ser tratados com 1 ou várias RNAs orquestradas por aplicações e seu ganho está na velocidade dos resultados aproximados. Os neurônios absorvendo a lógica das equações que regem todo ou parte do problema, (Ludwig, 2007). Os parágrafos a seguir mostram seu funcionamento.

As camadas "pastoreiam" os dados injetados até a camada de saída via hiperplanos ("cerças") (Silva, 2010) nos domínios do problema. Os neurônios de uma camada processam as entradas $x_j^{(k)}$ da anterior gerando saídas $y_j^{(k)}$ para a próxima (Fig.3). Cada neurônio processa entradas via dois processadores internos na etapa Free-forward (arranque à frente) que são o "somador" de entradas e pesos ($x_j^{(k)} \cdot w_{ji}^{(k+1)}$) (Eq.1 e Fig.3) e o "limiar de ativação" (Eq.2 e Fig.3-azul) que aplica uma função sobre essa soma. Na 2ª etapa, o Back-Propagation é o reverso (retro-propagação) usando os processadores, "avaliador" (Fig.4-azul) que calcula nos neurônios da última camada o erro ($d_j^n - y_j^{(k)}$) entre as saídas esperadas e calculadas, e o "difusor de erros" (Fig.4-azul) que distribui os quinhões de erro aos neurônios da camada e pulveriza os erros camada por camada anterior conforme os valores dos seus pesos $w_{ji}^{(k-1)}$, até chegar aos pesos da 1ª camada intermediária (Azevedo, 2000).

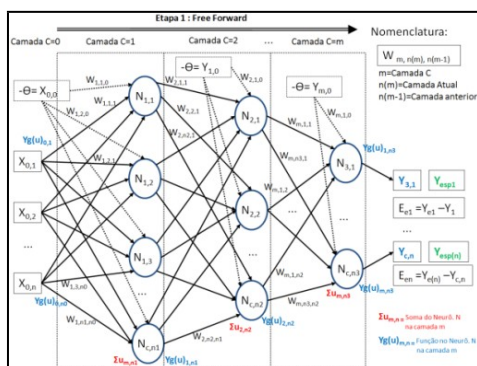


Fig.3: Arranque a frente Free-Forward.

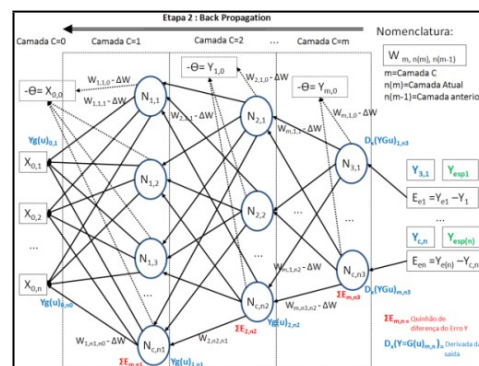


Fig.4: Ajuste Reverso de Retro-Propagação

Segunda Silva (2010), para que a RNA funcione corretamente, os valores reais de entrada precisam ser normalizados, reajustados à faixa de domínio da função de ativação, antes da Free-forward e da injeção das entradas das amostras, convertendo-se suas amplitudes reais para às das funções de ativação (Graf.1 e 2), se Tg. Hiperbólica (-1,1) ou se Logística (0,1), Moraes (2010). Segundo Silva (2010) e Ludwig (2007), a Free-forward é alimentada com amostras normalizadas $x_i^{(k-1)}$ injetadas na 1ª camada, em sequência ou aleatoriamente (Manzan, 2016). A seguir, o Back-

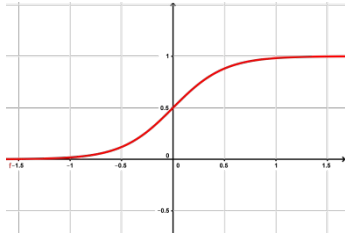
Propagation calcula o erro delta $\delta_j^{(k)}$ (Eq.3) entre saída esperada $d_j^{(n)}$ e calculada $y_j^{(k)}$. O difusor de erros reajusta os pesos $w_{ji}^{(k)}$ (Eq. 4) de cada neurônio j da camada k de conexão entre as entradas i de camada k em camada $k-1$ (Eqs. 3 e 4) somando parcelas η do gradiente $\delta_j^{(k)} y_j^{(k)}$. Os pesos $w_{ji}^{(k)}$ são reajustados de um Δw proporcional aos pesos de épocas t sucessivas ($w_{ji}^{(k)}(t+1) - w_{ji}^{(k)}(t)$) (Eq.4).

$$u_j^{(k)} = \sum_{i=1}^n (w_{ji}^{(k)} \cdot x_i^{(k-1)} - \Theta); y_j^{(k)} = g(u) \text{ com } u \in \mathbb{R} \text{ e } y \in \mathbb{R}(1) \quad (2)$$

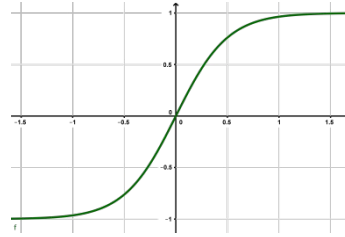
Sendo: $w_{ji}^{(k+1)}$ = peso da camada $k+1$, x_i = entrada, $g(u)$ = função de ativação.

$$\delta_j^{(k)} = (d_j^{(n)} - y_j^{(k)}) g'(l_j^{(k)}); w_{ji}^{(k)}(t+1) = w_{ji}^{(k)}(t) + m \Delta w + \eta \delta_j^{(k)} y_j^{(k-1)} \quad (3) \quad (4)$$

Sendo: $w_{ji}^{(k)}(t+1)$ = peso, $\delta_j^{(k)}$ = ErroDelta camada k , $g'(l_j^{(k)})$ = derivada da equação 2



Graf.1: Função Logística: $g(u) = \frac{1}{1 + e^{-\beta \cdot u}}$



Graf.2: Função Tangente Hiperb.: $g(u) = \frac{1 - e^{-\beta \cdot u}}{1 + e^{-\beta \cdot u}}$

A cada iteração t , é recalculado o erro quadrático médio $E_{(k)}$ (Eq.5 e 6) de cada saída $y_j^{(n)}$ nos neurônios da última camada (Eq.7) (Moraes, 2010), sendo acumulados no erro médio geral de todas as amostras E_M (Eq.7). Na retro-propagação da iteração $t+1$, o E_M é fragmentado em vários $\delta_j^{(k)}$ (Eq.3) e pulverizado nos processadores neurais e, após p iterações. O treino acaba quando a diferença do erro $\Delta E_{(n)}$ (Eq.7) de 2 iterações sucessivas atinge o erro ΔE_{adm} configurado, e as saídas $y_j^{(n)}$ são convertidas para valores reais (pós-processamento).

$$E_{(k)} = \frac{1}{2} \sum_{j=1}^n [d_j^{(k)} - y_j^{(k)}]^2; \Delta E_{(n)} = E_{(n)}(t+1) - E_{(n)}(t); E_M = \frac{1}{p} \sum_{k=1}^n [E_{(k)}] \quad (5) \quad (6) \quad (7)$$

Sendo: $d_j^{(k)}$ = saída esperada, $y_j^{(k)}$ = saída calculada cam. k , p = nº de iterações

Segundo Silva (2010) e Azevedo (2000), e comprovado na prática, os parâmetros da RNA precisam ser pré-ajustados em treinos iniciais, pois atuam no percentual e velocidade da convergência. Destacam-se 5 parâmetros que influenciam na convergência: 1º) **Nº de Atributos Entrada (significativos)**: RNAs exigem certa quantidade de atributos no domínio do problema, Manzan (2016). 2º) **Função de Ativação**: Funções adequadas geram melhor convergência (Manzan, 2016), o programa Orange Canvas^(TM)T comprova que Tg.Hip. é melhor que a Logística. 3º) **Nº de Neurônios nas Intermediária(s)**: O nº de neurônios/camada é empírico e os ajustes dos parâmetros de execução vão da experiência do pesquisador, Manzan (2016) e Ludwig (2007), mas em Silva (2010) há 3 técnicas para o cálculo: Witten-Flank (Eq.7), Fletcher-Gloss (Eq. 8) e Kolmogorov (Eq.9).

$$n_1 = \frac{n + n_c}{2}; 2\sqrt{n} + n_2 < n_1 < 2n + 1; 2n + 1 \quad (7) \quad (8) \quad (9)$$

Sendo: $n = n^{\circ}$ entradas, $n_1 = qt.$ Nros na cam. interm., $n_2 = qt.$ Nros na cam. saída, $n_c = n^{\circ}$ classes

4º) Faixa dos Aceleradores: Os aceleradores de aprendizado η e de direcionamento momentum m são importantíssimos, Manzan (2016). O acelerador η atua no gradiente $(\eta \delta_j^{(k)} y_j^{(k-1)})$ como bússola direcionando a convergência no caminho oposto ao sinal do gradiente (Silva, 2010). A 2ª parcela momentum $(m \Delta w)$ atua na velocidade da convergência até o erro mínimo, Manzan (2016). O erro Δw é repassado aos pesos $w_{ji}^{(k)}$ das camadas a cada Back-Propagation Ludwig (2007). Essas parcelas levam o EQM ($E_{(k)}$) entre saídas (Eq. 5) para o mínimo de sua variação $\Delta E_{(k)}$ (Eq. 7) entre iterações, Silva (2010). **5º) Convergência Mínima x Nº de Iterações:** Na prática é preciso um mínimo de iterações para boas convergências, evitando o "underfitting" (baixa convergência por parada antecipada), ou o "overfitting" (degradação da convergência por excesso de iterações), Silva (2010). Ex.: Altas convergências são demoradas e Baixas são rápidas. **6º) Erro Médio Quadrático:** Haverá necessidade de muitos treinos, pois que vários treinos são naturalmente descartados por se desviarem da solução, e segundo Neves (2021) algoritmos meta-heurísticos não garantem a obtenção da solução.

4. Apoio Multicritério com AHP (Analytic Hierarchy Process)

Orçar n propostas de projetos é complexa e demorado devido às muitas variações em "resistência x menor custo". Há várias técnicas de apoio a orçamento e uma delas é a Apoio à Decisão por Multicritério (ADMC) desenvolvida nos 90's por Saaty. O objeto de escolha (proposta) possui alternativas e critérios (Fig.5), Gomes (2009) podendo conter riscos (probabilidade) ou não (dominam-se as variáveis do cenário em questão). Cria-se então, uma tabela hierárquica iniciando pelo objetivo (1º nível) as alternativas (2º nível) e os critérios (3º). Os n critérios são as características comuns às alternativas, valorados conforme as mesmas. Os critérios são normalizados (0 a 1) para o mesmo domínio tornando-se utilidades. Esses também recebem um peso (relevância %) definido por especialistas e calculado na tabela cruzada 10 (Eqs. 10 e 11). O objetivo é então calculado (Eq.11) para cada alternativa. A coluna Gan-Proj indicará a alternativa objetivo de maior ganho e equilíbrio agregado. Esse tipo de cálculo é chamado de ADMC com AHP, (Hernandez, 2021).

	PCstConst	PNvAccto	PMgLucr	PROI	Crit	PCstConst	PNvAccto	PMgLucr	PROI		PesosNrm
PCstConst	1	0,5714286	1,14286	1,2857143	4	0,23	0,23	0,23	0,23	PCstConst	0,23
PNvAccto	1,75	1	2	2,25	7	0,40	0,40	0,40	0,40	PNvAccto	0,40
PMgLucr	0,875	0,5	1	1,125	3,5	0,20	0,20	0,20	0,20	PMgLucr	0,20
PROI	0,7777778	0,4444444	0,88889	1	3,1111111	0,18	0,18	0,18	0,18	PROI	0,18
	4,4027778	2,515873	5,03175	5,6607143	17,611111						

Tabela.1: Ex.Matriz de Calculo dos Pesos - Fonte: Autor , Base Ref. Hernandez (2021).

$$P c_{ij} = P c_{ij} : \sum_{i=1}^n \left[\sum_{j=1}^n (P c_{ij}) \right]; P N c_i = M é d i a G e o m (P c_{ij}) ; G a P j_i = \sum_{i=1}^n (P c_{ij} \cdot U c_{ij}) (10) (11) (12)$$

Proj.	CustConstru	UCstConst	PCstConst	NvAccto	UNvAccto	PNvAccto	MgLucr	UMgLucr	PMgLucr	ROI	UROI	PROI	Gan - Proj
8	1.957.311,70	0,48	0,23	1,05	0,33	0,40	0,72	0,32	0,20	1.577.593,23	0,50	0,18	0,39
9	990.230,46	0,74	0,23	1,08	0,34	0,40	0,78	0,34	0,20	913.388,58	0,29	0,18	0,42
10	836.619,46	0,78	0,23	1,03	0,33	0,40	0,77	0,34	0,20	688.621,47	0,22	0,18	0,41
	3.784.161,61	Peso->	0,7	3,16	Peso->	0,4	2,27	Peso->	0,8	3.179.603,28	Peso->	0,9	

Tabela.2: Ex.Matriz de Decisão Multicritério - Fonte: Autor, Base Ref. Hernandez (2021).

No caso da aferição de propostas de edificio com no exemplo da tabela 1, mostra que, se apenas o peso do critério ROI for relevante para o investidor, a alternativa "Projeto 8" seria a escolha para o mesmo aplicar seu capital, mas, observando todo o conjunto de critério via técnica AHP, a alternativa "Projeto 9" é a de ganho mais equilibrado em relação a todos os critérios pesados na escolha.

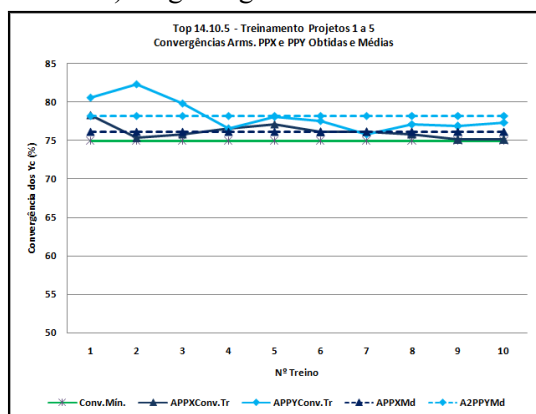
5. Métodos

O software possui os módulos em destaque: 2.1 (cadastro de projetos); 2.2 (mapeamento de plantas baixas); 3.1 (dimensionamento convencional); 3.2 (dimensionamento via RNA - pseudo-códigos sugeridos por Silva (2010) e Ludwig (2007)); 3.3 Orçamentos via AMD-AHP. Para a injeção das lajes, elencaram-se 14 atributos significativos do cálculo das armaduras. Para o treino / teste mapearam-se 10 projetos (5 p/ treinos e 5 p/ testes). Dimensionadas as lajes dos projetos (Ex.: Fig.2) no módulo 3.1 gerando os datasets para o 3.2, onde as dos projetos de 1 a 7 foram injetadas considerando-se 3 pavimentos (Térreo, Garagem, Pavto Tipo) num dataset com $(36+36+75+75+60+60+60)=402$ lajes para treino e $(60+75+75)=210$ para teste. A topologia tem 5 saídas y binárias (1 Byte de 5 bits) formando 6 diâmetros (Classes) comerciais de aço (5,6,8,10,12,16 mm) a serem obtidos pelo quinteto de neurônios da camada final, assim, cada Byte 1 diâmetro de 1 armadura de uma laje (+ e - nas direções x e y) de 1 pavimento tipo (Fig.1 e 2). A complexidade está na variação das lajes quanto à posição, dimensão, quantidade, carga e área/esp. de aço. Executadas ~560 prospecções treino/teste, conforme Silva (2010) só para encontrar as topologias e a parametrizações de convergências ideais.

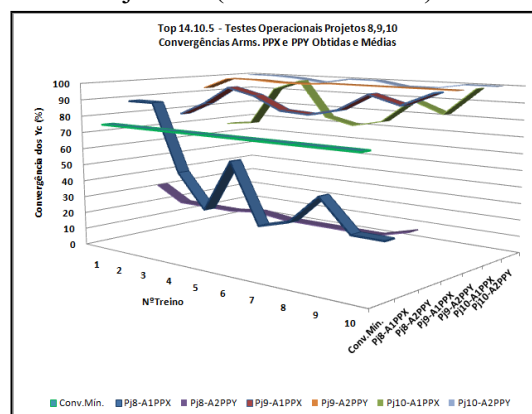
Topologias e Parâmetros adotados: 1) **Faixa Topológica melhor Convergência:** (8, 10 e 12) e adotada a Top.: 14-8-5 de maior %, conforme (Eq.7) a técnica de Witten-Flank; 2) **Atributos de Entrada:** Posições das 4 vizinhas, Tp.cômodo, Lado x, Lado y, Carga total/m², Altura; Direções das barras, Fck, Fyd, Área e espaçamento das barras; 3) **Função de Ativação:** Tangente Hiperbólica na camada intermediária (Fig.4); 4) **Faixas dos Aceleradores:** η (0,25 a 0,10) e m (0,15 a 0,05), valores acima destes causaram erro 0/0. Adotados $\eta=0,1$ e $m=0,05$ conforme prospecções prévias e sugestões de Silva (2010) Abdalla (2013); 5) **Convergência % Mín. / nº Máx. de Iterações:** 80% de convergência em até 100 iterações. 6) **Erro Médio Quadrático (EQM):** $\leq 10^{-2}$. O processamento teve 4 etapas (pré-processamento; free-forward; back-propagation; pós-processamento), Silva (2010). Dimensionadas as lajes via RNA, os resultados foram repassados ao módulo 3.3 onde as mesmas foram compor o orçamento dos projetos de 8 a 10 para a obtenção do mais vantajoso via tabelamento de seus critérios (Área/pavimento; Quantidade de material; Custo) via AMD.

6. Resultados

Foram testados 3 conjuntos de projetos, 1 (4 Treino / 1 Teste), 2 (5 Treino / 2 Teste) e 3 (7 Treino / 3 Teste). Segue alguns dos resultados obtidos com o conjunto 3 (Gráficos de 3 a 6):

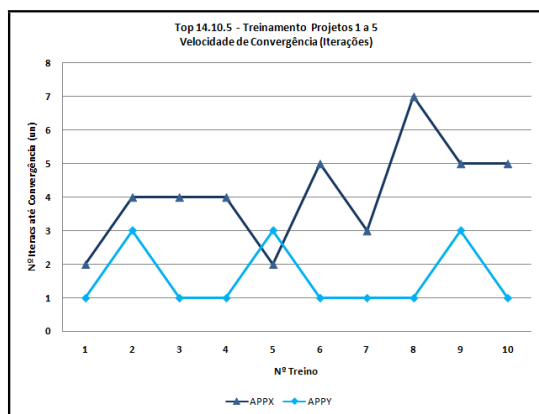


Gráf. 3: Ex. Treino Conv. Armads. PPX e PPY.

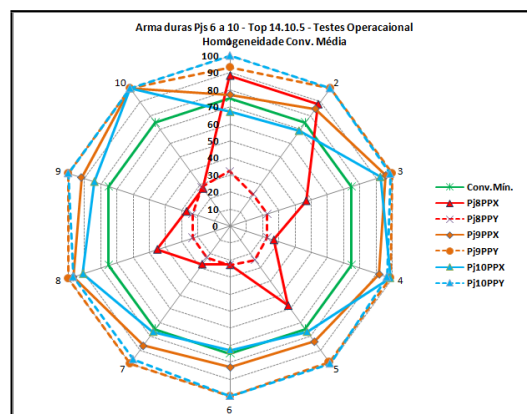


Gráf. 4: Ex. Testes Converg. Armad. PPX e PPY

O média de losts (descartes) nos treinos / testes foi de 2% para as armaduras principais positivas, e 70% para as principais negativas, justificando esse percentuais até mesmos com as velocidades de convergência, uma vez que a facilidade com que quanto mais variações de diâmetros estão nas amostras mais lenta é a RNA para convergir para altos percentuais de acertos entre treino e teste. Foi confirmado também que amostras com mais variações de diâmetros, mais neurônios são necessários.



Gráf. 5: Ex. Veloc. Conver. Armad. PPX e PPY



Gráf. 6: Ex. Veloc. Conver. Armad. PPX e PPY

7. Conclusões, Discussões e Trabalhos Futuros

Para o conjunto 3 de edifícios com os projetos 1 a 7 para treino, e 8 a 10 para testes, o algoritmo conseguiu treinar as convergências além da mínima (75%) da armadura direção x (PPX) em 76,2% e direção y (PPY) em 78,2% (Graf.3) o que demonstrou excelentes convergências para as armaduras positivas sendo a PPY mais rápida que a PPX com 70% das iterações entre 1 e 2 (Graf.4). Os testes operacionais (Graf.4), com base nos pesos w dos 10 treinos do conjunto 3 dos projetos de 1 a 5 (Gráf.3), geraram convergências médias Pj8 (PPX=45% PPY =24,6), (Pj9 PPX=89% PPY =99%) e (Pj9 PPX=83% PPY =99%) demonstrando que a RNA pode aprender a dimensionar as armaduras positivas x e y do conjunto 3, mas não conseguiu no projeto 8. Observa-se que, treino com conjuntos de projetos gerando matrizes com algum grau de heterogeneidade (numéricas e posicionais) entre as amostras pode dificultar o aprendizado da RNA, pois podem pontuar divergências nas aproximações matemáticas como a dos meta-heurísticos, gerando baixo aprendizado como, (Silva, 2010) (Kovács, 1996). Mesmo com poucas iterações uma RNA pode aprender se o treinamento for feito com expressivo número de amostras ($36+36+75+75+60+60+60=402$ lajes) como comprovado no gráfico 5. O gráfico de radar 6 de homogeneidade de mostra que nos projetos de 9 e 10, 80% dos testes convergiram além dos 75%, mesmo o treino tendo exigido apenas 75%, e, além disso, as velocidades de convergência atingiram 90% na faixa de 3 a 5 iterações para a PPX e 100% na faixa de 3 a 5 para a PPY respectivamente, o que mostra que o funcionamento do algoritmo está atingindo o objetivo (convergência elevada e rapidez), pois as entradas foram baseadas nos cálculos convencionais apresentando alta fidelidade com os cálculos reais de literatura de engenharia.

As armaduras Negativas não tiveram ainda o mesmo desempenho que as positivas, para o conjunto 3 de projetos, com convergências na ordem de 60% para as PNX e 55% para as PNY respectivamente, e com velocidades de iteração entre 22 e 25, muito mais lentas que as positivas. A maior heterogeneidade do conjunto 3 pode ser a causa da baixa performance. Comparando com programas como o Orange Canvas ^(c), este também conseguiu convergências de ~80% a ~90% com os mesmos *datasets*, confirmando que o software tem potencial mesmo ainda está recebendo melhorias tais como, aumentar a distinção entre diâmetros de armaduras positivas e negativas para atingir suas maiores convergências e encontrar os melhores atributos refletindo mais das equações do dimensionamento no meio e extremidade de vãos das lajes. As médias de losts (nº de treino/teste invalidados) para as armaduras positivas foram baixos (em torno de 2), mas para as armaduras negativas o mesmo se mostrou alto (5 a 6 por treino/teste). Para as armaduras positivas, o algoritmo aprendeu mais rápido, mas para as negativas foi lento, onde o maior redutor foi o alto grau de heterogeneidade das lajes do conjunto 3 de projetos, reduzindo ou limitando o aprendizado a baixas convergência e velocidade. Esta limitação pode estar ligada também à falta de um atributo para o algoritmo distinguir melhor entre o que são armaduras positivas e negativas. A parte final orçamentária por ADMC-AHP ainda está em desenvolvimento, não havendo resultados disponíveis

para julgamento ou discussão neste momento. É claro que o desempenho do algoritmo de RNA do software pode ser melhorado e para trabalhos futuros, sugere-se seu upgrade e também estender a combinação RNA+AHP para as vigas.

Agradecimentos

Agradeço à orientadora Eliane da Silva Christo, ao programa MCCT da EEIMVR-UFF pela oportunidade de participação no ERMAC, a Camila da secretaria MCCT e aos pesquisadores.

Referências

- ABDALLA, L. S.; VOLOTÃO C. S. Estudo da configuração de diferentes arquiteturas de redes neurais artificiais MLP para classificação de imagens ópticas. In: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, 13 a 18 de abril de 2013, INPE. 2013. p. 8200-8207.
- AZEVEDO, F. M.; BRASIL, L. M.; OLIVEIRA R. C. L. **Redes Neurais com Aplicações em Controle e em Sistemas Especialistas**. 1ª. ed. Florianópolis: Visual Books, 2000.
- BOTELHO, M. H. C.; MARCHETTI O. **Concreto Armado em Ti Amo - Volume I**. 7ª. ed. São Paulo: Editora Blucher, 2013.
- FERNANDEZ, J. C. **Uso de Redes Neurais Artificiais e AHP na Redução da Subjetividade na Seleção de Áreas com Potencial para Coleta de Água da Chuva**. 2020. 68 f. Dissertação (Mestrado em Engenharia Civil)–UFU, Uberlândia - MG, 2020.
- GOMES, L. A. M.; GOMES C.F.S.; ALMEIDA A.T. . **Tomada de Decisão Gerencial - Enfoque Multicritério** - 3ª. ed. São Paulo: Editora Atlas, 2009.
- HERNÁNDEZ, C. T. **Aula 24 - Auxílio Multicritério à Decisão**. Aulas da Disciplina de Análise Multicritério. Universidade Federal Fluminense, Volta Redonda, 2021. p.31.
- KOVÁCS, Z. L. **Redes Neurais Artificiais - Fundamentos e Aplicações**. 2. ed. São Paulo: Editora Collegium Cognition e Edição Acadêmica, 1996.
- LUDWIG JR, O. L.; MONTGOMERY. E. **Redes Neurais Artificiais para engenharia e ciências aplicadas**. 1. ed. São Paulo: Editora Ciência Moderna, 2007.
- MANZAN, J. R. G. **Análise de Desempenho de Redes Neurais Artificiais do Tipo Multilayer Perceptron por meio do Distanciamento dos Pontos do Espaço de Saída**. 2016. 129 f. Tese (Doutorado em Engenharia Elétrica)–UFU, Uberlândia - MG, 2016.
- MORAES, E. C. **Reconhecimento de Padrões e Redes Neurais Artificiais em Predição de Estruturas Secundárias de Proteínas**. 2016. 135 f. Tese (Doutorado em Ciências e Engenharia de Sistemas de Computação)–UFRJ, Rio de Janeiro - RJ, 2010.
- NEVES, T. S. **Inteligência Computacional Aplicada**. 2021. Disponível em: Aulas da disciplina ICA do curso do programa de mestrado MCCT da EEVR- UFF.
- SILVA, I. N.; SPATTI D, H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas**. 1ª. ed. São Paulo: Artlilber Editora, 2010.
- SPERANDIO,D.; MENDES. J. T.; SILVA L. H. M. **Cálculo Numérico**. 1ª.ed. São Paulo: Pearson Prentice Hall, 2003.
- UBERABA, L. R. H. **Projeto de Estruturas de Concreto Armado - Projetos de Estruturas de Edifícios à Luz da Informática**. 1ª. ed. São Paulo: Editora e Gráfica Publi, 2009.

MODELAGEM 3D DE CONCRETO SOBRE ESTRESSE TÉRMICO APLICANDO O MÉTODO DOS VOLUMES FINITOS

João Victor de Mello Marinho

*Escola de Engenharia Industrial Metalúrgica - Universidade Federal Fluminense
joaovmma@id.uff.br*

Victor de Oliveira Carvalho

*Escola de Engenharia Industrial Metalúrgica - Universidade Federal Fluminense
victoroliveiracarvalho@id.uff.br*

Jorge Juárez Trujillo

*Escola de Engenharia Industrial Metalúrgica - Universidade Federal Fluminense
jorgejuareztrujillo@id.uff.br*

José Adilson de Castro

*Escola de Engenharia Industrial Metalúrgica - Universidade Federal Fluminense
joseadilsoncastro@id.uff.br*

Resumo: O concreto civil de alta performance e os castables refratários apresentam ambos o problema de lascado térmico, em vezes violentamente explosivo. Na construção civil, esse fenômeno pode se apresentar em casos de incêndio, aumentando os riscos para sua mitigação. Por sua vez, os monólitos refratários industriais apresentam esse risco durante sua instalação, no primeiro aquecimento, onde precisa-se da definição de uma curva de temperaturas ajustada, a fim de eliminar a umidade presa nos poros o mais rápido possível. Atualmente na indústria, essas curvas são desenvolvidas principalmente de forma empírica. O presente estudo propõe a aplicação do método numérico dos volumes finitos para simular esse comportamento por computador, podendo ajustar os parâmetros envolvidos, seja para um concreto civil ou refratário. A solução é obtida através da linearização de um sistema de duas equações diferenciais parciais parabólicas. O modelo conseguiu descrever tridimensionalmente o comportamento transiente de um material cimentício submetido a estresse térmico, resultando valores aceitáveis das temperaturas internas, das pressões dos poros e da eliminação da água livre até a temperatura de 450°C.

Palavras-chave: Fenômenos de transporte. Método dos volumes finitos. Materiais cimentícios. Meio poroso.

1. Introdução

O concreto é uma mistura de cimento e agregados como pedra brita e areia sílica, amplamente utilizado em diversos setores como construção civil, túneis, ferrovias, usinas siderúrgicas e nucleares. Ao longo de recentes décadas, estudos na mistura de seus componentes desenvolveram técnicas como a otimização de tamanho de partícula, redução do conteúdo de água e o emprego de aditivos como o fumo de sílica que resultam em ganhos de propriedades como densidade, resistência mecânica e propriedades reológicas (MILLARD, 2019). Contudo, tais concretos, denominados concretos de alta performance, podem apresentar o fenômeno de lascado térmico ou explosivo, quando expostos ao estresse térmico. Acredita-se que é devido principalmente às propriedades de permeabilidade e condutividade térmica oriundas das novas misturas, podendo resultar em eventos catastróficos como aconteceu no incêndio do túnel da Mancha em 1996 e 1998 (ULM, 1999) e em acidentes industriais (PALMER et al., 2014).

Segundo Mindeguia (2010), as etapas de fragmentação do concreto podem ser definidas como: aquecimento rápido, que gera altas tensões unilaterais, eventualmente superando a resistência mecânica

do material, mudanças rápidas de volume e estrutura dos agregados e pressão interna gerada pela liberação de vapores de água e gases do agregado e da massa de cimento que ficam presos.

No setor siderúrgico, tem-se desenvolvido misturas para concretos refratários, capazes de sustentar altas temperaturas de trabalho para serem utilizados em ambientes extremos, por exemplo, nos canais de corrida dos altos fornos para transportar a escória e o metal fundido até os carros torpedos. Para evitar o aparecimento dos fenômenos responsáveis pela fragmentação do concreto, são desenvolvidas curvas de aquecimento gradual (TRUJILLO, 2020).

Atualmente, foram feitos grandes esforços para mensurar experimentalmente as propriedades envolvidas e a partir desses resultados propor relações constitutivas entre os parâmetros, com o intuito de prognosticar o comportamento em altas temperaturas do concreto, entendido como um meio poroso, mas, com baixa permeabilidade onde deve fluir vapor de água. Kalifa (2000) realizou medições experimentais de pressão, temperatura e perda de massa, aquecendo amostras de concretos de alta performance. Trujillo et al. (2023) desenvolveram uma solução computacional aplicada em concretos refratários, utilizando o método dos volumes finitos num domínio 3D (tridimensional) semelhante ao corpo de prova utilizado no experimento de Kalifa. Para adquirir maior conhecimento sobre o fenômeno, sem altos gastos, é possível desenvolver simulações numéricas, testando diferentes curvas de aquecimento em concretos com diferentes propriedades intrínsecas, como permeabilidade e condutividade térmica. Assim, este artigo busca aplicar o modelo 3D resolvido com o método dos volumes finitos desenvolvido por Trujillo et al. (2023), com uma curva de aquecimento até 450°C para simular o comportamento dos campos de temperatura, pressão interna dos poros e variação de massa num bloco de concreto aquecido.

2. Métodos Matemáticos e Numéricos

2.1 Modelo matemático

Este trabalho baseia-se no modelo matemático proposto por Bazant e Jirásek (2018), que consiste num sistema de duas equações diferenciais parciais parabólicas (EDP) que representam a conservação de massa (vapor de água) e a energia. Suas incógnitas são pressão P e temperatura T . O sistema representa o fluxo monofásico transiente de vapor de água num meio poroso que sofre alterações de permeabilidade em função do aumento de temperatura. A solução do sistema de equações é determinada através do método dos volumes finitos que converte as EDPs em equações lineares (CHAPRA, 2010) resolvidas por processos iterativos através de um código de programação em FORTRAN desenvolvido por Trujillo et al. (2023).

2.2 Equações de conservação

A conservação da massa considera as variações da água livre e da água quimicamente ligada que são liberadas durante o aquecimento, acordo a equação (1). A de Darcy serve para descrever o transporte de umidade gerada pelo gradiente de pressão dentro dos poros de concreto, acordo a equação (2).

$$\frac{\partial w}{\partial t} = -\nabla \cdot J + \frac{\partial w_d}{\partial t} \quad (1)$$

$$J = \frac{-a}{g} \nabla P \quad (2)$$

Sendo W_d a água liberada pela desidratação, W é a quantidade de água livre, em kg/m^3 , dependente da umidade h relativa dos poros, g é a gravidade atmosférica, a é a permeabilidade relativa do concreto em m/s , e J é o fluxo de vapor de água proporcional ao gradiente de pressão ∇P .

Já na equação do balanço de energia considera as entalpias do sistema. E a lei de Fourier para condução de calor via gradiente de temperatura e convecção por massa de vapor. Como o aquecimento envolve transporte de massa e calor, a conservação de massa é complementada com a equação do balanço de energia.

$$\rho C \frac{\partial T}{\partial t} - C_a \frac{\partial W}{\partial t} = -\nabla \cdot q + C_w J \cdot \nabla T \quad (3)$$

$$q = -\lambda \nabla T + C_w J \cdot \nabla T \quad (4)$$

Sendo ρ a densidade da matriz de concreto, C é o calor específico, C_a é o calor latente de vaporização da água livre, C_w é o calor específico da água, q é o fluxo de calor, e por fim λ é a condutividade térmica do concreto.

Juntando as equações de balanço de massa, (1) e energia (3) é criado um sistema de duas EDPs parabólicas com quatro variáveis de estado (W_d , W , T , P). Sendo W_d calculada através de uma função por partes obtida da análise termogravimétrica do concreto. Como W é uma função termodinâmica de P e T , sua variação no tempo pode ser substituída por sua variação sobre P e T . Com isso, duas funções de estado devem ser calculadas pelas EDPs, pressão e temperatura.

$$W_d(T) = \begin{cases} 0,0017 * T - 0,0171 (0 < T < 200) \\ 0,0407 * T - 7,5143 (200 \leq T < 400) \\ 0,0025 * T + 6,5253 (400 \leq T \leq 800) \end{cases} \quad (5)$$

2.3 Método de discretização

O domínio foi um bloco de concreto de $300 \times 300 \times 120$ mm dividido em volumes de controle regulares não sobrepostos. Cada volume forma um espaço 3D retangular, formando um estêncil de pontos na grade, representado pela Figura 1. Cada ponto da malha tem um nó, que pode representar, pressão, temperatura, ou conteúdo de água. Para se obter os valores dos nós, as equações (1) e (3) são integradas no volume de controle em intervalo de tempo definido.

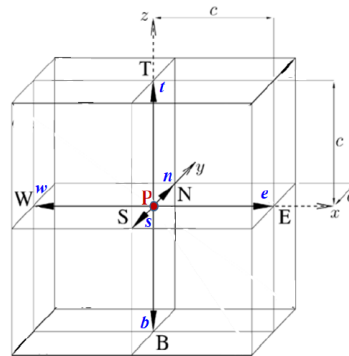


Figura 1 - Estêncil empregado para simulação 3-D. Fonte original: (TRUJILLO, 2020).

Para a discretização espacial, são consideradas as equações a seguir, fruto do desenvolvimento das equações de conservação de massa e energia, respectivamente, são integradas nos volumes de controle. Um perfil linear por partes é assumido com interpolação para P e T entre os nós, para calcular fluxos de massa. Para a discretização do tempo, foi usado um esquema implícito baseado em (PATANKAR, 1980). Com isso, foram desenvolvidos dois sistemas de equações lineares. Resolvendo primeiramente o campo de temperaturas, e com esses valores é calculado, numa seguinte iteração, o campo de pressões.

$$\frac{\partial P}{\partial t} = \nabla \cdot \left(\frac{a}{A1} \nabla P \right) + \frac{1}{A1} \cdot \frac{\partial W_d}{\partial t} - \frac{A3}{A1} \cdot \frac{\partial T}{\partial t} \quad (6)$$

$$\frac{\partial T}{\partial t} = \nabla \cdot \left(\frac{\lambda}{A4} \nabla T \right) + \frac{C_w}{A4} \cdot \left(\frac{a}{g} \cdot \nabla P \right) \cdot \nabla T - \frac{A2}{A4} \cdot \frac{\partial P}{\partial t} \quad (7)$$

2.4 Condições de contorno

Para o fluxo térmico, somente a face superior foi aquecida por meio de radiação, seguindo uma curva de aquecimento próxima à utilizada por (KALIFA, 2000). A condição de radiação foi definida pela lei de Stefan-Boltzmann (BOLTZMANN, 1884), expressa na equação:

$$\vec{q} \cdot n = \epsilon_c \cdot \sigma_{SB} (T^4 - T_{amb}^4) \quad (8)$$

Sendo n o vetor unitário normal ortogonal da superfície; $\vec{q}$ é o fluxo de calor por radiação; T é a temperatura do corpo radiante; T_{amb} é a temperatura ambiente; $\sigma_{SB} = 5,67 \times 10^{-8} \text{ W}/(\text{m}^2 \text{ K}^4)$ é a constante de Stefan-Boltzmann; $\epsilon_c = 0,9$ é a emissividade de calor da superfície do concreto (JONES, 2000).

As faces laterais simulam estar em contato com blocos isolantes de baixa condutividade térmica, então, foram consideradas superfícies com uma condição de contorno do tipo Neumann:

$$\vec{q} \cdot n = 0 \quad (9)$$

Na região inferior do bloco, para a condição de transferência de calor foi utilizada a lei de Newton de resfriamento:

$$\vec{q} \cdot n = B_T (T - T_{amb}) \quad (10)$$

Sendo T a temperatura da superfície do domínio; $B_T = 1,0 \text{ W}/(\text{m}^2 \text{ K})$ é o coeficiente de transferência de calor da superfície. Foi considerada convecção livre na superfície (CHAPMAN, 1987).

Para o fluxo de massa, o bloco de concreto foi definido com contornos permeáveis. Porém, a face inferior está em contato com uma chapa de aço, considerada impermeável, similar ao arranjo de (TRUJILLO, 2020), dessa forma, os fluxos normais a esse contorno desaparecem numa condição de contorno do tipo Neumann:

$$\vec{j} \cdot n = 0 \quad (11)$$

Sendo n o vetor unitário normal ortogonal da superfície de contorno; $\vec{j}$ é o fluxo da massa de água. A face superior do corpo de prova está em contato com a atmosfera e as faces laterais estão em contato com tijolos refratários isolantes de alta porosidade, nesses contornos a pressão P foi definida como a pressão atmosférica P_{amb} . A troca de vapor entre a superfície e o meio ambiente não é instantânea, definindo que o fluxo normal é proporcional à diferença de pressões:

$$\vec{j} \cdot n = B_w (P - P_{amb}) \quad (12)$$

Sendo $B_w = 1 \times 10^{-6}$ o coeficiente de transferência de vapor da superfície dependente da movimentação do ar, a iluminação ou o isolamento parcial.

2.5 Parâmetros iniciais

Em razão das propriedades não lineares, dependentes de temperatura, saturação, pressão e composição do concreto, os dados, relações e formulações das propriedades foram obtidas utilizando dados da literatura, como dispostos na Tabela 1:

Tabela 1: Principais parâmetros iniciais para a simulação

Propriedade	Nomenclatura	Origem	Valor	Unidade
Permeabilidade intrínseca	κ	(GUO, 2016)	$1,0 \cdot 10^{17}$	m^2
Condutividade térmica	λ	(GUO, 2016)	1,64	$W/m \cdot K$
Duração da simulação	t_r	(Definição própria)	40	hr
Tamanho do passo especial	$\Delta (x, y, z)$	(Definição própria)	0,25	mm
Tamanho do passo temporal	Δt	(Definição própria)	0,25	s
Temperatura inicial	T_0	(Definição própria)	20	°C

3. Resultados e discussão

3.1 Campo de Temperaturas

Na simulação foi usada uma curva de aquecimento desde 20°C até 450°C em um período de 6 horas, como mostra a Figura 2. Em todas as simulações foram considerados seis pontos dentro do material com profundidades de 10, 20, 30, 40, 50 e 120 mm. Se observa que quanto mais próximo da fonte quente, maior a temperatura do material, isso se dá pela condução de calor através do concreto. Vale ressaltar que, devido ao isolamento lateral, a simulação não considera a perda de calor do material para o ambiente. Porém, em uma situação real poderia haver uma saída mínima de calor para o ambiente, pelo que as curvas poderiam estar mais distantes entre si, o que mostra a necessidade dessa consideração.

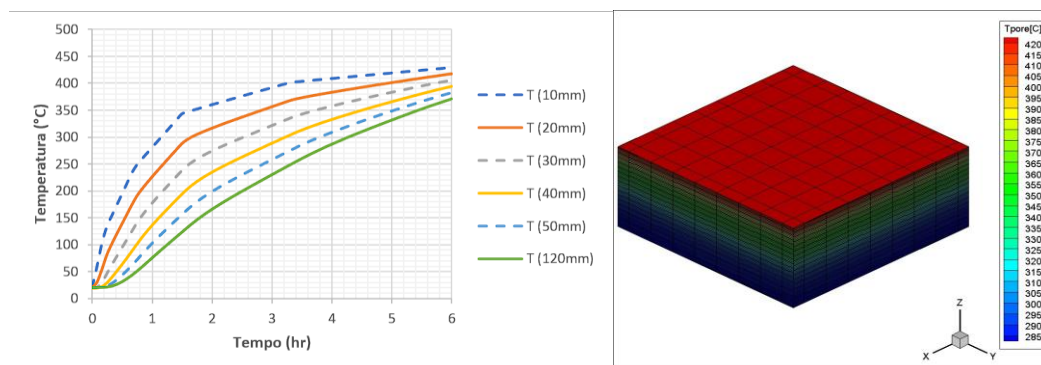


Figura 2: Curvas de temperaturas em diferentes profundidades e imagem 3d na hora 4 de aquecimento.

3.2 Campo de Pressões internas

A figura 3 a seguir mostra os resultados das pressões internas. Nota-se que durante o aquecimento, ao longo do tempo há um aumento gradual com picos de pressão, com máximo de 188 MPa atingido na posição de 120 mm localizada na extremidade mais oposta ao aquecimento, as elevadas pressões atingidas estão relacionadas a baixa permeabilidade dos concretos de alta performance. Observa-se que a pressão em 10 mm atingiu valores de pressão por abaixo das demais, apresentando um comportamento linear ao longo da simulação, acredita-se que por estar mais próxima à face aquecida,

não houve acúmulo de umidade suficiente para gerar pressões internas elevadas. É possível observar que os picos de pressão, com exceção da posição de 10mm, são atingidos em intervalos de tempo que coincidem com níveis de temperatura em torno de 300 °C, e onde há aumento de conteúdo de água, este fenômeno ocorre em condições próximas à temperatura crítica da água.

Na figura 3 à direita, apresentam-se os resultados das pressões simuladas em 3 dimensões, no instante de tempo $t = 0,5$ h em um corte de 50% do domínio. É possível observar no interior maiores níveis de pressão nas regiões centrais do domínio, a partir dos 20 mm de profundidade. Este fenômeno ocorreu ao início da simulação do aquecimento, quando esta região estava numa temperatura de 180 °C.

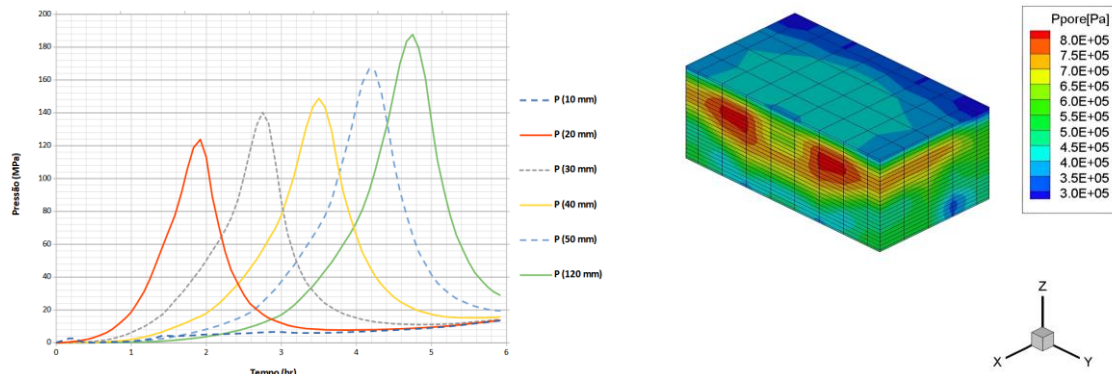


Figura 3: Curvas de pressão em diferentes profundidades e imagem 3d na hora 0,5 do aquecimento.

3.3 Conteúdo de água

A Figura 4 esquerda mostra o comportamento da saída de água livre; ao lado direito mostra o campo 3D da quantidade de água livre após 1h de aquecimento. Se percebe que quanto mais próximo da superfície quente, mais rápido a água livre é liberada, devido a esses pontos chegarem primeiro a temperaturas críticas, fazendo com que a água nessas regiões evapore primeiro. Em 1 hora a maior parte da umidade foi removida, em função do rápido aquecimento da fonte de calor. Em quase todas as profundidades simuladas ocorrem picos, acredita-se que podem ser da água quimicamente ligada que sai do concreto em maiores temperaturas. Às 5h se observa que toda água livre foi liberada do concreto.

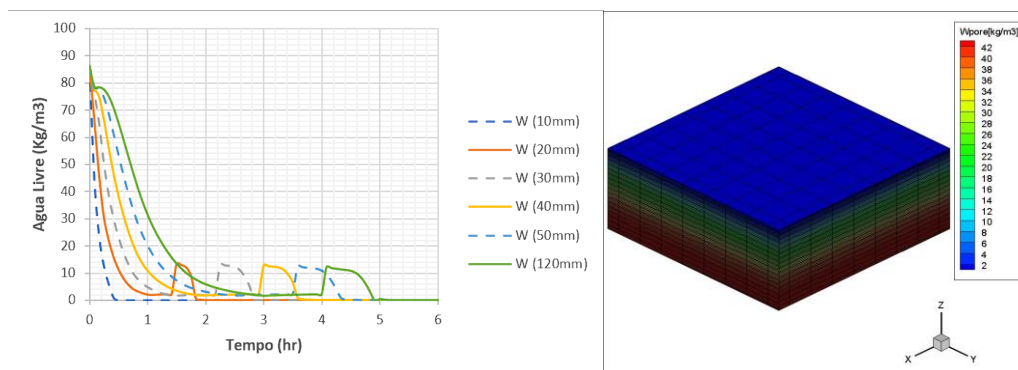


Figura 4: Curvas do conteúdo de água em diferentes profundidades e imagem 3d na hora 1,0 de aquecimento.

4. Conclusão

Neste trabalho foi desenvolvida uma aplicação do método dos volumes finitos, caracterizado como conservativo por descrever os fenômenos de transporte levando-se em consideração a conservação de massa e energia. Foi possível simular o comportamento do vapor dentro de um concreto de alta performance submetido a estresse térmico. As curvas de temperatura apresentaram resultados satisfatórios, não entanto, é preciso um maior refinamento e ajustes nas condições de contorno, a fim de

determinar de forma precisa condições com perdas de energia por saídas de calor que podem ocorrer devido a imperfeições do isolamento lateral nos cenários práticos. As curvas de pressão atingiram patamares consideravelmente elevados, isso é explicado pela baixa permeabilidade deste tipo de concreto, com tudo, a sua validação somente seria possível com experimentos do caso específico, os quais são sofisticados de realizar. O comportamento do conteúdo de água nas diferentes profundidades mostrou uma queda abrupta ao início da simulação, porém, também mostrou a necessidade de maior refinamento, pois foi seguida de um aumento elevado em momentos posteriores.

5. Referências

- BAZANT, Z. P.; JIRÁSEK, M. **Creep and hygrothermal effects in concrete structures**. Evanston: Springer. Solid Mechanics and its Applications, 2018. 960 p.
- BOLTZMANN, L. Derivation of Stefan's law: concerning the dependence of thermal radiation on the temperature of the electromagnetic light theory. **Annals of Physics and Chemistry**, v. 22, p. 291-294, 1884.
- CHAPMAN, A. J. **Fundamentals of heat transfer**. New York: Macmillan, 1987.
- CHAPRA, S.C.; Canale, R.P.; **Numerical Methods for Engineers**. 6. ed. New York: McGraw-Hill, 2010.
- GUO, F.; YUAN, Y.; MANG, H. A. Determination of the relative significance of material parameters for concrete exposed to fire. **International Journal of Heat and Mass Transfer**, v. 100, p. 191–198, 1 set. 2016.
- JONES, H. R. N. **Radiation heat transfer**. Oxford: Oxford University Press, 2000.
- KALIFA, P.; MENNETEAU, F. D.; QUENARD, D. Spalling and pore pressure in HPC at high temperatures. **Cement and Concrete Research**, v. 30, n. 12, p. 1915-1927, Agosto 2000.
- MILLARD, A.; PIMIENTA, P. **Modelling of concrete behaviour at high temperature**. State-of-the-Art Report of the RILEM Technical Committee 227-HPB. Cham, Switzerland, p. 100. 2019. (30).
- MINDEGUIA, J. C.; PIMIENTA, P.; NOUMOWÉ, A.; KANEMA, M. Temperature, pore pressure and mass variation of concrete subjected to high temperature: Experimental and numerical discussion on spalling risk. **Cement and Concrete Research**, v. 40, n. 3, p. 477-487, 2010.
- PALMER, G.; COBOS, J.; MILLARD, J.; HOWES, T. The accelerated drying of refractory concrete: Part I: A review of current understanding. **Refractories WORLDFORUM**, v. 6, n. 2, p. 75 – 83, April 2014.
- PATANKAR, S. V. Numerical heat transfer and fluid flow. **Hemisphere publishing corporation**, 1980.
- TRUJILLO, J. J.; **Contribuições para a determinação de curvas de secagem de refratários monolíticos com auxílio experimental e de simulação numérica**. 2020, 377 f. Tese (Doutorado em Ciências) –Escola de Engenharia de Lorena Universidade de São Paulo, Lorena, 2020.
- TRUJILLO, J.J. CASTRO, A. INNOCENTINI, M. VERNILLI, F. Combined transient 3D simulation and experimental methods to assess a slow heat-up curve used to dry a refractory concrete. **International Journal of Thermal Sciences**, v. 185, 2023.
- ULM, F. J.; COUSSY, O.; BAZANT, Z. P. The “channel” fire. I: chemoplastic softening in rapidly heated concrete. **Journal of Engineering Mechanics**, v. 125, n. 3, mar. 1999

A STATISTICAL ANALYSIS ON THE YIELD STRENGTH OF THE INCONEL 690 TT ALLOY AT ROOM TEMPERATURE

Alexandre Francisco
Universidade Federal Fluminense
afrancisco@id.uff.br

Matheus Felipe Silva Oliveira
Universidade Federal Fluminense
matheusfsoliveira@outlook.com

Abstract. *This article presents a statistical analysis on the yield strength of the Inconel 690 TT alloy used in steam generator tubes of pressurized-water reactor (PWR) nuclear power plants. The goal is to determine the probability distribution that best represents the yield strength data obtained from tensile tests. The analysis includes measures of statistical location and dispersion, as well as graphical and goodness of fit tests. The results show that the data exhibits a symmetric distribution, resembling a Normal distribution. The confidence interval for yield strength sample mean is calculated with a confidence level of 95%. The results demonstrate that the sample mean is within the limits specified by ASME-SB-163 and SANDVIK. A hypothesis test confirms that the sample mean can be considered the population mean. This statistical analysis contributes to understanding the mechanical properties of the Inconel 690 TT alloy and ensures the quality control of the steam generator tubes in nuclear power plants.*

Keywords: *Inconel 690 TT alloy, yield strength, statistical analysis, steam generator tubes, nuclear power plants.*

Introduction

Steam generators are heat exchangers that convert the feedwater into steam using the heat produced by the nuclear reactor. They are used in pressurized-water reactor (PWR) nuclear power plants, the most common type of nuclear power plant in the world. In these plants, there are two different water circuits, the primary and the secondary. The primary circuit contains the pressurized water that circulates through the reactor core, where it is heated due to the nuclear fission of uranium atoms. The heated radioactive water is taken to the steam generator, where it transfers its heat to the secondary circuit, without having direct contact with it. The secondary circuit contains pure water that is vaporized due to the heat received from the primary circuit. The steam is directed to the main turbine, where it makes the electric generator work, producing electricity. The extra steam is condensed in the condenser, where it loses its heat to the refrigerant water. The condensed water is pumped back to the steam generator, to restart the cycle (NRC, 2018).

The steam generator tubes are components that allow the heat exchange between the two water circuits, without mixing them. They are made from special metal alloys that should resist high temperatures, high pressures and corrosion involved with the process. The steam generator tubes are essential for safety and efficiency of PWR nuclear power plants. They prevent the radioactive contamination of the secondary circuit and the external environment, besides ensuring the quality of the produced steam. The steam generator tubes should be inspected and monitored periodically, to detect possible failures or faults that may interfere with their performance (NRC, 2018).

According to ASME SB-163, ASME III NB-2000 and ASME Code Case N-20-3, the Inconel 690 TT should be used in the steam generator tubes. The Inconel 690 TT is an iron-nickel-chromium alloy with high resistance to corrosion and creep, which is used in the Angra 1 nuclear power plant steam generator tubes. The yield strength of this alloy is a very important mechanical property to guarantee the performance under operation conditions. In this article, a statistical analysis of the mechanical data on yield strength of the Inconel 690 TT alloy was conducted, obtained from tensile tests.

To do this, the measurements of statistical location and dispersion of the mechanical data were calculated, such as the mean, median, standard deviation, variance, mode and the coefficient of

variation (Lewis, 1996). After that, the graphical method was used to compare the probability density histograms of the mechanical data with the curves of the Normal, Lognormal, Weibull and Exponential probability functions. This method allows a visual analysis of the probability function fit to the mechanical data, but doesn't offer any quantitative criteria to choose the best distribution (Lewis, 1996). Due to that, the Chi-squared, Anderson-Darling and Kolmogorov-Smirnov goodness of fit tests were applied, that evaluate the null hypothesis that the mechanical data follows a certain distribution based on the measures of the distance between the empirical and theoretical cumulative distribution functions (Montgomery, 2002).

After the definition of the probability distribution that better represents the mechanical data, the confidence interval for the yield strength mean was calculated, with confidence level of 95%. The confidence interval represents the interval in which the population mean is expected to be in, based on the sample mean and on the variability of the mechanical data. Besides that, a box plot (graphic that shows the median, quartiles and the extreme values of the data) was made to identify possible outliers, values that deviate from the central tendency of the mechanical data (Montgomery, 2002).

The statistical analysis realized in this article contributes to the knowledge of mechanical properties of the Inconel 690 TT alloy and for the quality control of the steam generator tubes of the Angra 1 nuclear power plant.

Methodology

The goal of this analysis was to identify the probability distribution that better represents the mechanical data, as well as to calculate the confidence interval for the mean and to detect possible outliers. Once the best probability function is found, it will be used for the reliability analysis of the steam generator tubes.

This paper contains statistical data from the steam generator tubes yield strength. These data were provided by SANDVIK to identify the uncertainty in such data. The statistical information includes the mean (μ), the standard deviation (σ), the coefficient of variation (COV) and the probability distribution that better represents the mechanical properties of the Inconel 690 TT alloy. Also, other important statistical information will be provided, such as the median, the bias and others.

According to ASME SB-163, the Inconel 690 TT alloy should have a minimum value of yield strength of 276 MPa, at room temperature (24 °C). The mechanical data provided contain 102 measurements of yield strength at room temperature obtained from a tensile test (ASME, 2000).

The first step to start the statistical analysis of these mechanical data is to build a frequency histogram that enables analyzing the distribution of the yield strength values obtained from the tensile test. To build the frequency histogram, a frequency table must be built, and for that, the class interval, or bins, (ΔS) may be defined by (Lewis, 1996):

$$\Delta S = r[1 + 3,3 \log_{10}(N)^{-1}] \quad (1)$$

where:

r is the range (difference between the maximum and minimum values of the yield strength data at room temperature and N is the Number of the yield strength data points at room temperature.

The next step after building the frequency histogram is to use the graphical method to define the probability distribution that better fits the mechanical data. This method consists of comparing the histogram with the curves of the theoretical probability density functions of each distribution. In this article, the Normal, Lognormal, Weibull and Exponential probability distributions will be used. The distribution that presents the best visual fit with the frequency histogram can be chosen to represent the mechanical data (Montgomery, 2002).

After performing the graphical analysis, goodness of fit tests were also performed to confirm the graphical method results and to obtain more accurate results. The goodness of fit test consists of comparing the mechanical data obtained from the tensile test with the expected data according to the chosen probability function. As previously said, the Normal, Lognormal, Weibull and Exponential probability distributions will be used in this test. The goodness of fit test provides a numerical value of how well the data fits the probability distribution, as well as the confidence level that indicates the probability of rejecting the null hypothesis that the mechanical data fits a certain distribution. In this

article, the Chi-squared, the Anderson-Darling and the Kolmogorov-Smirnov goodness of fit tests will be used to verify which distribution better fits the mechanical data (D'Agostino, 1986).

Once the graphical method and the goodness of fit tests are done, and the probability distribution that better fits the mechanical data is chosen, a confidence interval test for the mean is needed to ensure that all the calculations that were made previously are valid, and that all calculations that will be made in the future, involving the reliability analysis, will be using the right values of the mechanical properties of the Inconel 690 TT alloy. A level of confidence of 95% will be used for the confidence interval test for the mean (Montgomery, 2002).

Statistical analysis of the mechanical data

In this section, the results of the statistical characterization of the yield strength of the Inconel 690 TT alloy will be discussed. To characterize this data distribution, measurements of statistical location and dispersion must be calculated based on the data obtained from the tensile test. A box plot will also be plotted to show the maximum and minimum values of the mechanical data, as well as possible outliers. The measurements of statistical location calculated include the sample size (N), the sample mean ($\bar{M}$), the median (Med) and the mode (Mo). The measurements of statistical dispersion calculated include the sample standard deviation (s), the sample variance (s^2) and the coefficient of variation (COV) (Sekulski, 2019). Uncertainty measurements as the difference bias and ratio bias were calculated as well. The results of the calculations are shown in the Table 1.

Table 1. Statistical data of the yield strength at room temperature.

Sample size:	102
Sample mean:	312,58 MPa
Median:	313,00 MPa
Mode:	310,00 MPa
Standard deviation:	5,52 MPa
Coefficient of variation:	1,77 %
Variance:	30,44 MPa ²
(average) difference bias:	11,58 MPa
(average) ratio bias:	1,04

The results infer that the mechanical data has a symmetric distribution, with sample mean, median and mode with close values, which is characteristic of a Normal distribution (Lewis, 1996). The value of the sample mean is also above the minimum specified by ASME-SB-163. Both standard deviation and coefficient of variation got low values, which indicate the low variability of the mechanical data. The box plot (Figure 1) doesn't show any outlier, that means that there are no values that differ a lot from the sample mean.

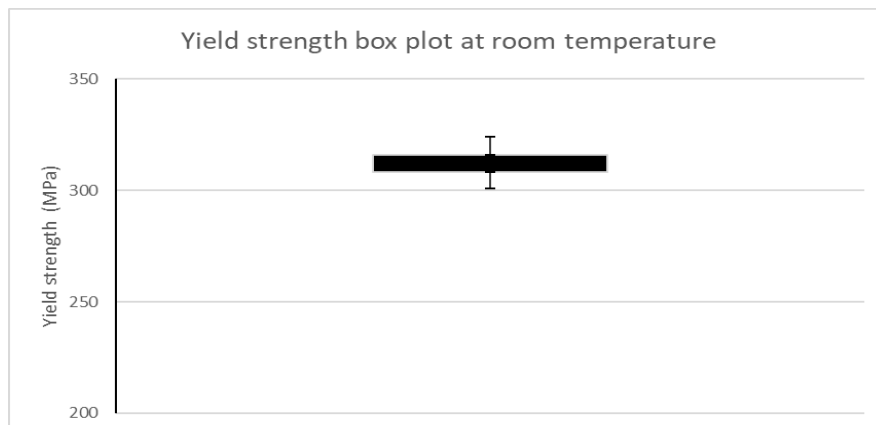


Figure 1. Yield strength box plot at room temperature.

These results suggest that all the mechanical data obtained from the tensile test of the Inconel 690 TT alloy were consistent and representative, and can be used for further analysis (Lewis, 1996).

Once the data characterization of the Inconel 690 TT mechanical data was done, a frequency table was built (Table 2). The frequency table must cover all of the data range, from the minimum to the maximum value (Lewis, 1996).

Table 2. Frequency table for yield strength at room temperature.

Bins	Frequency	Relative freq.	Relative freq. dens.
[301, 304[	6	0,0588	0,0147
[305, 308[	20	0,1961	0,0490
[309, 312[	24	0,2353	0,0588
[313, 316[	27	0,2647	0,0662
[317, 320[	18	0,1765	0,0441
[321, 324[	7	0,0686	0,0172

The frequency table was built to assist in the creation of a frequency histogram (Figure 2) to compare the mechanical data with the Normal, Lognormal, Weibull and Exponential distributions. The graphical method plots the theoretical distribution curves over the frequency histogram of the actual data and analyzes which theoretical distribution has the best fit to the shape of the mechanical data frequency histogram. The results presented in Figure 2 show that the Normal and Lognormal distributions have a good fit to the sample data, followed by the Weibull distribution. The exponential distribution has no fit for the data (Montgomery, 2002).

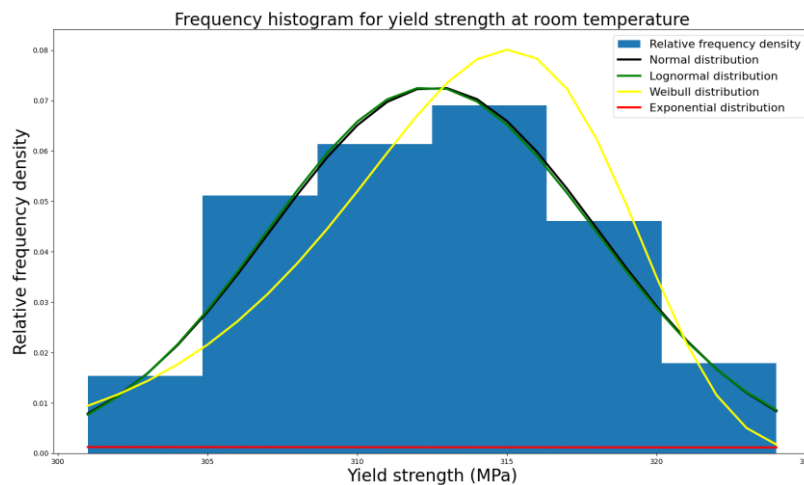


Figure 2. Frequency histogram at room temperature.

Due to the proximity of Normal and Lognormal distribution curves using the graphical method, it wasn't possible to determine which of these distributions have a better fit to the mechanical data of the Inconel 690 TT alloy. The results were inconclusive because the graphical method is based in a visual and subjective analysis of the fitting. A goodness of fit test was necessary to get more precise and conclusive results. The goodness of fit tests are statistical methods that allow to compare the probability distributions and choose the one that better fit the data according to the null hypothesis (Montgomery, 2002).

The Chi-squared, Kolmogorov-Smirnov and Anderson-Darling goodness of fit tests were performed to test which of the distributions has a better fit to the sample data of the mechanical properties of the Inconel 690 TT alloy. These tests are based on the difference between the empirical distribution of the sample data (test statistic) and the theoretical distribution function tested (critical value). The tests return a p-value, that indicates the probability of the data to fit in a certain

distribution assumed by the null hypothesis. The higher the value of p-value, the higher is the evidence favoring the null hypothesis. The results of the tests show that the Normal distribution presents higher values in the Chi-squared and Anderson-Darling tests. The difference between the results of the Normal and Lognormal distribution in the Kolmogorov-Smirnov tests is negligible as shown in Table 3 (D'Agostino, 1986).

Once the distribution that better fits the data is chosen, a confidence interval test for the mean was performed. This test was done to determine the true value of the population mean with a certain confidence level, based on the sample mean and standard deviation. The confidence interval provides a range of values for the yield strength, that contains the true value of the mean. The test is very useful to assess how precise the goodness of fit tests were and to analyze the variability of the data, to see how far from the sample mean the mean can be (Montgomery, 2002).

Table 3. Goodness of fit test results for yield strength at room temperature.

Chi-squared			
	χ^2_0	χ^2_{crit}	p-value (%)
Exponential	55,56	9,49	0,00
Normal	0,02	7,81	98,40
Lognormal	0,16	7,81	87,28
Weibull	0,184	7,81	85,40
Kolmogorov-Smirnov			
	D_n	D_{crit}	p-value (%)
Exponential	0,608	0,135	54,30
Normal	0,072	0,135	94,26
Lognormal	0,061	0,135	95,14
Weibull	0,129	0,135	89,74
Anderson-Darling			
	A^*	A_{crit}	p-value (%)
Exponential	45,517	0,752	0,00
Normal	0,415	0,752	67,81
Lognormal	1,85	0,752	6,43
Weibull	23,06	0,752	0,00

The maximum and minimum yield strength values were already established by ASME-SB-163 and also provided by SANDVIK in the results of the tensile test. For these values, some statistics were so low, to the point that they couldn't even be found in a standard Normal distribution table. Looking to find out what was happening, three confidence interval for the mean were performed: 90%, 95% and 99%. The results are shown in Table 4.

Table 4. Confidence interval for mean at room temperature.

Confidence level	Confidence interval (MPa)
90%	$311,68 < \mu < 313,48$
95%	$311,51 < \mu < 313,65$
99%	$311,17 < \mu < 313,99$

The results of the confidence interval test for the mean got a very low range in every confidence level tested. This is an evidence that the yield strength mean is according to the limits established by ASME-SB-163 and SANDVIK, which got higher ranges for the yield strength to be worked.

A lot of problems involving engineering may require accepting or rejecting a hypothesis about some statistical parameter. In this article, a hypothesis test for the mean was performed. Through this test, it is possible to check how strong is the evidence that the yield strength mean is according to the limits established by ASME-SB-163 and SANDVIK.

It was stated that the yield strength data follows a Normal distribution, with a test statistic Z_0 . The resulting probability is compared to a pre-defined significance level (usually $\alpha = 0.05$) to determine whether the null hypothesis should be rejected or not. The observed value of Z_0 is compared to tabulated values Z_{crit} that take into account the significance level. This allows engineers to make informed decisions about whether the material meets the requirements set by standards such as ASME-SB-163 and SANDVIK.

Two hypotheses, a null (H_0) and an alternative (H_1), are established regarding the mean of the yield strength at room temperature. The goal is to determine if these values are above a minimum value set by technical standards. The hypotheses are formulated by:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu > \mu_0 \end{cases} \quad (2)$$

where:

μ_0 is the supposed true value of the yield strength mean.

The results shown in Table 5 show that Z_0 values are greater than $Z_{crit} = 1.645$, so the null hypothesis must be rejected. Therefore, there is strong statistical evidence to assert that the mean of the yield strength of the Inconel 690 TT alloy at room temperature will be greater than the minimum values established by ASME-SB-163 and SANDVIK.

Table 5. Result of the hypothesis test for the mean at room temperature.

Parameter	Z_0	Z_{crit}
Yield strength mean	66,92	1,645

Conclusion

In conclusion, the statistical analysis conducted in this study provides valuable insights into the mechanical properties of the Inconel 690 TT alloy used in the steam generator tubes of the Angra 1 nuclear power plant. The analysis determined that the mechanical data exhibits a symmetric distribution, with the sample mean, median, and mode close in value, suggesting a Normal distribution. The standard deviation and coefficient of variation indicated low variability in the mechanical data.

Through graphical and goodness of fit tests, it was determined that the Normal and Lognormal distributions best fit the mechanical data. However, the confidence interval test for the mean revealed that the sample mean falls within the limits specified by ASME-SB-163 and SANDVIK, supporting the use of the Normal distribution for further analysis.

Furthermore, the hypothesis test for the mean confirmed that the sample mean can be considered as the population mean, validating all calculations and tests performed in this study. This statistical analysis contributes to the understanding of the Inconel 690 TT alloy's mechanical properties and quality control of the steam generator tubes in the nuclear power plant.

By ensuring the reliability and safety of the steam generator tubes, the findings of this analysis can aid in maintaining the efficiency and effectiveness of pressurized-water reactor (PWR) nuclear power plants, promoting the continued use of nuclear energy as a clean and sustainable energy source. The adoption of Inconel 690 TT alloy in steam generator tubes demonstrates its suitability for withstanding the harsh conditions of the nuclear power plant environment, ensuring the prevention of radioactive contamination and enhancing the safety of the plant and the surrounding environment.

Acknowledgements

This study was financed in part by Eletrobrás Termonuclear S/A through the scholarship for Matheus Felipe Silva Oliveira.

References

U.S. Nuclear Regulatory Commission (NRC). (2018). Pressurized Water Reactor (PWR) Systems. Disponível em: <https://www.nrc.gov/reading-rm/basic-ref/students/for-educators/04.pdf>

MONTGOMERY, R. *Applied Statistics and Probability for Engineers*. [S.l.]: John Wiley Sons, Inc., 2002.

Lewis, E. E. Introduction to Reliability Engineering. Wiley. 2ª edição, 1996.

Ralph B. D'Agostino. **GOODNESS-OF-FIT-TECHNIQUES**. New York: Marcel Dekker INC, 1986. 589 p. v. 68.

ASME (American Society of Mechanical Engineers). Specification for seamless nickel and nickel alloy condenser and heat-exchanger tubes. ASME SB-163. New York: ASME, 2022.

ASME (American Society of Mechanical Engineers). Alternative rules for qualification of welders and welding procedures for welding nickel alloy stainless steels. ASME Code Case N-20-3. New York: ASME, 1998.

ASME (American Society of Mechanical Engineers). Rules for construction of nuclear facility components. ASME III NB-2000. New York: ASME, 2000.

SEKULSKI, Zbigniew. Statistical Properties of the Yield Strength of Normal Strength Hull Structural Steel Plates. Poland, 2019.

Previsão de Enchentes no Rio Doce utilizando modelos SARIMA

Jean Sousa Mendes

*Programa de Modelagem Computacional em Ciência e Tecnologia,
Universidade Federal Fluminense
jeansousa@id.uff.br*

Eliane da Silva Christo

*Programa de Modelagem Computacional em Ciência e Tecnologia
Universidade Federal Fluminense
elianechristo@id.uff.br*

Kelly Alonso Costa

*Departamento de Engenharia de Produção
Universidade Federal Fluminense
kellyalonso@id.uff.br*

Resumo: As inundações fluviais representam uma ameaça grave em todo o mundo, com impactos financeiros e significativos para a sociedade. Este estudo concentrou-se na bacia do Rio Doce, especialmente na cidade de Governador Valadares, que enfrenta enchentes recorrentes durante o verão. O objetivo principal foi desenvolver um modelo estatístico de previsão de enchentes baseado no método SARIMA (*Seasonal Autoregressive Integrated Moving Average*). Uma característica fundamental deste estudo é a previsão com 6 horas de antecedência, para possibilitar uma resposta ágil às elevações do nível do rio e garantir tempo hábil para a implementação de medidas preventivas pelos órgãos competentes. O modelo obteve um erro médio absoluto percentual (MAPE) de 0,64% e um erro quadrático médio (RMSE) de 0,0224.

Palavras-chave: Previsão de Enchentes, SARIMA, Séries Temporais.

Introdução

As enchentes de rios são uma das catástrofes naturais mais custosas para a sociedade, tanto em termos financeiros quanto em perda de vidas (World Economic Forum, 2022). No século XX, as enchentes de rios causaram uma perda média global anual direta de US\$ 104 bilhões (UNDRR, 2015) e resultaram em cerca de sete milhões de vidas perdidas (Doocy et al., 2013). É compreensível, portanto, que as sociedades sempre tenham se esforçado para reduzir as catástrofes de inundações por meio de instrumentos de gerenciamento de enchentes.

O Rio Doce percorre extensas áreas dos estados de Minas Gerais e Espírito Santo, sendo uma importante bacia e fonte crucial de pesca e recursos para a cidade de Governador Valadares. No entanto, devido ao alto nível de assoreamento de suas margens, essa cidade enfrenta enchentes regulares com impacto significativo para os moradores das áreas na planície de inundação (Amorim et al., 2018).

Em 11 de janeiro de 2022, Governador Valadares registrou a terceira maior enchente histórica, onde o rio atingiu 4,35 metros. O maior nível do rio na cidade ocorreu em 1979, quando o rio Doce atingiu 5,18 metros, enquanto o segundo em 1997 com 4,77 metros. O alerta do risco de enchentes começa a partir de 3,2 metros, e a partir de 3,6 metros se caracteriza como inundação (SAAE, 2022).

Devido à gravidade dos danos causados pelas enchentes, a criação de ferramentas de previsão para o nível do rio tornou-se importante instrumento para mitigar impactos do fenômeno. Com base nessa necessidade, este artigo tem como objetivo desenvolver um modelo estatístico de previsão de enchentes do rio Doce na cidade de Governador Valadares, em Minas Gerais. Espera-se que essa abordagem proporcione informações mais precisas e oportunas para a gestão de situações críticas e contribua para a redução dos riscos associados às enchentes.

Revisão Bibliográfica

Nesta seção, discutem-se algumas pesquisas relacionadas à aplicação de modelos para previsões do fluxo de rios e prevenções de enchentes.

De acordo com Sorooshian et al. (2008), um modelo é uma representação simplificada de um sistema do mundo real. O melhor modelo é aquele que fornece resultados próximos à realidade com o uso do menor número de parâmetros e complexidade do modelo.

Na região Sul Fluminense, no interior do estado do Rio de Janeiro, Brasil, Meza (2021) estudou a previsão de inundações na cidade de Volta Redonda, que está localizada às margens do rio Paraíba do Sul e sofre com frequentes inundações. O ponto principal da pesquisa foi a utilização de um modelo de regressão linear generalizada, em inglês: "*generalized linear model*" (GLM), do tipo múltipla. As variáveis de entrada foram dados históricos de vazões de três pontos de medições diferentes do nível do rio. O autor também combinou a técnica de regressão múltipla com a otimização por enxame de partículas, "*particle swarm optimization*" (PSO), alcançando um menor erro estatístico de previsão. O erro percentual absoluto médio, "*mean absolute percentage error*" (MAPE), foi de 0,1758% para a combinação das técnicas. O MAPE é um tipo de cálculo de erro que expressa a porcentagem média dos desvios (em valor absoluto) entre os valores reais e as previsões. Além do disso, foi desenvolvido um aplicativo confiável para alertar sobre o risco de inundações do rio Paraíba do Sul em Volta Redonda.

Tadesse et al. (2017) estudaram o rio Waterval, na África do Sul, que é essencial para o planejamento e gerenciamento do sistema da operação do reservatório de Vaal Dam. Segundo o autor, não havia informações disponíveis na literatura sobre a previsão de fluxos de rios nessa região. Ele utilizou o modelo "*Seasonal Autoregressive Integrated Moving Average*" (SARIMA) de parâmetros (3,0,2)×(3,1,3)[12] para previsões do fluxo médio mensal. O modelo obteve um erro estatístico "*root-mean-square error*" (RMSE) de 1,342. O RMSE é um tipo de erro que calcula a média das raízes quadráticas dos desvios entre os valores reais e as previsões. O modelo selecionado apresentou bom desempenho, com previsões próximas aos valores observados, sendo útil para o planejamento e gerenciamento do sistema do rio Waterval e operação do reservatório Vaal Dam.

Coleta e tratamento de dados

O estudo utilizou dados de medições de vazão horária do Rio Doce na estação de Governador Valadares (código 56850000, registrada como "UHE GOVERNADOR VALADARES"). Esses dados são gerenciados pela Agência Nacional das Águas (ANA) e estão acessíveis por meio da plataforma "HidroWeb" (ANA, 2023). A análise abrangeu um período de quatro anos, de 2018 a 2022, excluindo os anos anteriores à 2017 devido à falta significativa de registros naquele ano. Os dados foram transformados em médias a cada seis horas para aprimorar a previsão de inundações, pensando em uma previsão com tempo hábil de resposta à mitigação de danos.

Uma análise exploratória envolveu a remoção dos horários sem valores (cerca de 1% dos registros), também se avalia a normalidade dos dados por meio de um histograma de frequência, além da aplicação do teste de Shapiro-Wilk, o qual avalia a hipótese nula de que os dados seguem uma distribuição normal (Royston 1992). Por fim, verifica-se a estacionariedade da série por meio do teste de Dickey-Fuller (Dickey et al., 1979). Essas etapas têm o objetivo de garantir a qualidade dos dados para as análises subsequentes.

SARIMA

O modelo estatístico "*seasonal autoregressive integrated moving average*" (SARIMA) foi escolhido para analisar os dados de fluxo de água do rio devido à sua capacidade de incorporar tendências e variações aleatórias nos dados, lidar com padrões sazonais ao longo do tempo e oferecer

simplicidade em comparação com métodos mais complexos. Isso o torna uma escolha eficaz para a análise de séries temporais de fluxo de água de rios. O SARIMA é um modelo sazonal que incorpora tanto fatores não sazonais quanto sazonais em um modelo multiplicativo. A Equação (1), segundo Morettin (2006), apresenta a notação abreviada para um modelo SARIMA(p, d, q)(P, D, Q)s:

$$\Phi_p(B^s)\varphi(B)\nabla_s^D\nabla_t^d x_t = \theta_q(B^s)\theta(B)w_t \quad (1)$$

Onde $\{w_t\}$ é a série temporal não estacionária, $\{x_t\}$ é o processo de ruído branco gaussiano, "s" é o período da série temporal. Os componentes autoregressivos e de média móvel ordinários são representados por polinômios $\varphi(B)$ e $\theta(B)$ de ordens "p" e "q". Os componentes autoregressivos e de média móvel sazonais são representados por $\Phi(B^s)$ e $\Theta(B^s)$ de ordens P e Q. Os termos ∇_d e ∇_s^D são componentes de diferença ordinária e sazonal de ordens d e D, B é o operador de deslocamento ou atraso da observação no tempo.

Todo o desenvolvimento do modelo será realizado através do “software” R (R Core Team, 2021).

Treino e teste do modelo

Após a definição e tratamento da base de dados, é necessário dividi-la em duas partes: treinamento e validação. O período escolhido para a fase de treinamento compreende de 01/01/2018 até 08/01/2022. Essa escolha se deve ao fato de que, em 11/01/2022, a cidade de Governador Valadares enfrentou uma das maiores enchentes registradas no município (SAAE, 2022).

Portanto, é relevante testar o modelo para capturar os eventos ocorridos próximos dessa data. Para a etapa de teste, o modelo gerará previsões futuras que serão comparadas com os dados reais. Após cada previsão, o modelo será atualizado com o dado real correspondente e, em seguida, uma nova previsão será gerada e comparada com o próximo dado real. Esse processo será repetido durante o período de 10 dias na base de validação, com o objetivo de analisar o comportamento do modelo nos dias mais críticos da inundação ocorrida em 2022.

A Tabela 1 apresenta um resumo da estrutura da base de dados e suas divisões entre conjunto de treinamento e conjunto de teste.

Tabela 1: Resumo da estrutura da base de dados do Rio Doce na estação 56850000

Dados	Frequência	Período	Adaptação	Base de treino	Base de teste
Nível em cm	Horária	2018-2022	Médias de 6h	<< 07/01/22	>> 08/01/2022

Definição de parâmetros

É crucial analisar a dependência temporal da série de dados usando a função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP) para determinar os parâmetros do modelo SARIMA. A FAC ajuda a identificar a dependência temporal geral, enquanto a FACP isola a correlação direta entre valores presentes e passados, auxiliando na escolha do termo autorregressivo, “*autoregressive*” (AR) e do termo de média móvel, “*moving average*” (MA) do modelo (Morettin 2006).

Além da análise gráfica das funções de autocorrelação, também utiliza-se a função “auto.arima” do pacote “forecast” no R (Hyndman e Khandakar 2008). Essa função automatiza a seleção de parâmetros com base em critérios estatísticos como “*Akaike information criterion*” (AIC) e “*Bayesian information criterion*” (BIC).

O período sazonal escolhido para o SARIMA foi de 4 passos devido à natureza horária dos dados organizados em intervalos de 6 horas. Dessa forma, os padrões sazonais de curto prazo diários são capturados para o fluxo do rio.

Validação do modelo

Para verificar a presença de autocorrelação serial nos resíduos para diagnosticar a adequação do modelo ajustado, aplica-se o teste Ljung-Box. Ele testa a hipótese nula de que não há autocorrelação nos resíduos até um determinado número de defasagens (Ljung et al. 1978).

Resultados

A Figura 1 traz a evolução do nível do rio Doce ao longo dos últimos anos, nota-se picos elevados durante os períodos de verão de cada ano, com destaque para o ocorrido no início de 2022.

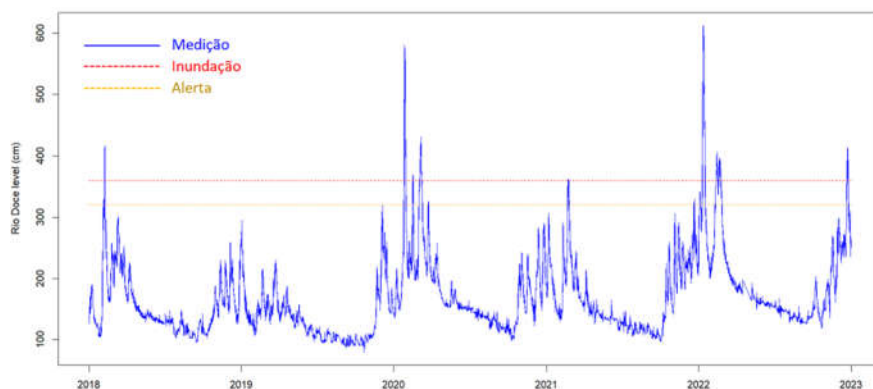


Figura 1: Nível do Rio Doce ao longo dos últimos anos

A Figura 2 compara a série original com a série logaritmizada através da transformação logarítmica na base 10. Na primeira imagem é possível notar uma distribuição assimétrica e positiva (cauda direita) que foi reduzida após a transformação.

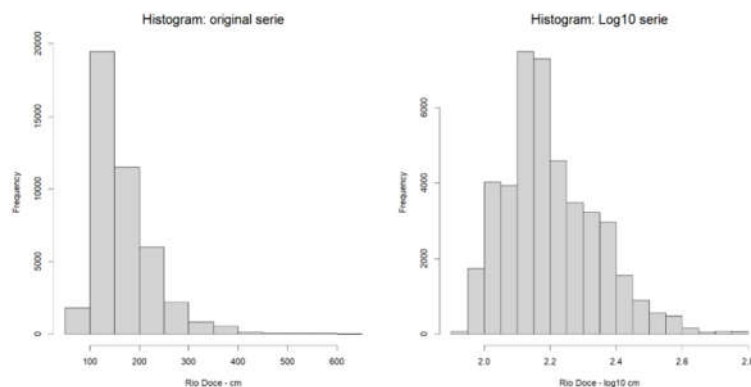


Figura 2: Comparação do nível do Rio Doce, dados reais versus simulação

O teste estatístico de Shapiro-Wilk confirma a distribuição normal dos dados após a logaritmização, resultando em um valor de p de 0,0631 para o teste, logo não é possível rejeitar a hipótese nula de que os dados possuem uma distribuição normal a um nível de significância de 95%.

A Figura 3 traz as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) geradas na análise.

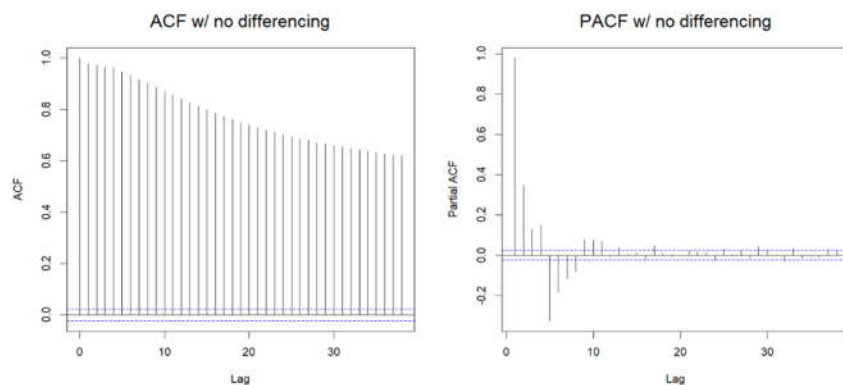


Figura 3: FAC e FACP da série temporal

O gráfico da FAC aponta um decaimento lento, o que indica que a série não é estacionária, propõe-se então uma diferenciação na própria série a fim de estabilizar a média.

A Figura 4 apresenta a FAC e FACP após a aplicação de uma diferenciação na série.

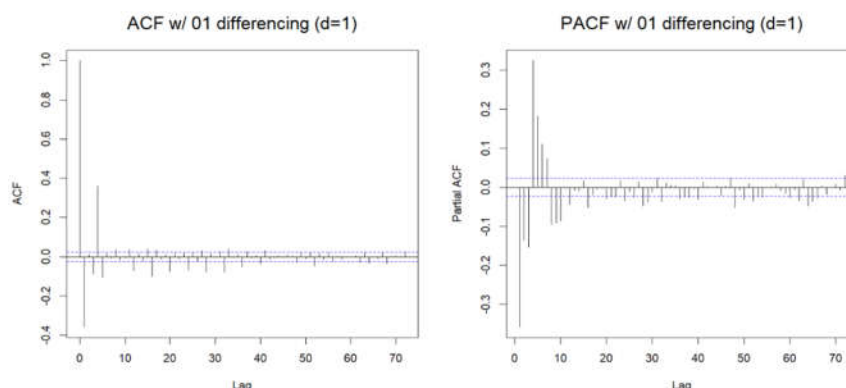


Figura 4: FAC e FACP da série temporal após uma defasagem

Após a diferenciação, é possível observar que uma diferenciação de ordem $d=1$ é suficiente. A estacionariedade é confirmada pelo teste de Dickey-Fuller, o qual resultou em um valor de p menor que 0,01, rejeitando a hipótese nula de não estacionariedade para os dados diferenciados.

A FAC também sugere uma ordem não sazonal $q=2$ e uma ordem sazonal $Q=1$, isso por causa da presença de picos significativos apenas nas posições sazonais múltiplas de 4. Quanto à FACP, os resultados sugerem uma ordem não sazonal $p=3$ e uma ordem sazonal $P=2$, considerando picos sazonais nas posições 4 e 8. A rápida taxa de decaimento nas funções FAC e FACP sugere um valor de $D=0$. Ao gerar o modelo SARIMA com esses parâmetros, obtém-se um erro MAPE de 0,6437%. No entanto, análises estatísticas mais precisas são necessárias para confirmar a escolha dos parâmetros em relação à análise gráfica.

Com base na função *auto.arima* do pacote "forecast" do R, o modelo SARIMA (2, 1, 3)(2, 0, 2)₄ foi sugerido para a previsão dos fluxos do rio Doce. Este modelo apresentou um erro MAPE de 0,6428%, valor menor do que o modelo estimado pela análise gráfica, sendo, portanto, o critério para a escolha do modelo neste estudo.

O valor do BIC para o modelo foi de -27458,32, enquanto o AIC foi de -27524,96. Estes testes AIC são utilizados na avaliação de modelos. O teste estatístico de Ljung-Box realizado no modelo resultou em um valor de p de 0,9913, indicando que não há autocorrelação significativa nos resíduos, pois, não é possível rejeitar a hipótese nula de independência.

Os termos φ_1 e φ_2 representam os coeficientes autorregressivos dos componentes não sazonais, enquanto θ_1 , θ_2 , θ_3 representam os de média móvel dos componentes não sazonais. Já os termos Φ_1 e Φ_2 representam os coeficientes autorregressivos sazonais enquanto Θ_1 e Θ_2 são os de média móvel dos componentes sazonais.

A Tabela 2 permite a comparação do erro deste modelo em relação aos outros modelos discutidos na seção de revisão bibliográfica.

Tabela 2: Resumo dos resultados deste e de alguns outros trabalhos

Rio	Autor	Método	Previsão	MAPE	RMSE
Paraíba do Sul (BR)	Meza (2021)	GLM & PSO	Horária	0,18%	
Waterval (ZA)	Tadesse et al. (2017)	SARIMA	Mensal		1,3420
Rio Doce (BR)	Autores	SARIMA	6 horas	0,64%	0,0224

O estudo revelou um alto potencial de previsão ao adotar dados organizados em intervalos de 6 horas, uma escolha adequada para lidar com o tipo de elevação do nível do Rio Doce. Embora tenha apresentado um desempenho ligeiramente inferior em comparação com o modelo de previsão horária do Paraíba do Sul, é fundamental compreender que o comportamento do Rio Doce se difere desse rio, pois o aumento do volume se caracteriza por ser gradual e lento por causa do seu nível de assoreamento.

A escolha da organização dos dados (mensais, horários, etc.) depende dos objetivos da pesquisa. O estudo realizado na África do Sul optou por previsões mensais, uma vez que seu foco estava na avaliação do volume de água do reservatório local para geração de energia. Por outro lado, a previsão horária no Rio Paraíba do Sul se justifica devido à proximidade da cidade de Volta Redonda em relação ao leito do rio e sua rápida subida.

A previsão com 6 horas de antecedência possibilita uma resposta ágil à elevação do nível do rio, enfatizando a importância de selecionar modelos de previsão adequados para as particularidades de cada situação.

A Figura 5 ilustra o desempenho de validação do modelo de previsão horária durante os 10 dias próximos aos momentos mais críticos da inundação em 2022 na cidade de Governador Valadares.

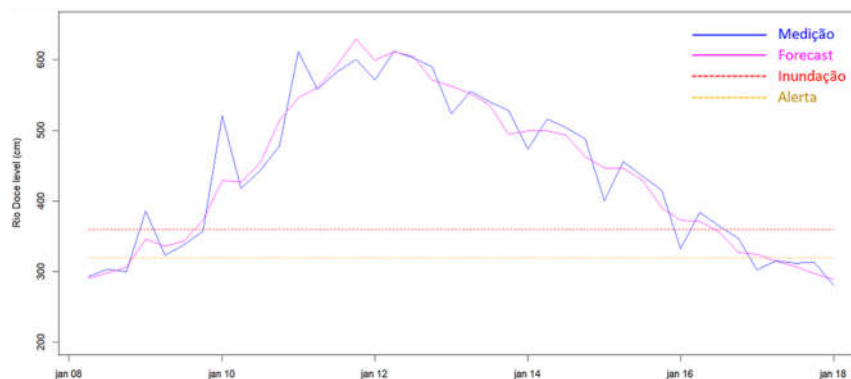


Figura 5: Nível do Rio Doce, dados reais versus dados simulados

É evidente que o modelo demonstrou um ajuste notável em relação às medições reais, conseguindo prever com precisão as fases de ascensão, picos e declínio na curva.

Conclusões

Este estudo destaca a eficácia do modelo SARIMA na previsão hidrológica, apresentando baixo erro de modelagem (MAPE de 0,64% e RMSE de 0,0224) para o Rio Doce. Uma característica fundamental deste estudo é a previsão com 6 horas de antecedência. Essa abordagem difere dos modelos

de previsão horária tradicionais, sendo adotada com base nas características específicas do Rio Doce e considerando o impacto das enchentes na cidade de Governador Valadares.

O SARIMA se revela uma ferramenta valiosa para prever enchentes de maneira eficaz. Recomenda-se expandir a pesquisa para incluir previsões horárias e mensais, permitindo comparações adicionais com estudos anteriores. Além disso, pretende-se explorar a implementação do método GLM com PSO nos dados do Rio Doce e comparar seus resultados com o desempenho do SARIMA, com o objetivo de aprimorar ainda mais as previsões em benefício da comunidade local.

Referências

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (ANA). **HIDROWEB**. 2023. Disponível em: <<https://www.ana.gov.br/>>. Acesso em: 7 set. 2023.

DICKEY, D. A., & FULLER, W. A. **Distribution of the estimators for autoregressive time series with a unit root**. Journal of the American Statistical Association, v. 47, p. 427-431, 1979.

DOOCY, S., DANIELS, A., MURRAY, S., and KIRSCH, T. D.. **The human impact of floods: a historical review of events 1980–2009 and systematic review**. PLoS Currents, v. 5, 2013.

HYNDMAN RJ, KHANDAKAR Y. **Automatic time series forecasting: the forecast package for R**. Journal of Statistical Software, v. 26, n. 3, p. 1-22, 2008.

LJUNG, G. M., & BOX, G. E. P.. **On a measure of a lack of fit in time series models**. Biometrika, v. 65, n. 2, p. 297-303, 1978.

MEZA, L. A. L.. **Modelo de Aprendizado de Máquina Combinado com Algoritmo de Exame de Partículas para a Emissão de Alerta de Inundações no Município de Volta Redonda**. 2021. 110 f. Pós-Graduação em Modelagem Computacional em Ciência e Tecnologia–Uff, Volta Redonda, 2021. Dissertação de Mestrado.

MORETTIN, P. A., & Toloi, C. M.. **Análise de Séries Temporais**. São Paulo: Blucher, 2006.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Viena, 2021.

SERVIÇO AUTÔNOMO DE ÁGUA E ESGOTO (SAAE). **Nível do Rio Doce**. 2022. Disponível em: <<https://www.saaev.com.br/>>. Acesso em: 12 dez. 2022.

TADESSE, K. B., & DINKA, M. O.. **Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa**. Journal of Water and Land Development, v. 35, p. 229-236, 2017.

UNITED NATIONS OFFICE FOR DISASTER RISK (UNDRR). **Global Assessment Report on Disaster Risk Reduction 2015**. 2015. Disponível em: <<https://www.preventionweb.net/>>. Acesso em: 8 set. 2023.

WORLD ECONOMIC FORUM. **The Global Risks Report 2022**. 2022. Disponível em: <<https://www.weforum.org/>>. Acesso em: 8 set. 2023.

UTILIZAÇÃO DE SÉRIES TEMPORAIS PARA PREVER VALORES DE BITCOIN COM MODELOS ARIMA E LSTM REVISÃO SISTEMÁTICA

Thiago Brum Leite

Universidade Federal Fluminense

thiagobrum@id.uff.br

Eliane da Silva Christo

Universidade Federal Fluminense

elianechristo@id.uff.br

Cecilia Toledo Hernandez

Universidade Federal Fluminense

ctoledo@id.uff.br

Resumo: O mercado de investimento em criptomoedas tem experimentado um crescimento exponencial desde o surgimento do Bitcoin em 2009. No entanto, a previsão dos valores associados a esse novo ativo financeiro, um aspecto fundamental para orientar os investidores sobre o melhor momento para comprar e vender seus Bitcoins, permanece como um grande desafio. Desde 2017, observa-se um aumento significativo no número de publicações que empregam análises de séries temporais para projetar valores de diversas criptomoedas. O objetivo deste artigo é conduzir uma revisão sistemática destinada a identificar os modelos que têm sido objeto de estudo e que têm demonstrado melhor desempenho, bem como implementá-los. Após identificá-los, são realizados experimentos de modelagem, análise de desempenho e comparação entre modelos ARIMA e LSTM. Os resultados mais promissores surgiram da sinergia entre essas duas abordagens, indicando que uma abordagem híbrida pode proporcionar vantagens na previsão dos valores do Bitcoin. Existem diversos modelos de previsão de séries temporais, tanto modelos estatísticos quanto modelos de aprendizagem de máquina, por isso é proposto que sejam utilizadas diferentes combinações de modelos a fim de otimizar os resultados na previsão dos valores do Bitcoin.

Palavras-chave: Séries Temporais. Bitcoin. Previsão. Revisão Sistemática. ARIMA. LSTM.

1. Introdução

A análise de séries temporais compreende a avaliação de conjuntos de observações ao longo do tempo, abrangendo diversos contextos, como dados climáticos, vendas empresariais, resultados médicos e valores financeiros. Suas principais finalidades, conforme descritas por (MORETTIN e TOLOI, 2006), são: 1) compreender a geração da série; 2) realizar previsões de curto e longo prazo; 3) descrever os comportamentos da série, identificando tendências e padrões; e 4) identificar periodicidades relevantes nos dados.

Em meio a esse cenário de avanço tecnológico, (NAKAMOTO, 2008) publicou o artigo "Bitcoin: A Peer-to-Peer Electronic Cash System", cujo objetivo foi criar uma moeda digital descentralizada e um sistema de pagamento eletrônico que não dependesse de uma autoridade central, como um governo ou um banco central. Nakamoto criou o Bitcoin, a primeira criptomoeda, e introduziu o conceito de Blockchain. Atualmente, existem mais de 20.000 criptomoedas, mas este mercado ainda é volátil e os preços das criptomoedas apresentam oscilações severas, o que apresenta um grande risco aos investidores. O Bitcoin e o Ethereum são as principais criptomoedas, representam 40% e 20% do total do mercado, respectivamente, enquanto as outras criptomoedas dividem os 40% restantes.



Figura 1: Gráfico do preço do Bitcoin de 20/12/2016 a 17/12/2018

Fonte: TradingView

Em 2017, o Bitcoin registrou um notável crescimento de 2132,42%, seguido por uma subsequente retração de 83,13% no ano seguinte, exibindo assim um grau exorbitante de volatilidade intrínseca. Este acentuado aumento experimentado em 2017 instigou um notório incremento na pesquisa e investigação relacionados ao Bitcoin, inclusive no escopo da análise de séries temporais, com o propósito de antecipar seu valor futuro. Em vista deste contexto marcado por flutuações extremas, torna-se imperativo para os investidores dispor de instrumentos analíticos que lhes permitam antever tanto os movimentos ascendentes quanto os declinantes desta criptomoeda.

2. Metodologia

2.1. Teste de Shapiro-Wilk

De acordo com (SHAPIRO, 1965), o teste de Shapiro-Wilk avalia a normalidade de uma amostra de dados. A hipótese nula (H_0) presume que os dados seguem uma distribuição normal, enquanto a hipótese alternativa (H_1) sugere o contrário. A estatística do teste (W) baseia-se nas discrepâncias entre os valores observados e os esperados sob uma distribuição normal. Calcula-se um valor-p, e se for menor que um nível de significância geralmente definido em 0,05, rejeita-se a hipótese nula, indicando que os dados não são normalmente distribuídos.

2.2. Transformação de Box-Cox

A transformação de Box-Cox (BOX, 1981) é uma técnica estatística empregada para estabilizar a variância e aproximar os dados de uma distribuição normal e uma variância constante.

A formulação da transformação Box-Cox é a seguinte:

$$\begin{aligned} Y_i &= x_i^\lambda, & \text{se } \lambda > 0 \\ Y_i &= \ln(x_i), & \text{se } \lambda = 0 \\ Y_i &= (x_i^\lambda), & \text{se } \lambda < 0 \end{aligned} \quad (1)$$

Onde:

$Y(\lambda)$ é a série transformada.

Y são os dados originais.

λ é o parâmetro de transformação. Este parâmetro pode assumir qualquer valor real, e a escolha do valor ótimo de λ é fundamental para a eficácia da transformação.

2.3. ARIMA

ARIMA (Auto Regressive Integrated Moving Average), é um modelo estatístico amplamente utilizado para análise e previsão de séries temporais, conforme explicado por (BOX, 2008), é uma abordagem matemática essencial para a analisar e entender séries temporais e previsões.

De acordo com (BAYER, 2010), ele representa séries temporais como um processo estocástico expresso matematicamente.

Conforme explicado por (ZHANG, 2003), um modelo ARIMA é denotado como ARIMA(p, d, q), onde "p" é a ordem do componente autorregressivo, "d" é a ordem de diferenciação e "q" é a ordem do componente de média móvel.

(PINTO MOREIRA DE SOUZA, 2017) demonstram que a escolha desses valores é crucial na modelagem ARIMA, e cada parâmetro é um número inteiro maior ou igual a zero.

2.4. Redes LSTM

As Redes Neurais de Memória de Longo Prazo (LSTM), introduzidas por (HOCHREITER e SCHMIDHUBER, 1997), são uma evolução das Redes Neurais Recorrentes (RNN) que abordam a limitação relacionada à memória. Enquanto as RNN tradicionais enfrentam desafios ao capturar conexões de longo alcance na rede, (GREFF, 2016) demonstra que as LSTMs foram projetadas para resolver esse problema incorporando um módulo de repetição especial com quatro camadas de redes neurais interagindo de maneira única.

O elemento central das LSTMs é o estado da célula, que percorre toda a estrutura da rede e pode ser modificado por meio de portões. Esses portões consistem em uma camada de rede neural com função de ativação sigmoide e operação de multiplicação, permitindo controlar a quantidade de informações que entra ou sai do estado da célula LSTM. Essa arquitetura foi projetada para lidar eficazmente com dependências de longo prazo em dados sequenciais.

3. Resultados

Foram escolhidos dois softwares, sendo um para a Revisão Sistemática e outro para a modelagem. Para a Revisão Sistemática foi escolhido o StArt® (State of the Art through Systematic Review), por ser um Software gratuito, desenvolvido com intuito de organizar e auxiliar o processo de Revisão Sistemática. Para a modelagem foi escolhido o software R® com os pacotes de extensão *forecast*, *MASS*, *tseries*, *keras*, *tensorflow* e *reticulate* que fornecem as funções necessárias aos testes realizados.

3.1. Análise da Revisão Sistemática

Foi realizada uma Revisão Sistemática, pesquisando artigos, teses e dissertações, publicados a partir de 2017, contendo as palavras-chave: “séries temporais”, previsão, bitcoin, volatilidade; “time series”, prediction, cryptocurrency, forecast, volatility.

Após aplicação dos critérios, foram selecionados 13 trabalhos para compor a Revisão Sistemática. Onde foi possível identificar quais os métodos que têm apresentado melhores resultados e os mais utilizados dentre estes.

MODELO	ARTIGOS	QUANTIDADE
ARIMA	(AMJAD e SHAH, 2017), (GARCIA, 2021), (SILVA, 2021), (DERBENTSEV, 2019), (DESEV, 2019)	5

LSTM	(ADEGBORUWA, 2019), (GARCIA, 2021), (SILVA, 2021)	3
RNN	(KODAMA, 2017), (ALMEIDA, 2019), (ARGUELLES, 2018)	3
HMM	(KODAMA, 2017), (SILVA, 2021)	2
SVM	(GARCIA, 2021), (ARGUELLES, 2018)	2

Tabela 1: Modelos mais utilizados

Dentre os métodos utilizados nos artigos, o que apresenta o maior número de aparições são: o ARIMA, com um total de 5 utilizações entre os artigos. Em segundo lugar temos o LSTM e o RNN empatados, com 3 aparições cada e em terceiro lugar temos o HMM e o SVM com duas aparições cada. Portanto foram selecionados os modelos ARIMA e LSTM para serem implementados.

3.2. Análise da Série Temporal

Foram capturados dados de valores diários do Bitcoin, no período de 17/09/2014 até 24/05/2023. Para uma primeira análise da série, foi construído o gráfico da mesma.



Figura 2: Gráfico do preço do Bitcoin diário de 17/09/2014 a 24/05/2023

Apesar de grandes oscilações, é possível verificar que a série apresenta forte tendência de crescimento. No teste de sazonalidade, usamos a função *decompose* no R®. Esta função pode nos retornar o período em que a sazonalidade ocorre ou simplesmente um valor zero quando não há sazonalidade. No caso da série que estamos analisando, a função retornou zero, o que significa que não há sazonalidade presente.

3.3. Análise da Normalidade

Foi realizado o teste de Shapiro-Wilk no R®, através da função *shapiro.test* para verificar a normalidade da série. O valor da estatística de teste (W) foi 0.786, porém o valor-p ($2.2e-16$) é praticamente zero, o que indica que os dados não seguem uma distribuição normal. Com base neste resultado, podemos concluir que os dados não são normalmente distribuídos.

3.4. Tentativa de Normalização da Série - Transformação de Box-Cox

Para tentar normalizar a série foi utilizada a transformação de Box-Cox, através da função *BoxCox* do R®. Após aplicar a transformação de Box-Cox, o teste de Shapiro-Wilk apresentou resultados semelhantes ao anterior com um valor-p praticamente zero. Indicando que a série não foi normalizada.

3.5. Análise de Modelos ARIMA

Foram realizados testes utilizando o *auto.arima* do R. O melhor modelo selecionado para a série foi o ARIMA(0,1,0). Isso significa que não há termos de autorregressão ($p = 0$) ou termos de média móvel ($q = 0$) no modelo. O termo de diferenciação é $d = 1$, indicando que a série temporal foi diferenciada uma vez para torná-la estacionária.

```
> auto.arima(serie_temporal, trace=TRUE, approximation = FALSE)
ARIMA(2,1,2)(1,0,1)[365] with drift : Inf
ARIMA(0,1,0) with drift : 49692.67
ARIMA(1,1,0)(1,0,0)[365] with drift : Inf
ARIMA(0,1,1)(0,0,1)[365] with drift : Inf
ARIMA(0,1,0) : 49691.01
ARIMA(0,1,0)(1,0,0)[365] with drift : Inf
ARIMA(0,1,0)(0,0,1)[365] with drift : Inf
ARIMA(0,1,0)(1,0,1)[365] with drift : Inf
ARIMA(1,1,0) with drift : 49692.9
ARIMA(0,1,1) with drift : 49692.93
ARIMA(1,1,1) with drift : 49694.92

Best model: ARIMA(0,1,0)

Series: serie_temporal
ARIMA(0,1,0)

sigma^2 = 642981: log likelihood = -24844.5
AIC=49691.01 AICC=49691.01 BIC=49697.03
```

Figura 3: Função Auto.Arima para recomendação do modelo

Fonte: Autor

De acordo com os resultados dos testes de ajuste do modelo, o mais aderente apresentou somente uma diferenciação, sem coeficientes significativos para a parte autorregressiva e de média móvel. Por isso, será feito um modelo híbrido com LSTM.

3.6. Análise de Redes LSTM

Foi desenvolvido um código em R® para criar, treinar e avaliar um modelo LSTM (Long Short-Term Memory). O protocolo de desenvolvimento se deu nas seguintes etapas:

- Carregar e Pré-processar Dados: Para evitar que números muito grandes ou muito pequenos causem instabilidade nos cálculos internos da rede, os dados históricos do Bitcoin foram carregados e submetidos a um processo de normalização. Isto é, os extremos da série foram ajustados para se situarem dentro da faixa entre 0 e 1.
- Criar Sequências de Dados: Os dados foram divididos em sequências de entrada e saída para o treinamento do modelo, usando uma janela de tempo deslizando.
- Dividir em Treinamento e Teste: Os dados foram divididos em conjuntos de treinamento e teste para avaliar o modelo.
- Construir o Modelo LSTM: Um modelo LSTM foi construído com uma camada LSTM, uma camada de *dropout* para evitar *overfitting* e uma camada de saída densa.
- Compilar o Modelo: O modelo foi compilado com uma função de perda e otimizador especificados.
- Treinar o Modelo: O modelo foi treinado com os dados de treinamento, especificando o número de épocas e tamanho do lote.
- Fazer Previsões: O modelo treinado foi usado para fazer previsões dos preços do Bitcoin.
- Avaliar o Modelo: O desempenho foi avaliado utilizando as métricas MSE, RMSE, MAE e MAPE.

3.6.1 Modelo LSTM com dados originais

Assim, procedeu-se à inclusão da série temporal contendo os preços diários do Bitcoin, registrados no período de 17/09/2014 a 24/05/2023, na Rede LSTM. Os resultados desse processo estão documentados na tabela 2.

Erro	LSTM
MSE	0.01929618
RMSE	0.1389107
MAE	0.1108154
MAPE	23.6%

Tabela 2: Cálculo dos Erros do LSTM

3.6.2 Modelo LSTM com dados diferenciados

Após constatar que o modelo ARIMA(0,1,0), abordado no item 3.5, se restringe a apenas uma diferenciação da série, ausente de componentes AR e MA, a decisão foi tomada para examinar o desempenho do modelo LSTM empregando esta série diferenciada. Os resultados dessa análise são detalhados na tabela 3.

Erro	LSTM
MSE	0.008603095
RMSE	0.09275287
MAE	0.06124054
MAPE	13.7%

Tabela 3: Cálculo dos Erros do LSTM utilizando dados diferenciados

Na Tabela 3, observa-se uma melhora significativa nos resultados quando o modelo é executado com os dados diferenciados. Evidencia-se um aprimoramento de aproximadamente 40% em comparação com o uso dos dados em sua forma original.

4. Conclusões

Através da Revisão Sistemática, identificou-se que, entre os modelos que apresentaram os melhores resultados, ARIMA e LSTM se destacaram como os mais pesquisados. Entretanto, ao implementá-los, uma abordagem ARIMA(0,1,0) foi sugerida como a mais aderente, indicando uma simples diferenciação na série temporal. Notavelmente, essa diferenciação serviu como base para o aprimoramento dos resultados do modelo LSTM, que apresentou um importante ganho de desempenho, atingindo um aumento de aproximadamente 40%.

Esses resultados revelam que as redes neurais, quando combinadas com modelos estatísticos, podem demonstrar melhorias notáveis em sua capacidade preditiva. Vale ressaltar que, embora a Revisão Sistemática tenha mencionado a existência de pesquisas que exploraram outras arquiteturas de redes neurais, como o RNN, a integração desses modelos com o ARIMA não pôde ser realizada devido a restrições de tempo.

Como sugestão para trabalhos futuros, indica-se a realização de análises abrangentes, explorando diferentes combinações entre modelos estatísticos e algoritmos de aprendizado de máquina, a fim de aprofundar o entendimento das sinergias potenciais entre essas abordagens distintas.

Bibliografia

- ADEGBORUWA, T. I. . A. S. A. . & B. M. M. Time series analysis and prediction of bitcoin using long short term memory neural network. **In 2019 15th International Conference on Electronics, Computer and Computation (ICECCO)**, Dezembro 2019. 1-5.
- ALMEIDA, C. L. Predição de séries temporais aplicada ao mercado de criptomoedas. **UFOP**, 2019.
- AMJAD, M.; SHAH, D. Trading bitcoin and online time series prediction. **NIPS 2016 time series workshop**., 2017. p. 1-15.

- ARGUELLES, D. C. M. Predicting the direction, maximum, minimum and closing price of daily/Intra-daily bitcoin exchange rate using batch and online machine learning techniques. **UFSCAR**, 2018.
- BAYER, F. M. Wavelets e modelos tradicionais de previsão: um estudo comparativo. **Rev. Bras. Biom**, 2010. p. 40–61.
- BIOLCHINI, J. E. A. Systematic review in software engineering. **System engineering and computer science department COPPE/UFRJ, Technical Report ES v. 679, n. 05**, 2005. 45.
- BOX, G. E. . & C. D. R. An analysis of transformations revisited, rebutted. **Mathematics Research Center**, University of Wisconsin-Madison, 1981. p. 0008.
- BOX, G. E. . J. G. M. . R. G. C. . & L. G. M. Time series analysis: forecasting and control. **John Wiley & Sons**, 2008.
- DERBENTSEV, V. . D. N. . S. O. . & B. V. Forecasting cryptocurrency prices time series using machine learning approach. **SHS Web of Conferences**, 2019. Vol. 65, p. 02001.
- DESEV, K. . K. S. . & D. D. Forecasting cryptocurrency markets through the use of time series models. **Business and Economic Horizons**, 2019. 242-253.
- FIGUEIREDO FILHO, D. B. . & S. J. J. A. Desvendando os Mistérios do Coeficiente de Correlação de Pearson. **Revista Política Hoje**, 2009. 18(1), 115-146..
- GARCIA, V. D. S. F. Séries temporais para predição de finanças no contexto de criptomoedas. **UFOP**, João Monlevade, 2021.
- GREFF, K. . S. R. K. . K. J. . S. B. R. . & S. J. LSTM: A search space odyssey. **IEEE transactions on neural networks and learning systems**, 2016. p. 2222-2232.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, 1997. p. 1735-1780.
- KITCHENHAM, B. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. 2004, 2004. p. 1-26.
- KODAMA, O. . P. L. . & K. Regime change and trend prediction for Bitcoin time series data. **CBU International Conference Proceedings**, 2017. pp. 384-388.
- MARGARIDO, M. A. . & D. M. J. H. Teste para mais de uma raiz unitária: uso do software sas® na elaboração de uma rotina para o teste Dickey-Pantula. **Pesquisa & Debate**, 2006. 29.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Blucher, 2006.
- NAKAMOTO, S. Bitcoin: A peer-to-peer electronic cash system. **Decentralized business review**, 2008.
- PINTO MOREIRA DE SOUZA, D. . D. S. C. E. . & R. A. A. Location of faults in power transmission lines using the ARIMA method. **Energies**, 2017. p. 1596.
- SHAPIRO, S. S. . & W. M. B. An analysis of variance test for normality (complete samples). **Biometrika**, 1965. 52(3/4), 591-611.
- SILVA, L. P. Q. . D. A. A. P. . C. D. O. . C. G. O. . & R. A. A. D. A. Utilizando HMM para previsão de preço e estratégia de investimento em criptomoedas BitCoin. **Anais do IV Workshop em Blockchain: Teoria, Tecnologias e Aplicações**, Agosto 2021. 134-147.
- STIGLER, S. Francis Galton's account of the invention of correlation. **Statistical Science**, 1989. 4(2), 73–79.
- WANG, Q.; SONG, X.; LI, R. A novel hybridization of nonlinear grey model and linear ARIMA residual correction for forecasting US shale oil production. **Energy**, v. 165, 2018. p. 1320-1331.
- ZHANG, G. P. Time series forecasting using a hybrid arima and neural network model. **Neurocomputing**, Elsevier, 2003. p. 159–175.

Comportamento do Número de Condicionamento na Formulação Variacional Ultra Fraca, com Funções de Bessel como Base, para a Equação de Helmholtz Não Homogênea

Julius Monteiro de Barros Filho

CEFET-RJ/campus Nova Iguaçu - PPGI/UFRJ

julius.filho1@cefet.br

Fernanda Lúcia Sá Ferreira

CEFET-RJ/campus Nova Iguaçu - PPGI/UFRJ

fernanda.ferreira@cefet.br

Amaury Alvarez Cruz

PPGI /UFRJ

amaury@dcc.ufrj.br

Daniel Gregorio Alfaro Vigo

PPGI /UFRJ

dgalfaro@ic.ufrj.br

Resumo: A Formulação Variacional Ultra Fraca (UWVF) se apresenta como uma metodologia promissora para a simulação de vários fenômenos ondulatórios. No entanto, o sistema linear oriundo da discretização dessa formulação pode ser bastante mal condicionado. Nessa pesquisa, fazemos uma análise do condicionamento do sistema linear subjacente em relação à escolha de certas famílias de funções de base

Palavras-chave: Formulação Variacional Ultra Fraca. Sistemas Mal Condicionados. Ondas Cilíndricas. Funções de Bessel.

Introdução

A aproximação numérica de problemas de propagação de ondas é um tema de grande relevância científica. A fidelidade da modelagem matemática em relação ao fenômeno físico e a complexidade computacional dos algoritmos utilizados são pontos de discussão fundamentais. Os métodos numéricos que utilizam ondas planas como funções de base têm sido amplamente estudados em física computacional e análise numérica, com o objetivo de aproximar a solução da equação de Helmholtz. O interesse por métodos que utilizam funções de base mais sofisticadas aumentou recentemente, visando os desafios encontrados em problemas de alta frequência.

O método da formulação variacional ultra fraca (UWVF) com funções de base de ondas planas, introduzido por Cessenat, mostrou-se uma ferramenta computacional útil para a solução numérica da equação de Helmholtz.

Nesta pesquisa, investigamos o uso de ondas cilíndricas como funções de base na UWVF para a equação de Helmholtz não homogênea, buscando soluções aproximadas altamente precisas. Em particular, consideramos as funções de Bessel de primeiro tipo como funções de base, explorando o comportamento do número de condicionamento do sistema linear em relação ao tipo e número de funções de base escolhidas, bem como o tipo de elemento finito utilizado na discretização do domínio computacional.

O Problema de Helmholtz

Seja Ω um meio não homogêneo. Consideremos o seguinte problema de Helmholtz:

$$\nabla \cdot \left(\frac{1}{\rho} \nabla u \right) + \frac{k^2}{\rho} u = f, \quad \text{em } \Omega$$

$$\left(\frac{1}{\rho} \frac{\partial u}{\partial \mathbf{n}} - i\sigma u \right) = Q \left(-\frac{1}{\rho} \frac{\partial u}{\partial \mathbf{n}} - i\sigma u \right) + g \quad \text{em } \Gamma$$

Note que, se $Q = -1$, a condição de fronteira é do tipo Dirichlet. Se $Q = 1$ a condição de fronteira é do tipo Neumann e se $Q=0$, a condição de fronteira é do tipo Robin. Na segunda equação, g é uma função conhecida.

Formulação Variacional Ultra Fraca (UWVF)

Na Formulação Variacional Ultra Fraca (UWVF) o domínio do problema é particionado em um conjunto de elementos finitos $\{K_1, K_2, \dots, K_N\}$, dois a dois disjuntos, formando uma malha. Vamos denotar por $\Sigma_{k,j}$ a borda entre os elementos K_k e K_j , por $\mathbf{n}_k$ o vetor normal em ∂K_k e a borda exterior por $\Gamma_k = \partial\Omega \cap \partial K_k$.

Além das características geométricas, a malha para a UWVF é construída de tal forma que, em cada elemento K_k , os parâmetros físicos ρ e k sejam constantes.

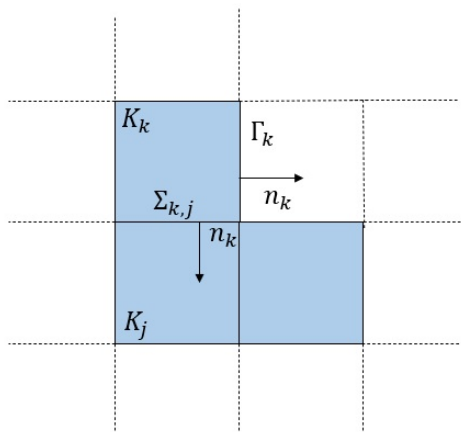


Figura 1: Malha particionada em elementos finitos.

Essa abordagem de particionar o domínio em elementos finitos permite uma representação mais localizada e adaptável da solução, permitindo uma melhor resolução do problema em regiões de interesse. A malha pode ser refinada, adicionando mais elementos finitos onde necessário, para obter uma solução mais precisa.

Com isso, o problema de Helmholtz pode ser decomposto em subproblemas, todos em meios homogêneos.

$$\Delta u_k + k_k^2 u_k = 0, \quad \text{em } K_k$$

$$\left(\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} - i\sigma u_k \right) = Q \left(-\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} - i\sigma u_k \right) + g \quad \text{em } \Gamma_k$$

Com condições de transmissão através de uma borda $\Sigma_{k,j}$ dadas por

$$u_k = u_j$$

$$\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} = -\frac{1}{\rho_j} \frac{\partial u_j}{\partial \mathbf{n}_j}.$$

Essas condições garantem a continuidade de u através das interfaces entre os elementos da malha. Então para cada elemento K_k temos o seguinte problema:

$$\Delta u_k + k_k^2 u_k = 0, \quad \text{em } K_k$$

$$\left(\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} - i\sigma_k u_k \right) = Q \left(-\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} - i\sigma_k u_k \right) + g \quad \text{em } \Gamma_k$$

$$u_k = u_j \quad \text{em } \Sigma_{k,j}$$

$$\frac{1}{\rho_k} \frac{\partial u_k}{\partial \mathbf{n}_k} = -\frac{1}{\rho_j} \frac{\partial u_j}{\partial \mathbf{n}_j} \quad \text{em } \Sigma_{k,j}$$

Supondo $u_k \in H^2(K_k)$ e definindo uma nova função no “esqueleto” da malha:

$$\chi_k = \left(\left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) u_k \right) |_{\partial K_k} \quad k = 1, 2, 3, \dots, N$$

obtemos, por integração, que

$$\begin{aligned} \sum_{k=1}^N \int_{\partial K_k} \frac{1}{\sigma_k} \overline{\left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} = \\ \sum_{k=1}^N \int_{\partial K_k} \left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) u_k \overline{\left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} - \sum_{k=1}^N -2i \int_{K_k} f \overline{v_k}. \end{aligned}$$

para todas as funções teste v_k que satisfaçam a equação de Helmholtz adjunta, ou seja,

$$\Delta \overline{v_k} - K_k^2 \overline{v_k} = 0.$$

Como $\partial K_k = \cup_j \Sigma_{k,j} \cup \Gamma_k$, podemos escrever

$$\begin{aligned} \sum_{k=1}^N \int_{\partial K_k} \frac{1}{\sigma_k} \overline{\left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} = \\ \sum_{k=1}^N \sum_j \int_{\Sigma_{k,j}} \frac{1}{\sigma_k} \overline{\left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} + \\ \sum_{k=1}^N \left[\int_{\Gamma_k} \frac{Q}{\sigma_k} \overline{\left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} + \int_{\Gamma_k} \frac{g}{\sigma_k} \overline{\left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k} \right] - \sum_{k=1}^N -2i \int_{K_k} f \overline{v_k}. \end{aligned}$$

Essa é a formulação variacional ultra fraca para o problema de Helmholtz em meios heterogêneos. Definamos ainda $F_k : L^2(\partial K_k) \rightarrow L^2(\partial K_k)$ por

$$F_k(y_k) = \left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k |_{\partial K_k},$$

onde $v_k \in H^1(K_k)$ satisfaz a equação de Helmholtz e $\left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) v_k = y_k$ em ∂K_k .

Com isso a UWVF se apresenta como:

$$\sum_{k=1}^N \int_{\partial K_k} \frac{1}{\sigma_k} \chi_k \overline{y_k} - \sum_{k=1}^N \sum_{j=1}^N \frac{1}{\sigma_k} \chi_k \overline{F_K(y_K)} -$$

$$\sum_{k=1}^N \int_{\Gamma_k} \frac{Q}{\sigma_k} \chi_k \overline{F_K(y_K)} = \sum_{k=1}^N \int_{\Gamma_k} \frac{g}{\sigma_k} \overline{F_K(y_K)} - 2i \sum_{k=1}^N \int_{K_k} f \overline{v_k}$$

Nessa formulação para UWVF as funções desconhecidas χ_k são calculadas em ∂K_k usando as funções teste y_k , que também são funções em ∂K_k . Uma vez calculados os χ_k 's, temos que, pelas condições de transmissão

$$\chi_k + \chi_j = -2i\sigma u_j \quad \text{em } \Sigma_{k,j}.$$

Portanto,

$$u|_{\Sigma_{k,j}} = \frac{1}{2\sigma} (\chi_k + \chi_j).$$

Discretização

A UWVF é discretizada usando o método de Galerkin. Então devemos escolher subespaços de dimensão finita dos espaços $L^2(\partial K_k)$, $k = 1, 2, \dots, N$.

Para cada elemento K_k da malha, uma família finita de funções $\phi_{k,l}$, $l = 1, 2, \dots, n_k$ é escolhida satisfazendo a equação de Helmholtz adjunta em K_k e $\phi_{k,l} = 0$ em $\Omega - \overline{K_k}$. O espaço aproximante discreto de $L^2(\partial K_k)$ é formado por todas as funções y_k^a tais que

$$y_k^a = \sum_{l=1}^{N_k} y_{k,l} \left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) \phi_{k,l} \quad \text{em } \Sigma_{k,j} \quad k = 1, 2, \dots, N.$$

Analogamente

$$\chi_{k,l} = \sum_{l=1}^{N_k} \chi_{k,l} \left(-\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) \phi_{k,l} \quad \text{em } \Sigma_{k,j} \quad k = 1, 2, \dots, N,$$

onde $\{\chi_{k,l}\}_{l=1}^{N_k}$ são os parâmetros a serem calculados. Note que $F_k(y_k^a) = \sum_{l=1}^{N_k} \chi_{k,l} \left(\frac{1}{\rho_k} \frac{\partial}{\partial \mathbf{n}_k} - i\sigma_k \right) \phi_{k,l}$. O problema aproximado para a UWVF é obtido substituindo o χ_k por χ_k^a e y_k por y_k^a .

O problema aproximado para a UWVF tem a seguinte escrita matricial:

$$(D - C)X = b,$$

onde $X = (\chi_{11}, \dots, \chi_{1N_k}, \chi_{21}, \dots, \chi_{2N_k}, \dots, \chi_{N1}, \dots, \chi_{NN_k})^T$ e

$$D = \begin{bmatrix} D_1 & 0 & 0 & \dots & 0 \\ 0 & D_2 & 0 & \dots & 0 \\ 0 & 0 & D_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & D_N \end{bmatrix},$$

onde é uma matriz Hermitiana e Positiva Definida dada por:

$$D_k^{l,m} = \int_{\partial K_k} \frac{1}{\sigma_k} \left(-\frac{1}{\rho_k} \frac{\partial \phi_{k,m}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,m} \right) \overline{\left(-\frac{1}{\rho_k} \frac{\partial \phi_{k,l}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,l} \right)}.$$

A matriz C também é estruturada por blocos

$$C = \begin{bmatrix} C_{11} & \dots & C_{1N_k} \\ \vdots & \ddots & \vdots \\ C_{N_k 1} & \dots & C_{N_k N_k} \end{bmatrix}$$

onde

$$C_{k,j}^{l,m} = \int_{\Sigma_{k,j}} \frac{1}{\sigma_k} \left(\frac{1}{\rho_k} \frac{\partial \phi_{j,m}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{j,m} \right) \overline{\left(\frac{1}{\rho_k} \frac{\partial \phi_{k,l}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,l} \right)} + \int_{\Gamma_k} \frac{Q}{\sigma_k} \left(-\frac{1}{\rho_k} \frac{\partial \phi_{k,m}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,m} \right) \overline{\left(\frac{1}{\rho_k} \frac{\partial \phi_{k,l}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,l} \right)}.$$

E por último temos que

$$b_j^k = \int_{\Gamma_k} \frac{g}{\sigma_k} \overline{\left(\frac{1}{\rho_k} \frac{\partial \phi_{k,l}}{\partial \mathbf{n}_k} - i\sigma_k \phi_{k,l} \right)} - 2i \int_{K_k} f \overline{\phi_{k,l}}.$$

Uma vez resolvido o sistema, a solução aproximada fica:

$$u|_{K_k} = \sum_{l=1}^{N_k} \chi_{k,l} \phi_{k,l} \quad \text{em } K_k.$$

Funções de Base

A seguir, apresentamos as famílias de funções de base que usamos para os experimentos numéricos nessa pesquisa.

- Base de ondas planas

$$\phi_{k,l}(\mathbf{x}) = \exp(i\overline{K_k} d_{k,l} \cdot \mathbf{x}) \quad \text{em } \Omega_k \quad \text{e} \quad 0 \quad \text{em } \Omega - \Omega_k,$$

$$d_{k,l} = \left(\cos \left(2\pi \frac{l-1}{p_k} \right), \sin \left(2\pi \frac{l-1}{p_k} \right) \right) \quad l = 1, 2, \dots, p_k$$

- Base de ondas cilíndricas (Funções de Bessel)

$$\phi_{k,l}(\mathbf{x}) = J_l(\overline{K_k} |\mathbf{x} - \mathbf{x}_{0,k}|) e^{il\theta} \quad \text{em } \Omega_k \quad \text{e} \quad 0 \quad \text{em } \Omega - \Omega_k,$$

onde l é a ordem da função de Bessel J_l e θ é o ângulo polar sobre $x_{0,k}$.

- Base de ondas cilíndricas com *Scaling*

$$\phi_{k,l}(\mathbf{x}) = \frac{J_l(\overline{K_k} |\mathbf{x} - \mathbf{x}_{0,k}|)}{J_l(\overline{K_k} L_k)} e^{il\theta} \quad \text{em } \Omega_k \quad \text{e} \quad 0 \quad \text{em } \Omega - \Omega_k,$$

onde L_k é o fator de *Scaling*.

- Base de Trefftz

$$\phi_{k,l}(\mathbf{x}) = H_0^1(K_k |\mathbf{x} - \mathbf{y}_{k,l}|) \quad \text{em } \Omega_k \quad \text{e} \quad 0 \quad \text{em } \Omega - \Omega_k,$$

onde

$$\mathbf{y}_{k,l} = \left(x_k^C + R \cos \left(2\pi \frac{l}{p_k} \right), y_k^C + R \sin \left(2\pi \frac{l}{p_k} \right) \right) \quad l = 1, 2, \dots, p_k$$

e H_0^1 é a função de Hankel.

Experimentos Numéricos

Vamos considerar o seguinte problema particular:

- A malha e os parâmetros
- $k = 0.5$;
- $u(x) = \frac{1}{4}H_0^1(K|x - x_0|)$.

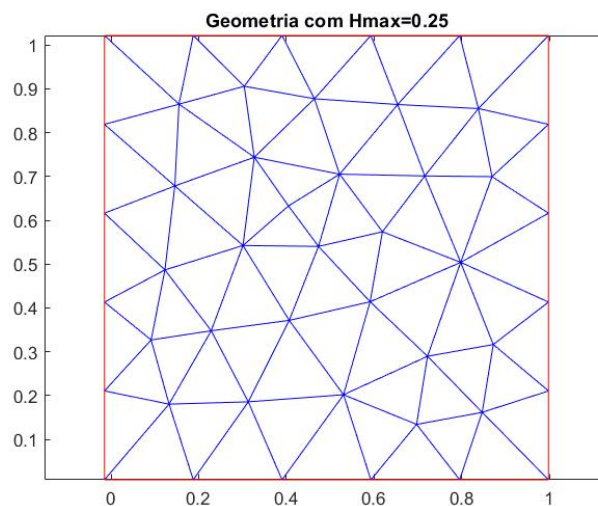


Figura 2: Malha inicial com aresta máxima 0.25.

Experimento: Número de Condicionamento x Malha

Aqui, analisamos o número de condição do sistema linear em função do tamanho da malha e do tipo de função de base. Os fatores de scaling para as ondas cilíndricas foram tomados como a maior das arestas em cada triângulo da malha. Já os sistemas lineares, foram resolvidos pelo método do Gradiente BiConjugado Estabilizado e o número de condição das matrizes foi avaliado na norma 2.

Em todos os casos, os gráficos revelam que o fator de scaling aqui adotado nas funções de Bessel foi capaz de reduzir drasticamente o número de condição dos sistemas lineares subjacentes, quando usamos malhas cada vez mais finas. Já em malhas menos refinadas, vemos que a mudança no tipo de função de base não é relevante e, considerando-se que as ondas planas admitem integração analítica, a escolha destas últimas seria mais indicado.

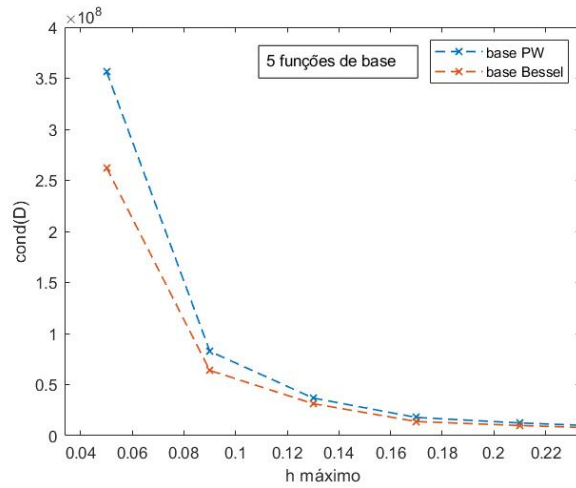


Figura 3: ondas planas versus ondas cilíndricas.

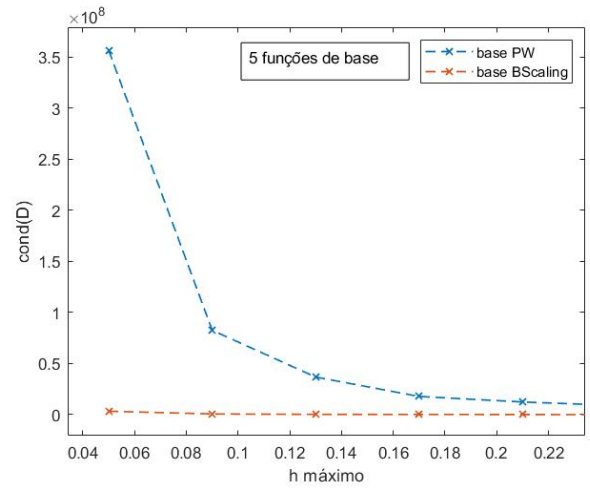


Figura 4: ondas planas versus ondas cil. scaling.

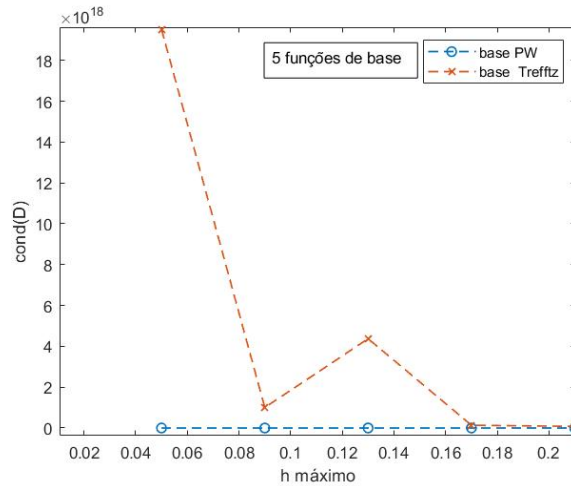


Figura 5: ondas planas versus base Trefftz.

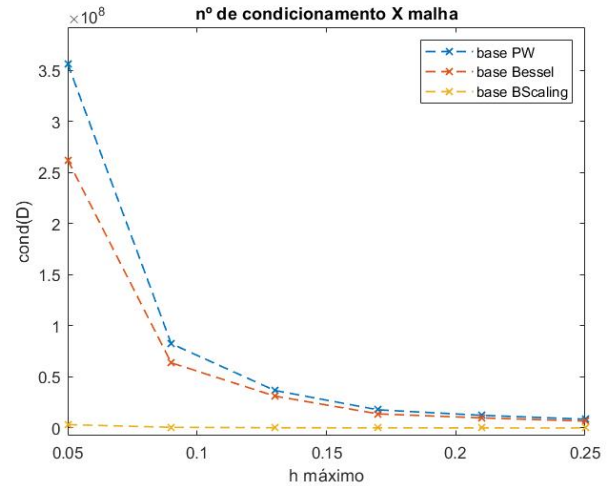


Figura 6: os 3 tipos de base.

Conclusões

Produzimos um método baseado em UWVF com funções de base usando funções de Bessel com scaling. Os experimentos numéricos apontam para uma grande vantagem no uso do fator de scaling adotado aqui, quando o foco é a redução do mal condicionamento do sistema linear que surge naquela metodologia. Como o mal condicionamento pode ser um limitante para algumas aplicações da UWVF, concluímos que o desenvolvimento de análise numérica para a formulação com funções de Bessel (com scaling) é bastante relevante.

Referências

- Alvarez, A. C. and García, G. C. and Sarkis, M. The ultra weak variational formulation for the modified mild-slope equation. **Applied Mathematical Modelling**, v. 52 p.28-41, 2017
- Cessenat, O. and Després, B. Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem. **SIAM Journal on Numerical Analysis**, 1998.

Computação numérica aplicada à Cosmologia: modelando o universo de Brans-Dicke-Rastall

Manoel Vicente de Souza Filho
Universidade Federal Fluminense - UFF
manoelvicente@id.uff.br

Thiago Roberto da Possa Caramês
Universidade Federal Fluminense
trpcarames@id.uff.br

Resumo: Neste trabalho, utilizamos computação numérica para construir um modelo cosmológico oriundo de uma teoria de gravidade alternativa à teoria da Relatividade Geral (RG), que constitui o cenário padrão em que se descreve os fenômenos gravitacionais. A teoria aqui utilizada, conhecida como teoria de Brans-Dicke-Rastall (BDR), consiste numa tentativa de conjugar aspectos da gravidade de Brans-Dicke com a propriedade não-conservativa da teoria de Rastall. Nela, o campo gravitacional é resultado tanto da distorção da geometria do espaço-tempo como da presença de um campo escalar dinâmico, denominado campo de Brans-Dicke ou dilaton. Além disso, assume-se que o tensor energia-momento não esteja mais sujeito à mesma lei de conservação da RG. No presente estudo, o propósito principal é o de obter a solução numérica do sistema de equações não-lineares que descrevem a dinâmica do modelo. Além disso, fazemos uma comparação da evolução do universo obtida a partir das equações dinâmicas com os dados observacionais do chamado parâmetro de Hubble.

Palavras-chave: Brans-Dicke-Rastall. Relatividade Geral. Parâmetro de Hubble. Dilaton. Fundo cosmológico

Introdução

O modelo padrão da cosmologia, o chamado Λ CDM (Λ *Cold Dark Matter*), tem se revelado muito bem sucedido na tarefa de descrever o universo observável. Esse modelo não apenas oferece uma descrição satisfatória da atual fase de aceleração cósmica que nosso universo atravessa, como também explica corretamente a dinâmica da formação das estruturas em grande escala que os dados observacionais revelam [1]. Apesar de suas virtudes, esse modelo pressupõe a existência de duas componentes materiais exóticas, cuja natureza ainda desconhecemos, a saber, a energia escura e a matéria escura. Essas duas componentes têm papéis muito bem definidos dentro do modelo: enquanto a energia escura é responsável pela aceleração cósmica, a matéria escura governa o processo de formação de estruturas no universo. Para entendermos o significado da sigla Λ CDM, lembremos que neste cenário, a energia escura é representada por um termo que está presente nas equações de campo da RG, a saber, a constante cosmológica, denotada pela letra grega Λ . Além disso, para que seja eficiente na tarefa de conduzir o processo de formação de estruturas, a matéria escura deve ser não-relativística, aproximadamente sem pressão ou, em outras palavras, "fria", de onde deriva o termo *Cold Dark Matter*, literalmente, "matéria escura fria".

A natureza desses dois constituintes "escuros" ainda é desconhecida, razão pela qual tem sido comum tentativas de substituir os papéis de ambos por outros graus de liberdade, como por exemplo a modificação da gravidade, com propostas que sugerem que a RG trate-se apenas de um limite de uma teoria gravitacional mais geral.

Tentativas de ir além da RG foram algo comum ao longo de todo o século XX, desde que Einstein formulou essa importante teoria gravitacional. Uma das mais notáveis foi, certamente, a teoria de Brans-Dicke, proposta na década de 1960 por Carl H. Brans e Robert H. Dicke [2]. Trata-se de uma teoria escalar-tensorial, na qual o campo gravitacional é descrito tanto pelo tensor métrico, quanto por um campo escalar dinâmico. Nessa teoria, a constante gravitacional de Newton é substituída por um campo escalar dinâmico, e a partir da análise PPN (sigla para formalismo pós-Newtoniano parametrizado), os

resultados da RG são recuperados quando o parâmetro livre da teoria de BD tende ao infinito, $\omega \rightarrow \infty$. As motivações para formulação de teorias gravitacionais alternativas ao longo do século foram as mais diversas [3]. No caso da teoria de Brans-Dicke, trata-se de uma tentativa de incorporar o Princípio de Mach à RG [4].

Uma outra teoria alternativa surge em 1972, proposta por P. Rastall. Nela, o autor sugere uma generalização da lei da conservação padrão, na qual o tensor energia-momento está sujeito na RG. Assim, na teoria de Rastall [5], a lei de conservação usual, $\nabla^\mu T_{\mu\nu} = 0$, é substituída por uma relação fenomenológica cuja forma é $\nabla^\mu T_{\mu\nu} \propto \nabla_\nu R$, ou seja, proporcional ao divergente do escalar de curvatura de Ricci. Note que na ausência de curvatura, ambas as leis acima coincidem com àquela válida no espaço-tempo plano, isto é, $\nabla^\mu T_{\mu\nu} = 0$.

Em 1974, foi introduzida uma versão da teoria gravitacional de Brans-Dicke [6], em que se incorpora o aspecto não-conservativo da gravidade Rastall. Essa teoria, que chamamos de teoria de Brans-Dicke-Rastall (BDR), foi explorada com mais detalhes décadas mais tarde [7]. Nesse trabalho, os autores desenvolvem, no contexto dessa teoria BDR, estudos que abrangem consequências cosmológicas, soluções esfericamente simétricas e análise do formalismo PPN dessa teoria.

Nosso propósito no presente trabalho foi o de aplicar ferramentas de computação numérica para construir um modelo cosmológico alternativo ao modelo padrão Λ CDM. Como sabemos, a presença de uma constante cosmológica no Λ CDM possibilita um ajuste com os dados observacionais, mas, por outro lado, traz discordância com as previsões da teoria quântica de campos sobre a ordem de grandeza da densidade de energia do vácuo. Assim, ao propor o modelo em questão, nosso objetivo foi o de examinar como podemos dispensar a necessidade de uma constante cosmológica, sem descuidarmos das restrições que os dados observacionais impõem a qualquer modelo cosmológico alternativo.

1 Teoria de Brans-Dicke-Rastall

A teoria BDR [8, 9] é uma extensão da RG que conjuga aspectos de duas outras teorias gravitacionais, a saber, a teoria de Brans-Dicke [8] e a teoria de Rastall [9].

Conforme visto em [10], as equações dinâmicas de BDR são dadas por

$$T_{;\mu}^{\mu\nu} = \frac{(1-\lambda)\phi}{16\pi} R^{,\nu}, \quad (1)$$

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = & \frac{8\pi}{\phi} \left[T_{\mu\nu} - \frac{1-\lambda}{2(1-2\lambda)}g_{\mu\nu}T \right] \\ & + \frac{\omega}{\phi^2} \left[\phi_{;\mu}\phi_{;\nu} + \frac{\lambda}{2(1-2\lambda)}g_{\mu\nu}\phi_{;\rho}\phi^{;\rho} \right] \\ & + \frac{1}{\phi} \left[\phi_{;\mu;\nu} + \frac{(1+\lambda)}{2(1-2\lambda)}g_{\mu\nu}\square\phi \right], \end{aligned} \quad (2)$$

e

$$\square\phi = \frac{8\pi\lambda}{3\lambda-2(1-2\lambda)\omega}T - \frac{\omega(1-\lambda)}{3\lambda-2(1-2\lambda)\omega} \frac{\phi^{;\rho}\phi_{;\rho}}{\phi}, \quad (3)$$

sendo “,” a derivada ordinária e “;” a derivada covariante. Note que as equações acima formam um sistema de equações diferenciais não-lineares, que como tal não admite solução analítica, mas apenas numérica. Devemos notar ainda que as equações-padrão da RG são recuperadas a partir do sistema acima, quando tomamos $\lambda = 1$ e $\phi = 1/G$.

As equações acima valem para um sistema gravitante arbitrário. Se considerarmos que o sistema em questão é o nosso próprio universo, precisamos assumir a validade do Princípio Cosmológico, segundo

o qual o universo respeita a homogeneidade e isotropia em grandes escalas. Matematicamente, essa condição pode ser traduzida por duas hipóteses. A primeira é que adotemos a chamada métrica de Friedman-Lemaître-Roberson-Walker (FLRW), que pode ser expressa pelo elemento de linha abaixo:

$$ds^2 = dt^2 - a(t)^2(dx^2 + dy^2 + dz^2). \quad (4)$$

E a segunda pressupõe que na escala cosmológica a matéria se comporte como um fluido perfeito, para o qual o tensor energia-momento associado é

$$T^{\mu\nu} = (p + \rho)u^\mu u^\nu - pg^{\mu\nu}, \quad (5)$$

sendo p , ρ e u^μ a pressão, densidade de energia e o campo de velocidades do fluido, respectivamente.

Usando (4) e (5) no sistema de equações (1) - (3), chegamos às equações que descrevem a evolução cosmológica do modelo [7]:

$$\dot{\rho} + 3\frac{\dot{a}}{a}(1+w)\rho = -\frac{3(1-\lambda)}{8\pi}\phi \left[\frac{\ddot{a}}{a} + \frac{\dot{a}}{a}\frac{\ddot{a}}{a} - 2\left(\frac{\dot{a}}{a}\right)^3 \right], \quad (6)$$

$$3\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi\rho}{\phi} \left[\frac{1-3\lambda}{2(1-2\lambda)} + \frac{3(1-\lambda)}{2(1-2\lambda)}w \right] + \omega \left[\frac{2-3\lambda}{2(1-2\lambda)} \right] \left(\frac{\dot{\phi}}{\phi}\right)^2 + \left[\frac{3(1-\lambda)}{2(1-2\lambda)}\frac{\ddot{\phi}}{\phi} + \frac{3(1+\lambda)}{2(1-2\lambda)}\frac{\dot{a}}{a}\frac{\dot{\phi}}{\phi} \right], \quad (7)$$

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{\phi} \left[\frac{1-\lambda-(1+\lambda)w}{2(1-2\lambda)} \right] \rho + \omega \frac{\lambda}{2(1-2\lambda)} \left(\frac{\dot{\phi}}{\phi}\right)^2 + \frac{1+\lambda}{2(1-2\lambda)}\frac{\ddot{\phi}}{\phi} + \frac{5-\lambda}{2(1-2\lambda)}\frac{\dot{a}}{a}\frac{\dot{\phi}}{\phi}, \quad (8)$$

e

$$\frac{\ddot{\phi}}{\phi} + 3\frac{\dot{a}}{a}\frac{\dot{\phi}}{\phi} = \frac{8\pi\lambda}{3\lambda-2(1-2\lambda)\omega}(1-3w)\frac{\rho}{\phi} - \omega \frac{1-\lambda}{3\lambda-2(1-2\lambda)\omega} \left(\frac{\dot{\phi}}{\phi}\right)^2, \quad (9)$$

onde ".", denota a derivada com respeito ao tempo cósmico e $w \equiv p/\rho$ o chamado parâmetro da equação de estado. Como estamos tratando nossa componente material como matéria fria, sem pressão, então adotaremos $w = 0$.

Por conveniência, em nosso modelo definimos a evolução temporal do universo em termos do parâmetro *redshift*, que pode ser expresso em termos do tempo cósmico, através da inversão da equação abaixo [8]

$$t = \frac{2H_0^{-1}}{1+(1+z)^2}. \quad (10)$$

Podemos expressar a densidade de matéria ρ em termos de uma quantidade adimensional conhecida como parâmetro de densidade que, para a teoria BDR, pode ser definido como:

$$\Omega_m = \frac{8\pi\rho}{3H_0^2\phi} \frac{(1-3\lambda)}{2(1-2\lambda)}. \quad (11)$$

Usando (11) na equação da continuidade (6), podemos determinar a evolução temporal da matéria no nosso modelo em termos de Ω_m como

$$\Omega'_m(z) = \frac{3}{1+z}\Omega_m(z) + \frac{3(1-\lambda)H}{8\pi}\frac{\phi}{\rho_c} \left[H''(1+z) - 3H' + \frac{H'^2}{H}(1+z) \right], \quad (12)$$

em que a linha denota a derivada com respeito ao redshift, z . Além disso, podemos usar a equação de Friedmann modificada (20), para escrever a evolução temporal do parâmetro de densidade associado ao campo de Brans-Dicke, ϕ . Em termos do redshift, a equação que descreve essa evolução é dada por:

$$\Omega_\phi = \omega \frac{2-3\lambda}{6(1-2\lambda)} (1+z)^2 \left(\frac{\phi'}{\phi} \right)^2 + \frac{(1-\lambda)}{2(1-2\lambda)} (1+z)^2 \frac{\phi''}{\phi} - \frac{\lambda}{1-2\lambda} (1+z) \frac{\phi'}{\phi} + \frac{1-\lambda}{2(1-2\lambda)} \frac{H'}{H} (1+z)^2 \frac{\phi'}{\phi}, \quad (13)$$

que definimos a partir da equação (8), de modo a respeitar a condição $\Omega_m + \Omega_\phi = 1$, visto que no nosso modelo só temos duas componentes, matéria e energia escura, sendo esta última descrita por (13).

2 Integração numérica via Python

A partir do sistema de equações diferenciais fortemente acopladas que descrevem a dinâmica da teoria BDR e a descrição quantitativa da presença de matéria no universo dadas pelas equações (19 - 23), reescrevemos esse sistema em termos do redshift e o integramos numericamente usando computação em Python. Para tanto, foi utilizado a biblioteca interna de integração de equações diferenciais ordinárias da referida linguagem, *odeint*.

Visando a integração numérica do sistema e com base em motivações físicas verificáveis na teoria BDR (conforme pode ser visto em [10]), fixamos as seguintes condições iniciais:

$$\phi_0 = \frac{2[2\lambda + (3\lambda - 2)\omega]}{3\lambda - 2(1 - 2\lambda)\omega}, \quad (14)$$

$$0 \leq \phi'_0 \leq \frac{2[2\lambda + (3\lambda - 2)\omega]}{3\lambda - 2(1 - 2\lambda)\omega}, \quad (15)$$

$$H_0 = 70 \text{ Km s}^{-1} \text{ Mpc}^{-1}, \quad (16)$$

$$\rho_0 = 0.6 \rho_{c,0} \phi_0 \frac{1-3\lambda}{1-2\lambda}, \quad (17)$$

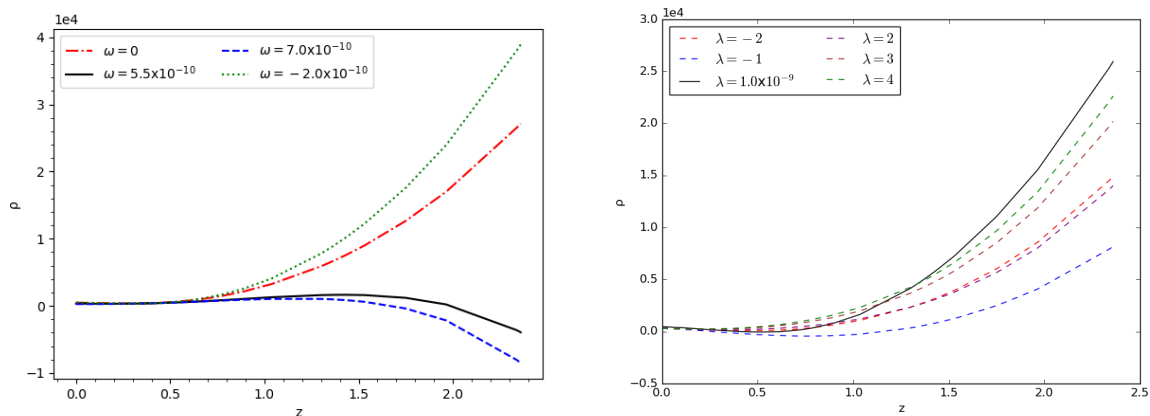
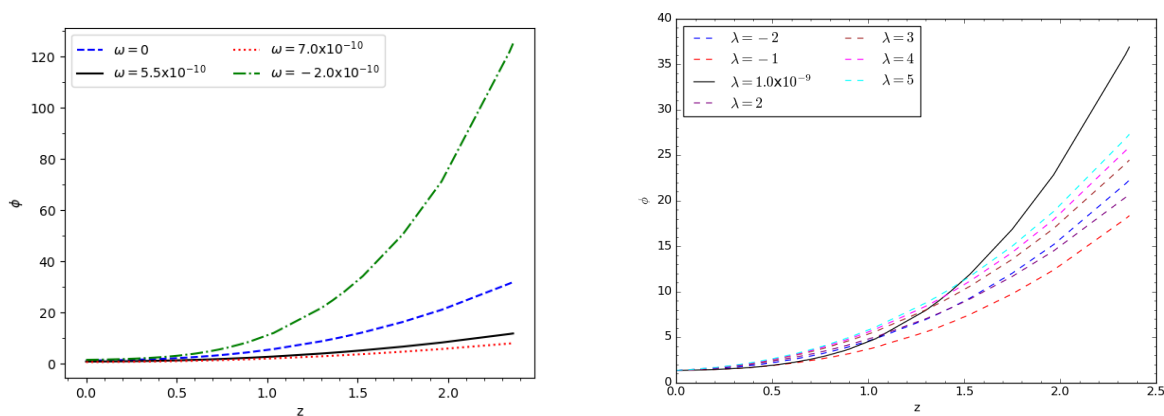
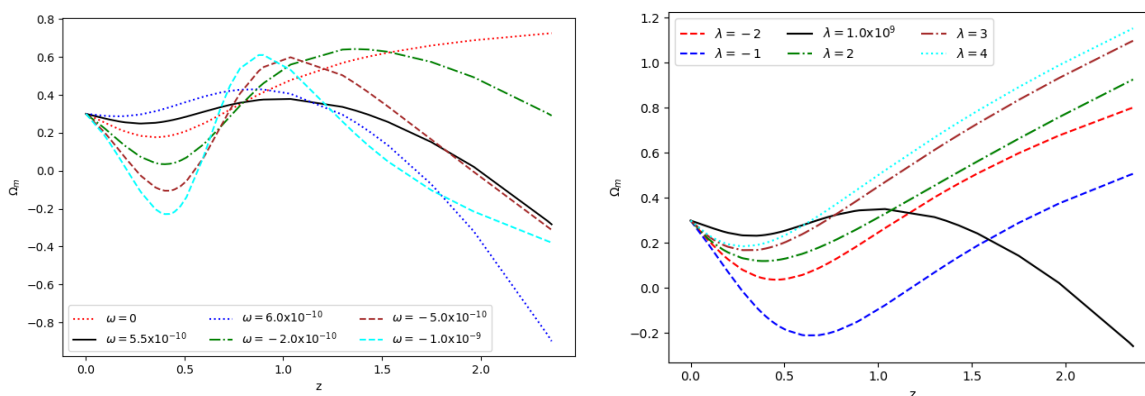
$$H'_0 = \frac{3}{2} \Omega_{m0} H_0, \quad (18)$$

e

$$\Omega_{m0} = 0.3, \quad (19)$$

de forma que o parâmetro de Hubble é $H = \dot{a}/a$, a constante gravitacional modificada para teoria BDR é $G = 2[2\lambda + (3\lambda - 2)\omega]/(3\lambda - 2(1 - 2\lambda)\omega)\phi$ e o parâmetro de densidade crítico corresponde a $\rho_{c,0} = 3H_0^2/8\pi G$.

Dessa maneira, os resultados da integração numérica do sistema supracitado são representados nos gráficos das Figuras [1], [2] e [3], correspondentes aos valores de ρ , ϕ e Ω_m , respectivamente. No gráfico à esquerda, fixamos $\lambda = 1.0 \times 10^{-9}$, enquanto no gráfico à direita, fixamos $\omega = 5.5 \times 10^{-10}$.


 Figura 1: Gráficos que representam a solução numérica da densidade de matéria, ρ .

 Figura 2: Gráfico que representa a solução numérica para o campo de Brans-Dicke, ϕ .

 Figura 3: Gráfico que representa a solução numérica para o parâmetro de densidade da matéria, Ω_m .

Entre todas as curvas acima representadas, é interessante notarmos que o caso $\omega = 0$, na coluna da esquerda, aponta na direção de um modelo cosmológico potencialmente viável. No primeiro plot, esse caso mostra que a densidade de matéria cai à medida que o universo atinge o presente (ou seja $z = 0$), que é justamente o que observamos ter acontecido no universo. No segundo plot, vemos o papel do campo de Brans-Dicke exercer uma importância no passado e se aproximar de uma constante no presente, conforme se espera que ocorra com campo de Brans-Dicke. E no terceiro, vemos o comportamento do parâmetro de densidade de matéria, que nos mostra claramente a queda da influência de Ω_m , ilustrando, assim, o

fim da era da matéria, dando lugar à atual era da energia escura. Os demais resultados não são aceitáveis, pois trazem valores negativos para ρ ou Ω_m em algum momento do passado, o que não é fisicamente aceitável.

Comparação com os dados observacionais de $H(z)$

Para plotarmos o gráfico da Figura [4], ajustamos as condições iniciais do sistema mencionadas anteriormente e variamos o parâmetro livre ω , de modo a obtermos resultados dentro das barras de erro provenientes dos dados observacionais para o parâmetros de Hubble $H(z)$. Além disso, espera-se também que o parâmetro de densidade do campo de Brans-Dicke ($\Omega_{\phi 0}$) e o parâmetro de densidade da matéria (Ω_{m0}) atuais sejam 70% e 30%, respectivamente, de forma que correspondam à mesma taxa de contribuição de energia escura e matéria na composição material do universo.

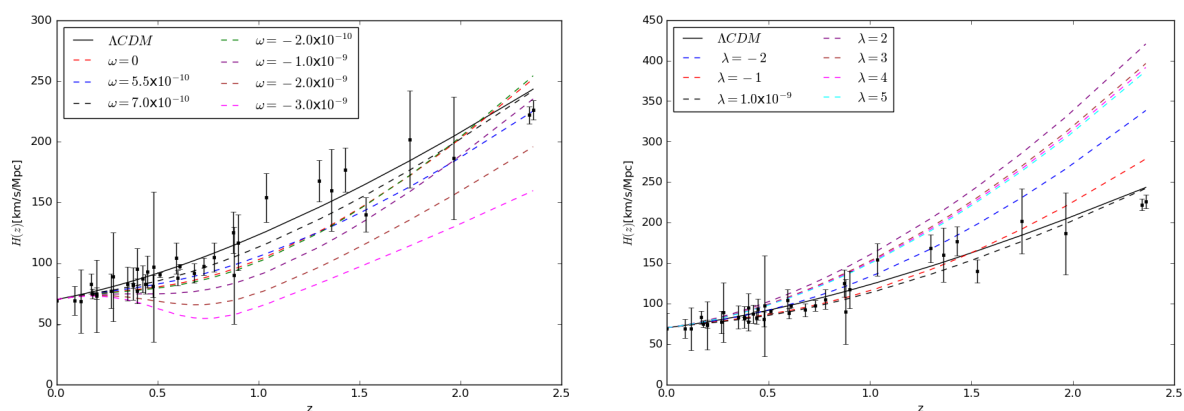


Figura 4: Curvas que descrevem o comportamento do parâmetro de Hubble para diferentes valores de ω são apresentadas em preto sólido para a RG e em tracejado para o modelo Brans-Dicke-Rastall. A curva do parâmetro de Hubble ($H(z)$) é uma representação gráfica da evolução da taxa de expansão do universo em diferentes redshifts. Essa curva é importante para entender a dinâmica do universo e para testar diferentes modelos cosmológicos, como os mencionados na discussão anterior sobre os modelos de Brans-Dicke e Rastall.

Conclusões

Ao combinar o tensor energia-momento da teoria de Brans-Dicke com a violação da conservação de energia da teoria de Rastall, emerge um modelo cosmológico que depende de dois parâmetros independentes: ω para Brans-Dicke e λ para Rastall, onde este último representa o grau de não-conservação do sistema. A evolução do fundo cósmico do universo, descrita pela curva do parâmetro de Hubble ($H(z)$), demonstra uma notável sensibilidade em relação a esses parâmetros.

Realizamos a integração numérica do sistema de equações não-lineares que caracterizam a dinâmica do modelo. Como o modelo possui dois parâmetros livres, optamos por dois caminhos: fixar um e deixar o outro livre e depois inverter o procedimento. Fazendo isso, fixando $\lambda = 10^{-9}$ e variando ω , chamou nossa atenção o caso $\omega = 0$ para o qual obtivemos resultados compatíveis com o que se espera da dinâmica cosmológica em grande escala.

Na tentativa de confrontarmos nossos resultados com dados observacionais, notamos que, à medida que ω diminui, a curva $H(z)$ torna-se mais sensível às mudanças em ϕ'_0 , afastando-se dos resultados previstos pela RG para um universo em estágio tardio. Porém, o caso $\omega = 0$, do qual falamos anteriormente, possui uma curva de evolução que se enquadra nos limites impostos pelas barras de erro desse conjunto de dados observacionais.

Há uma conexão experimental entre os parâmetros ω e λ , e ao fixar $\lambda = 0$, o modelo exibe instabilidade, resultando em curvas descontínuas em todo o espaço. No entanto, ao considerar λ próximo a zero e seguindo as restrições impostas pela teoria, é possível recuperar os testes locais da RG. Adicionalmente, quando o valor do parâmetro de Rastall, ω , é fixado em um módulo maior que 7, o modelo novamente se torna instável, resultando em curvas $H(z)$ pouco comportadas.

Agradecimentos

Expressamos nossos sinceros agradecimentos ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) por fomentar a nossa pesquisa através de apoio financeiro.

Referências

- [1] Dodelson, Scott, Evalyn I. Gates, and Michael S. Turner. "Cold dark matter." *Science* 274.5284 (1996): 69-75.
- [2] C. Brans e R. Dicke, *Phys. Rev.* 124, 925 (1961).
- [3] S. Nojiri, S.D. Odintsov, V.K. Oikonomou, *Phys.Rept.* 692, 1-104 (2017).
- [4] Weinberg, Steven. "Gravitation and cosmology: principles and applications of the general theory of relativity." (1972).
- [5] P. Rastall, *Phys. Rev. D* 6, 3357 (1972).
- [6] L. Smalley, *Phys. Rev. D* 9, 1635 (1974).
- [7] Caramês, Thiago RP, et al. "The brans–dicke–rastall theory." *The European Physical Journal C* 74.11 (2014): 3145.
- [8] L. Smalley, *Phys. Rev. D* 9, 1635 (1974).
- [9] Thiago R. P. Caramês, Mahamadou H. Daouda, Júlio C. Fabris, Adriano M. de Oliveira, Oliver F. Piattella e Vladimir Stokov, *Eur. Phys. J. C* 74, 3145 (2014).
- [10] Thiago R. P. Caramês, Mahamadou H. Daouda, Júlio C. Fabris, Adriano M. de Oliveira, Oliver F. Piattella e Vladimir Stokov, *Eur. Phys. J. C* 74, 3145 (2014).

IMPLEMENTAÇÃO EM VOLUMES FINITOS DE UM MÉTODO TIPO GODUNOV PARA SISTEMAS 2×2 DE LEIS DE CONSERVAÇÃO

Eron Oliveira do Nascimento
Universidade Federal Fluminense
eronoliveira@id.uff.br

Panters Rodríguez Bermúdez
Universidade Federal Fluminense
pantersrb@id.uff.br

Wagner Rambaldi Telles
Universidade Federal Fluminense
wtelles@id.uff.br

Resumo: No presente trabalho apresenta-se uma forma de implementação computacional do método de Godunov, o qual utiliza uma discretização em volumes finitos em 1D, e é direcionado para a resolução de sistemas 2×2 de equações diferenciais parciais hiperbólicas denominadas leis de conservação. O desenvolvimento é feito à partir de um processo de linearização para obter-se o valor do fluxo para os volumes, em conjunto, de um *Riemann solver* aproximado e condições de correção de entropia. A validação do código foi realizada tanto para as equações de águas rasas como para um modelo de escoamento horizontal e vertical de fluidos imiscíveis em meios porosos, para isto foram utilizados resultados analíticos e numéricos encontrados em distintos artigos científicos de impacto internacional. Neste artigo, por limitação de espaço, mostramos apenas o caso de escoamento trifásico vertical (considerando efeitos gravitacionais) em meios porosos, as soluções numéricas obtidas neste trabalho podem ser consideradas as primeiras aproximações numéricas das soluções exatas semi-analíticas encontradas e descritas qualitativamente em Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013).

Palavras-chave: Volumes finitos. Método de Godunov. Leis de conservação. Implementação computacional.

Introdução

Sistemas de leis de conservação estão presentes em diversos modelos de problemas da engenharia e fenômenos naturais. Estas são um conjunto de equações diferenciais parciais (comumente não lineares) de primeira ordem de caráter hiperbólico e representam fenômenos de transporte de distintos tipos. Devido à não linearidade destas equações diferenciais, a resolução das mesmas costuma apresentar dificuldades matemáticas e computacionais que, em geral, limitam a compreensão e a obtenção das soluções destes problemas da engenharia. Uma destas dificuldades matemáticas presente, é o conhecido fenômeno de formação de ondas de choque (frentes descontínuas que se transportam a velocidade constante), ondas de rarefação, etc., que se produzem como consequência do cruzamento entre as curvas características do sistema hiperbólico. Para poder estudar tais fenômenos, o conceito de solução de EDP deve ser estendido (soluções generalizadas ou fracas). Porém, soluções fracas costumam não ser únicas, e critérios de entropia adicionais para selecionar as soluções fisicamente corretas precisam ser considerados. Por outro lado, para sistemas de duas ou mais leis de conservação, a teoria existente de Lax (1973) e Liu (1976) sobre problemas de Riemann é muito mais complexa que para modelos escalares, o que dificulta enormemente a obtenção de soluções analíticas ou exatas para esses tipos de problemas. Complicações que surgem nos problemas físicos como heterogeneidades no meio, problemas multidimensionais no espaço (2D ou 3D), variações de condições como temperatura e pressão, efeitos de capilaridade ou interação entre partículas, entre outras acrescentam as dificuldades enfrentadas. Deste modo, a implementação de métodos numéricos robustos e precisos, se faz necessário para obter as soluções numéricas para os

problemas de Riemann para leis de conservação. Durante anos, a comunidade científica tem dedicado esforços ao estudo, criação, aprimoramento de esquemas numéricos acurados e à implementação de códigos computacionais (simuladores) para resolver estes problemas. Segundo LeVeque (2004), existem basicamente três tipos de formulações para implementar esses métodos (diferenças finitas, volumes finitos, elementos finitos). Neste artigo, é apresentada a implementação de um método tipo Godunov, baseado em volumes finitos, para a solução de problemas unidimensionais de sistemas 2×2 de leis de conservação. Utiliza-se a construção de matrizes lineares nas interfaces dos volumes, em substituição da matriz jacobiana das funções de fluxo. Como resultado, o problema é transformando em diversos sistemas locais, resolvidos pelo *Riemann solver* de Roe (1981). A condição de entropia para correção das soluções é definido pelo critério de Harten e Hyman (1983).

O código construído foi utilizado para resolver alguns problemas de escoamento trifásico em meios porosos com influência da gravidade, já previamente publicados na tese de Rodríguez-Bermúdez (2010) e no artigo de Rodríguez-Bermúdez e Marchesin (2013). Dadas condições iniciais de Riemann, os resultados consistem em perfis numéricos para as saturações das fases cuja descrição qualitativa é validada.

A perspectiva de construção do simulador visou a possibilidade de expandir o código para outros tipos de problemas e números de equações envolvidas, desde que modeladas por sistemas hiperbólicos.

Volumes finitos

A equação diferencial da lei de conservação pode ser formulada através da discretização do domínio em volumes finitos e usando o princípio da conservação de massa. Segundo LeVeque (2004), para o caso unidimensional:

$$\frac{d}{dt} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} Q(x,t) dx = f(Q(x_{i-\frac{1}{2}},t)) - f(Q(x_{i+\frac{1}{2}},t)), \quad (1)$$

onde Q é o vetor que agrupa as $\mathbf{m}$ variáveis conservadas e $f(Q)$ o vetor das $\mathbf{m}$ funções de fluxo do modelo. Ou seja, a variação no tempo da quantidade Q no volume i é igual à diferença entre os valores do fluxo nas interfaces de saída $x_{i+\frac{1}{2}}$ e entrada $x_{i-\frac{1}{2}}$.

Reduzindo o intervalo $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ em uma escala infinitesimal e assumindo certas condições de diferenciabilidade, a relação (1) é simplificada para a forma diferencial, sendo:

$$Q_t + f(Q)_x = 0. \quad (2)$$

em que os subíndices x, t indicam as derivadas parciais em relação ao espaço e o tempo, respectivamente. A expressão (2) descreve um sistema de EDP hiperbólicas na sua forma conservativa.

Retornando à forma integral (1), em intervalos infinitesimais, a função $Q(x,t)$ pode ser aproximada por um valor constante em cada volume $\mathbf{i}$, tornando as soluções das integrais apenas o produto do valor de $Q(x,t)$ pela variação temporal (Δt) ou espacial (Δx). Obtém-se assim uma expressão discretizada em que o cálculo do valor de cada Q_i no volume, na iteração de tempo seguinte $\mathbf{n} + 1$, é dado apenas por termos da iteração atual $\mathbf{n}$:

$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{\Delta x} (F_{i+\frac{1}{2}}^n - F_{i-\frac{1}{2}}^n), \quad (3)$$

onde $F_{i-\frac{1}{2}}^n$ e $F_{i+\frac{1}{2}}^n$ são os valores que a função de fluxo assume nas interfaces. A expressão (3) é a base usada no código implementado.

Método de Godunov

Segundo apresenta Godunov e Bohachevsky (1959), o sistema (2) pode ser resolvido através de um algoritmo no qual para cada volume i : aproxima $Q(x,t)$ por uma função polinomial, resolve o sistema

para as condições delimitadas nas interfaces e calcula a média do valor ao longo do próprio volume, processo esse chamado *REA: reconstruct-evolve-average*. Já LeVeque (2004) propõe uma adaptação deste conceito para calcular o valor dos fluxos $F_{i-\frac{1}{2}}^n$ e $F_{i+\frac{1}{2}}^n$ na expressão (3).

Fazendo a aproximação da função $Q(x, t)$ por um valor constante, os diversos sistemas locais em cada volume agora têm como condição inicial específica um problema de Riemann. Sendo a descontinuidade (interface) considerada a origem, $x_{i+\frac{1}{2}} = 0$ para esse problema local, os valores constantes Q_E e Q_D dos volumes adjacentes correspondem à definição da condição. Os valores $F_{i-\frac{1}{2}}^n$ e $F_{i+\frac{1}{2}}^n$ obtém-se a partir da solução do problema de Riemann avaliada nas respectivas interfaces, e depois avaliando a função de fluxo; denominam-se fluxo de Godunov (numérico).

Para a implementação desenvolvida neste trabalho é usada a forma não conservativa (4), requerido pelo *Riemann solver* aplicado, em que reescreve-se (2) sendo $A(Q)$ a matriz jacobiana do sistema hiperbólico.

$$Q_t + A(Q)Q_x = 0. \quad (4)$$

Dependendo do modelo estudado, a matriz $A(Q)$ pode apresentar complexidades, dificultando a resolução direta do problema de Riemann. Para a implementação, entretanto, foi usado um solucionador numérico aproximado para o problema de Riemann (*Riemann solver*). De acordo com Roe (1981), a matriz A pode ser reescrita com suas componentes sendo valores constantes, apenas sendo necessário atender às condições: (i) preservar o caráter hiperbólico do sistema, (ii) ser consistente à matriz original e (iii) satisfazer a condição de choque de Rankine-Hugoniot. Isto transforma o sistema em um caso linear, com maior facilidade de solução. A nova matriz $\tilde{A}$ é definida em cada interface, em substituição de $A(Q)$. A definição de $\tilde{A}$ é abrangente e apenas se faz necessário que atenda aos critérios acima mencionados. Como sugerido por Roe (1981) e Toro (2009), foi usado apenas a média simples entre os valores que a matriz assume para Q nas adjacências, facilitando a implementação e reduzindo o trabalho computacional: $\tilde{A} = \frac{1}{2} [A(Q_E) + A(Q_D)]$.

Assim, os sistemas resolvidos nas interfaces serão:

$$\begin{cases} Q_t + \tilde{A}(Q)Q_x = 0, \\ Q(x, 0) = \begin{cases} Q_E, & x < 0, \\ Q_D, & x > 0. \end{cases} \end{cases} \quad (5)$$

A solução dos sistemas (5), conforme desenvolve Toro (2009), é obtida pela expansão dos estados Q em função dos autovalores λ_j e autovetores $\kappa^{(j)}$ de $\tilde{A}$ e avaliada para $x = 0$ (interface). Aplicando à função de fluxo, a forma geral do fluxo de Godunov para um sistema de m equações, obtém-se:

$$F_{i+\frac{1}{2}} = \frac{1}{2} \left[f(Q_i) + f(Q_{i+1}) - \sum_{j=1}^m (\beta_j - \alpha_j) h(\lambda_j) \kappa^{(j)} \right], \quad (6)$$

$$F_{i-\frac{1}{2}} = \frac{1}{2} \left[f(Q_{i-1}) + f(Q_i) - \sum_{j=1}^m (\beta_j - \alpha_j) h(\lambda_j) \kappa^{(j)} \right], \quad (7)$$

onde α_j e β_j são os coeficientes da combinação linear de Q_i e Q_{i+1} , respectivamente.

A utilização do *Riemann solver* de Roe (1981), introduz erros decorrentes da aproximação da matriz $\tilde{A}$ em linear e gera soluções apenas na forma de choques, devido à condição de Rankine-Hugoniot. O critério de correção de entropia de Harten e Hyman (1983) foi escolhido e aplicado segundo a definição da função (8):

$$h(\lambda_j) = \begin{cases} \delta_j & \text{se } |\lambda_j| < \delta_j, \\ |\lambda_j| & \text{se } |\lambda_j| \geq \delta_j, \end{cases} \quad (8)$$

$$\delta_j = \max \{0, \lambda_j - \lambda_j(Q_E), \lambda_j(Q_D) - \lambda_j\}. \quad (9)$$

Sendo o fator δ_j um parâmetro para mensurar o quanto a solução aproximada desvia da solução correta, acrescentando um valor intermediário que suaviza os choques para formar rarefações corretas.

Tratando-se de um método explícito, está sujeito a um critério de estabilidade para a convergência da solução no código. Courant et al. (1967) define que para valores entre $0 < v < 1$ o método é convergente e estável e este princípio foi usado no código para calcular o intervalo de tempo Δt de cada iteração, conforme (10).

$$\Delta t = \frac{v \Delta x}{\lambda_{max}^n}. \quad (10)$$

Nos sistemas em estudo, para cada matriz $\tilde{A}$ na interface, existem dois autovalores tal que $\lambda^+ > \lambda^-$ indicando os grupos de onda *fast-family* e *slow-family*, respectivamente. A determinação do maior valor para cada iteração pode ser feita, segundo LeVeque (2004), pela escolha (11):

$$\lambda_{max}^n = \max_i \left(\left| \lambda_{i-\frac{1}{2}} \right|, \left| \lambda_{i+\frac{1}{2}} \right| \right). \quad (11)$$

Simulações numéricas

Para demonstrar a acurácia e robustez do método de Godunov implementado, foram selecionados alguns exemplos de escoamento trifásico em meios porosos assim como exemplos do modelo de águas rasas (shallow water). Neste artigo, apenas apresentamos alguns exemplos de escoamento em meios porosos. As variáveis conservadas de interesse neste problema são as saturações de cada fase: $Q = (q_1, q_2, q_3)$. Consideramos um escoamento com algumas simplificações, tais como: permeabilidade constante do meio (rocha homogênea), as densidades e viscosidades das fases são constantes, temperatura constante, entre outras, veja Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013).

Pelo fato de considerarmos o escoamento imiscível e o meio poroso totalmente saturado pelas fases, teremos que o problema pode ser descrito por duas variáveis e suas funções de fluxo correspondentes, tornando-se um sistema 2×2 . Neste estudo foram escolhidos: q_1 e q_3 como variáveis conservadas e q_2 calculada pela relação $q_2 = 1 - q_1 - q_3$.

O modelo proposto têm funções de fluxo obtidas à partir da lei de Darcy para infiltração e relaciona as permeabilidades relativas das fases com as suas respectivas saturações, dadas segundo uma expressão quadrática, modelo de Corey et al. (1956). As funções em (12) são descritas por Rodríguez-Bermúdez (2010):

$$f_1(q_1, q_3) = \frac{\frac{q_1^2}{\mu_1} \left(\alpha + \frac{q_3^2}{\mu_3} \rho_{13} + \frac{q_2^2}{\mu_2} \rho_{12} \right)}{\frac{q_1^2}{\mu_1} + \frac{q_2^2}{\mu_2} + \frac{q_3^2}{\mu_3}}, \quad f_3(q_1, q_3) = \frac{\frac{q_3^2}{\mu_3} \left(\alpha + \frac{q_1^2}{\mu_1} \rho_{31} + \frac{q_2^2}{\mu_2} \rho_{32} \right)}{\frac{q_1^2}{\mu_1} + \frac{q_2^2}{\mu_2} + \frac{q_3^2}{\mu_3}} \quad (12)$$

onde α é o fator adimensional que indica a razão entre efeitos de convecção por gradientes de pressão e convecção por gravidade. Quando esse valor é pequeno, os efeitos gravitacionais são predominantes, enquanto um valor maior, a gravidade é desprezível.

A diferença entre as densidades das fases i e j (nesta ordem) é dado por ρ_{ij} e μ_i representa a viscosidade dinâmica. O desenvolvimento das funções, das variáveis utilizadas e do tempo da solução estão em unidades adimensionalizadas.

A implementação tem como estrutura um código principal onde se calcula de forma iterativa a variável Q_i usando a expressão (3) junto do fornecimento, pelo usuário, das propriedades das fases e dos parâmetros numéricos (valor v , tempo de simulação, condições iniciais). As funções de fluxo (12),

a formação da matriz linearizada $\tilde{A}$ e os cálculos necessários para o valor do fluxo de Godunov (6) e (7) são escritos em formas de funções independentes, em um arquivo de código separado, permitindo a fácil alteração (adaptação) para outros tipos de problemas.

A malha utilizada apresenta 1000 pontos correspondentes aos volumes em domínio computacional unitário, o número de estabilidade CFL utilizado é $\nu = 0.1$, parâmetros estes determinados por tentativa. As condições iniciais são de tipo de Riemann (constantes por partes), fornecendo os valores das saturações das fases presentes na descontinuidade inicial (poço de injeção). Como condição de contorno, embora não necessária do modelo físico real, é usada a repetição dos valores calculados nas adjacências do contorno computacional. O código foi construído e simulado utilizando a linguagem (software) MATLAB.

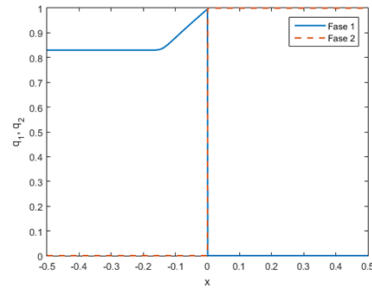
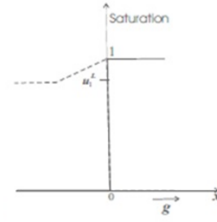
Resultados e discussões

Consideramos o caso de escoamento trifásico, 1D, vertical puramente gravitacional. Foram realizadas simulações com base em resultados previamente publicados à fim de validação qualitativa da implementação. Para o escoamento trifásico vertical (considerando efeitos gravitacionais) em meios porosos, as soluções numéricas obtidas neste trabalho podem ser consideradas as primeiras aproximações numéricas das soluções exatas semi-analíticas encontradas e descritas qualitativamente em Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013).

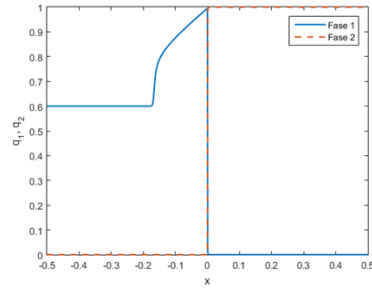
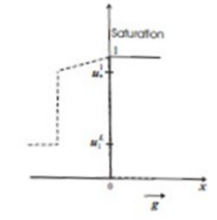
Escoamento trifásico com gravidade: Escoamentos que ocorrem em uma direção vertical (ou inclinada), os efeitos da gravidade são considerados a partir dos distintos valores de densidade entre as fases. Considera-se um problema onde inicialmente três fases diferentes, presentes no reservatório, estão separadas por uma membrana fixada na origem do meio em estudo. Os estados iniciais serão as saturações destas fases que estão à esquerda (acima) e à direita (abaixo) da membrana, Q_E e Q_D respectivamente. Ao iniciar a simulação a membrana é retirada e a acomodação das fases se inicia; o efeito gravitacional pode ser exemplificado como se acontecesse na vertical. Para obter as soluções de Riemann da figura 1 são usadas os seguintes parâmetros: $\alpha = 0$; $\mu_1 = \mu_2 = \mu_3 = 1$, a densidade varia caso a caso e a aceleração da gravidade atua no sentido positivo do eixo x . São simulados alguns dos exemplos presentes em Rodríguez-Bermúdez (2010) e Rodríguez-Bermúdez e Marchesin (2013), onde os autores obtiveram (descreveram) qualitativamente as soluções de Riemann utilizando métodos semi-analíticos e a teoria de leis de conservação para sistemas. As simulações numéricas obtidas no presente artigo estão de acordo com as soluções descritas nas referências acima citadas.

Nos casos das figuras 1(a) e 1(c), as densidades foram de $\rho_1 = \rho_2 = 1, \rho_3 = 0.8$. As fases 1 e 3 estão na região “superior” do meio, a tendência será da fase 3 (menos densa) ir ficando acima da fase 1, este fenômeno é representado por uma onda de rarefação no sentido de $x < 0$ no qual a saturação da fase 1 aumenta (de esquerda a direita). A formação da onda choque que precede a onda de rarefação na solução da figura 1(c) vem da maior saturação inicial da fase 3, o que resulta no surgimento dessa descontinuidade característica. Como a fase 2 só está presente na região “inferior” do meio, sua saturação não se altera já que não se desloca. O tempo de simulação é $t = 2.15$.

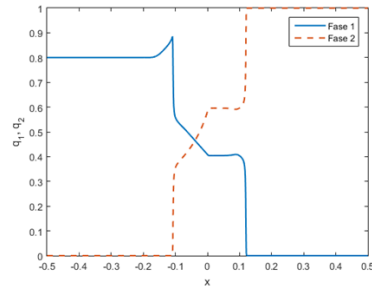
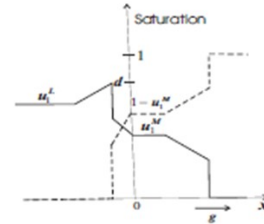
Nos casos das figuras 1(e) e 1(g), as densidades foram de $\rho_1 = 1, \rho_2 = \rho_3 = 0.8$. São utilizadas condições iniciais semelhantes às dos exemplos anteriores. Agora a fase 2 tem menor densidade e também se deslocará para as regiões “superiores”. A estrutura do grupo de ondas *slow-family* que viaja para acima é chamada de *double-characteristic-shock*, isto é, uma choque com rarefações adjacentes a ambos lados, e é uma estrutura atípica nas soluções de Riemann para os modelos estudados na literatura. O grupo de ondas *fast-family* que viaja pra abaixo é composta por uma onda de choque no qual essencialmente a fase 1 desloca-se para o fundo do reservatório. O tempo de simulação também é de $t = 2.15$.


 (a) $Q_E = (0.83, 0, 0.17)$ e $Q_D = (0, 1, 0)$


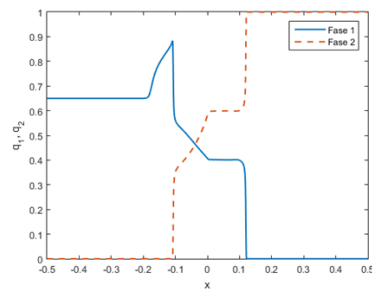
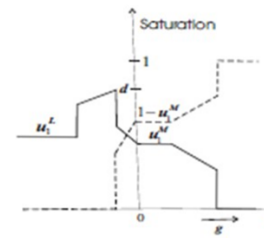
(b) Resultado semi-analítico esperado


 (c) $Q_E = (0.6, 0, 0.4)$ e $Q_D = (0, 1, 0)$


(d) Resultado semi-analítico esperado


 (e) $Q_E = (0.8, 0, 0.2)$ e $Q_D = (0, 1, 0)$


(f) Resultado semi-analítico esperado


 (g) $Q_E = (0.65, 0, 0.35)$ e $Q_D = (0, 1, 0)$


(h) Resultado semi-analítico esperado

Figura 1: Perfis de saturação para as fases 1 e 2 obtidas para problemas de Riemann em escoamentos verticais produzidos pela ação da gravidade, com a resposta semi-analítica correspondente de Rodríguez-Bermúdez (2010).

Conclusões

Foi implementado um método tipo Godunov em volumes finitos para sistemas 2×2 de leis de conservação e que pode ser estendidos para o caso de um maior número de equações. Foi validado o código para distintos modelos descrevendo fenômenos de transporte e logo utilizado para obter soluções

numéricas de um problema ainda em estudo como o caso de escoamento trifásico com gravidade em meios porosos.

Na implementação utilizou-se a construção de matrizes lineares nas interfaces dos volumes, em substituição da matriz jacobiana das funções de fluxo. Como resultado, o problema é transformando em diversos sistemas locais, resolvidos pelo *Riemann solver* de Roe (1981). A condição de entropia para correção das soluções é definido pelo critério de Harten e Hyman (1983). Do ponto de vista computacional, esta técnica reduz o custo ao obter as solução. Os resultados fornecidos pelo código e descritos na literatura se correlacionam adequadamente.

O tempo computacional de execução do código está relacionado ao número de iterações, à complexidade das funções de fluxo e o valor v usado. Os perfis das soluções de Riemann, apresentam uma formação bem definida dos grupos de onda (estruturas de choques e rarefações) e ausência de oscilações espúrias, mostrando a correta escolha dos parâmetros numéricos. Além disso, a utilização do *Riemann solver* e da configuração da matriz $\tilde{A}$ não influenciaram a solução e a correção de entropia se faz necessária e importante para o ajuste.

Referências

- AZEVEDO, A. V. et al. The solution by the wave curve method of three-phase flow in virgin reservoirs. **Transport in Porous Media**, v. 83, n. 01, p. 99–125, 2010.
- COREY, A. et al. Three-Phase Relative Permeability. **Journal of Petroleum Technology**, v. 8, n. 11, p. 63–65, 1956.
- COURANT, R.; FRIEDRICHS, K.; LEWY, H. On the partial difference equations of mathematical physics. **IBM journal of Research and Development**, v. 11, n. 02, p. 215–234, 1967.
- GODUNOV, S. K.; BOHACHEVSKY, I. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. **Matematicheskij sbornik**, v. 47, n. 03, p. 271–306, 1959.
- HARTEN, A; HYMAN, J. M. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. **Journal of computational Physics**, v. 50, n. 02, p. 235–269, 1983.
- LAX, P. D. **Hyperbolic systems of conservation laws and the mathematical theory of shock waves**. 1. ed. SIAM, 1973.
- LEVEQUE, R. J. **Finite-Volume Methods for Hyperbolic Problems**. 1. ed. New York: Cambridge University Press, 2004.
- LIU, T. -P. The entropy condition and the admissibility of shocks. **Journal of Mathematical Analysis and Applications**, v. 53, n. 01, p. 78–88, 1976.
- RODRÍGUEZ-BERMÚDEZ, P. **Buoyancy driven three-phase flow in porous media**. 2010. 112 f. Tese (Doutorado em Matemática) – IMPA, Rio de Janeiro, 2010.
- RODRÍGUEZ-BERMÚDEZ, P.; MARCHESIN, D. Riemann Solutions for vertical flow of three phases in porous media: simple cases. **Journal of Hyperbolic Differential Equations**, v. 10, n. 02, p. 335–370, 2013.
- ROE, P. L. Approximate riemann solvers, parameter vector, and difference schemes. **Journal of Computational Physics**, v. 43, n. 02, p. 357–372, 1981.
- TORO, E. F. **Riemann solvers and numerical methods for fluid dynamics: a practical introduction**. 3. ed. Berlin: Springer Science and Business Media, 2009.

NOVOS ESQUEMAS DE DIFERENÇAS FINITAS PARA A EQUAÇÃO DE HELMHOLTZ

Gustavo Benitez Alvarez
Universidade Federal Fluminense
benitez.gustavo@gmail.com

Helder da Fonseca Nunes
Universidade Federal Fluminense
helderfn@id.uff.br

Resumo: A equação de Helmholtz descreve os harmônicos-temporais de ondas acústicas. É bem conhecido que métodos de diferenças finitas e elementos finitos apresentam o efeito de poluição do erro para número de onda médio e alto. Nesse trabalho são analisados dois novos esquemas consistentes de diferenças finitas centradas de segunda ordem de precisão em uma e duas dimensões. A análise de dispersão, o comportamento do erro e os resultados numéricos mostram o bom desempenho do Novo Esquema 2.

Palavras-chave: Equação de Helmholtz. Diferenças Finitas. Análise de dispersão. Poluição do erro.

Introdução

A equação de Helmholtz escalar com condições de contorno de Dirichlet é dada por:

$$\Delta u + k^2 u = f \text{ em } \Omega, \quad (1)$$

$$u = g \text{ em } \partial\Omega, \quad (2)$$

onde k é o número de onda, Ω é o interior do domínio limitado com contorno Lipschitz $\partial\Omega$, f é um termo fonte, e u é um campo escalar que descreve os harmônicos-temporais de ondas acústicas. Pela sua grande importância, tal equação é objeto de muitos estudos. Contudo, grande parte das aplicações práticas exigem que se tenha um valor de número de onda elevado e sabe-se que a qualidade da solução numérica depende significativamente desse parâmetro, sendo sua precisão deteriorada com o aumento do valor de k (LOULA et al., 2007).

Para que uma aceitável aproximação possa ser obtida, a resolução da malha utilizada deve ser ajustada de acordo com o número de onda. Esse ajuste é obtido através da regra do dedão (*rule of thumb*), que prescreve uma quantidade mínima de elementos por comprimento de onda. (IHLENBURG; BABUŠKA, 1995b). Contudo, é de conhecimento que, mesmo quando a regra do dedão é obedecida, a performance dos métodos de diferenças finitas e elementos finitos diminui drasticamente com o aumento do número de onda (SINGER; TURKEL, 1998; IHLENBURG; BABUŠKA, 1995a). Esse problema é conhecido por efeito de poluição do erro, e está relacionado à diferença entre o número de onda k da solução exata e o número de onda k^h da solução numérica. Para problemas unidimensionais, já existem métodos de elementos finitos que eliminam esse efeito, tal como o método GLS (Galerkin-Mínimos Quadrados) apresentado por Harari e Hughes (1991). Para problemas bidimensionais, sabe-se que é impossível eliminar totalmente esse efeito de poluição, como pode ser visto em Babuška e Sauter (1997).

Buscaram-se desenvolver, portanto, métodos que minimizassem essa poluição do erro, como o Método de Elementos Finitos Quase Estabilizado (QSFEM) descrito por Babuška, Ihlenburg et al. (1995). Alguns trabalhos tem sido desenvolvidos também com a utilização de diferenças finitas, de modo a obter uma melhor qualidade nas soluções. Podemos citar Sutmann (2007) que desenvolveu um novo esquema de sexta ordem, Singer e Turkel (1998) que desenvolveram esquemas de quarta ordem baseados nas generalizações das aproximações de Padé. Nabavi, Siddiqui e Dargahi (2007) desenvolveram um novo

esquema compacto de sexta ordem utilizando nove pontos. Wu (2017) desenvolveu um esquema de quarta ordem que minimiza o erro de dispersão, e Wu e Xu (2018) apresentam um esquema de sexta ordem.

No presente trabalho são apresentados dois novos esquemas de diferenças finitas centrados de segunda ordem de precisão para a equação de Helmholtz em uma e duas dimensões. Esses esquemas são consistentes, ou seja, conforme há o refinamento da malha a solução numérica tende para solução exata. Os novos esquemas foram obtidos realizando novas aproximações do termo k^2u da equação 1.

O trabalho apresenta a seguinte organização: inicialmente são apresentados os esquemas de diferenças finitas e a análise de dispersão para os casos 1D e 2D. Em sequência tem-se um breve estudo do erro seguido dos resultados numéricos e discussões. Por fim, tem-se as conclusões.

Esquemas de Diferenças Finitas em Uma e Duas Dimensões

Três esquemas de diferenças finitas são apresentados: Centrado Clássico (CC), Novo Esquema 1 (NE-1) e Novo Esquema 2 (NE-2). Todos são esquemas centrados de segunda ordem de precisão, onde a aproximação do termo k^2u na equação de Helmholtz pode ser entendida como uma média da solução em torno do ponto central. A análise realizada aqui apenas considera malhas uniformes, onde $h = \Delta x = \Delta y$ é o espaçamento da malha.

Em Uma Dimensão (1D)

Para todos os esquemas a derivada segunda foi aproximada com diferenças finitas centradas de segunda ordem de precisão:

$$\frac{d^2u}{dx^2} \approx \frac{U_{i-1} - 2U_i + U_{i+1}}{h^2}, \quad (3)$$

onde, $U_{i-1} = U(x_i - h)$, $U_i = U(x_i)$, $U_{i+1} = U(x_i + h)$.

Centrado Clássico (CC)

Neste esquema o segundo termo da equação de Helmholtz é aproximado como $k^2u(x) \approx k^2U_i$. Isto resulta na equação matricial

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{h^2} + k^2U_i = f_i \text{ ou } \frac{1}{h^2} (AU_{i-1} + BU_i + AU_{i+1}) = f_i, \quad (4)$$

onde $A = 1$ e $B = -2 + (kh)^2$.

Novo Esquema 1 (NE-1)

Neste esquema utilizou-se a aproximação $k^2u(x) \approx k^2 \frac{U_{i-1} + U_{i+1}}{2}$, resultando na equação matricial

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{h^2} + k^2 \left(\frac{U_{i-1} + U_{i+1}}{2} \right) = f_i \text{ ou } \frac{1}{h^2} (AU_{i-1} + BU_i + AU_{i+1}) = f_i, \quad (5)$$

onde $A = 1 + \frac{(kh)^2}{2}$ e $B = -2$.

Novo Esquema 2 (NE-2)

Este esquema consiste numa média dos esquemas CC e NE-1 resultando na equação do estêncil

$$\frac{U_{i-1} - 2U_i + U_{i+1}}{h^2} + k^2 \left(\frac{U_i + \frac{U_{i-1} + U_{i+1}}{2}}{2} \right) = f_i \text{ ou } \frac{1}{h^2} (AU_{i-1} + BU_i + AU_{i+1}) = f_i, \quad (6)$$

onde $A = 1 + \frac{(kh)^2}{4}$ e $B = -2 + \frac{(kh)^2}{2}$.

Em Duas Dimensões (2D)

Analogamente, para todos os esquemas as derivadas segundas foram aproximadas com diferenças finitas centradas de segunda ordem de precisão:

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} \text{ e } \frac{\partial^2 u}{\partial y^2} \approx \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2}. \quad (7)$$

A equação do estêncil clássico de 5 pontos para cada esquema resulta numa equação matricial do tipo:

$$\frac{1}{h^2} (AU_{i-1,j} + AU_{i+1,j} + BU_{i,j} + AU_{i,j-1} + AU_{i,j+1}) = f_{i,j}. \quad (8)$$

Centrado Clássico (CC)

Temos a equação do estêncil:

$$\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2} + k^2 U_{i,j} = f_{i,j}, \quad (9)$$

que pode ser escrita na forma da equação 8 com $A = 1$ e $B = -4 + (kh)^2$.

Novo Esquema 1 (NE-1)

$$\begin{aligned} \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2} + \\ + k^2 \left[\left(\frac{U_{i-1,j} + U_{i+1,j}}{4} \right) + \left(\frac{U_{i,j-1} + U_{i,j+1}}{4} \right) \right] = f_{i,j}, \end{aligned} \quad (10)$$

ou na forma da equação 8 com $A = 1 + \frac{(kh)^2}{4}$ e $B = -4$.

Novo Esquema 2 (NE-2)

$$\begin{aligned} \frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{h^2} + \frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{h^2} + \\ + k^2 \left[\frac{\left(\frac{2U_{i,j} + U_{i-1,j} + U_{i+1,j}}{4} \right) + \left(\frac{2U_{i,j} + U_{i,j-1} + U_{i,j+1}}{4} \right)}{2} \right] = f_{i,j}, \end{aligned} \quad (11)$$

ou na forma da equação 8 com $A = 1 + \frac{(kh)^2}{8}$ e $B = -4 + \frac{(kh)^2}{2}$.

Análise de Dispersão

A diferença entre o número de onda da solução exata (k) e da solução numérica (k^h) pode ser estimada através da análise de dispersão. Esta diferença provoca a perda de fase da solução numérica e impacta o efeito de poluição do erro. A análise consiste em considerar a equação de Helmholtz homogênea ($f = 0$) e supor soluções numéricas na forma de ondas planas que, substituídas na equação do estêncil para cada esquema, permite obter uma expressão para k^h (número de onda discreto) em função de k (FERNANDES, 2009).

Caso Unidimensional

Neste caso, a solução numérica na forma $U_i = e^{ik^h x_i}$ é substituída na equação do estêncil 4 obtendo:

$$k^h = \frac{1}{h} \arccos\left(-\frac{B}{A}\right). \quad (12)$$

Substituindo os valores de A e B de cada esquema em 12 e realizando a expansão em série de Taylor em torno de $kh = 0$, temos as expressões de $k^h - k$ para cada método:

$$\text{CC: } k^h - k = \frac{1}{24}(k^3 h^2) + \frac{3}{640}(k^5 h^4) + \mathcal{O}(k^6 h^5), \quad (13)$$

$$\text{NE-1: } k^h - k = -\frac{5}{24}(k^3 h^2) + \frac{43}{640}(k^5 h^4) + \mathcal{O}(k^6 h^5), \quad (14)$$

$$\text{NE-2: } k^h - k = -\frac{1}{12}(k^3 h^2) + \frac{1}{80}(k^5 h^4) + \mathcal{O}(k^6 h^5). \quad (15)$$

Caso Bidimensional

Neste caso, a onda plana discreta na direção θ é $U_{i,j} = e^{ik^h(x_i \cos \theta + y_j \sin \theta)}$. Substituindo na equação homogênea 8 obtemos:

$$B + 2A(\cos(k^h h \cos \theta) + \cos(k^h h \sin \theta)) = 0. \quad (16)$$

Substituindo os valores de A e B para cada esquema e realizando a expansão em série de Taylor em torno de $kh = 0$, após algumas manipulações algébricas com o auxílio do *software* Mathematica¹ obtemos:

$$\text{CC: } k^h - k = \frac{3 + \cos 4\theta}{96}(k^3 h^2) + \frac{235 + 162 \cos 4\theta + 35 \cos^2 4\theta}{92160}(k^5 h^4) + \mathcal{O}(k^6 h^5), \quad (17)$$

$$\text{NE-1: } k^h - k = \frac{-9 + \cos 4\theta}{96}(k^3 h^2) + \frac{1315 - 198 \cos 4\theta + 35 \cos^2 4\theta}{92160}(k^5 h^4) + \mathcal{O}(k^6 h^5), \quad (18)$$

$$\text{NE-2: } k^h - k = \frac{-3 + \cos 4\theta}{96}(k^3 h^2) + \frac{235 + 162 \cos 4\theta + 35 \cos^2 4\theta}{92160}(k^5 h^4) + \mathcal{O}(k^6 h^5). \quad (19)$$

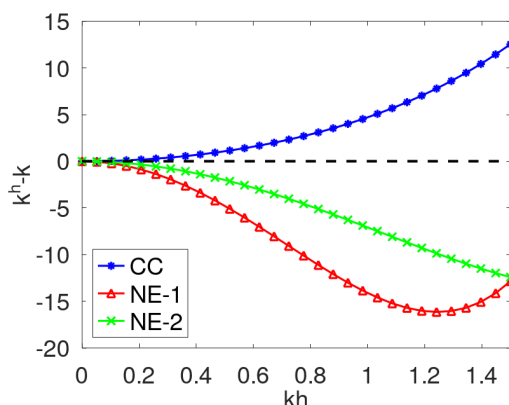


Figura 1: Dispersão 1D para $k = 100$.

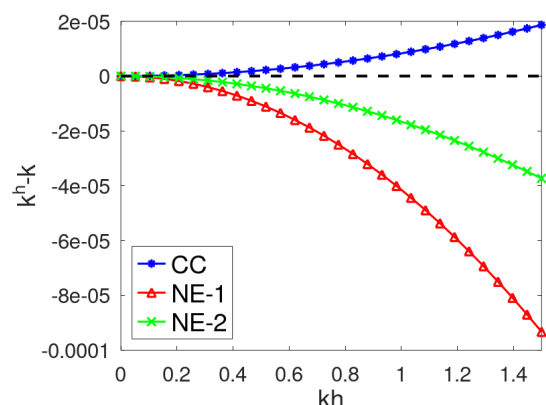
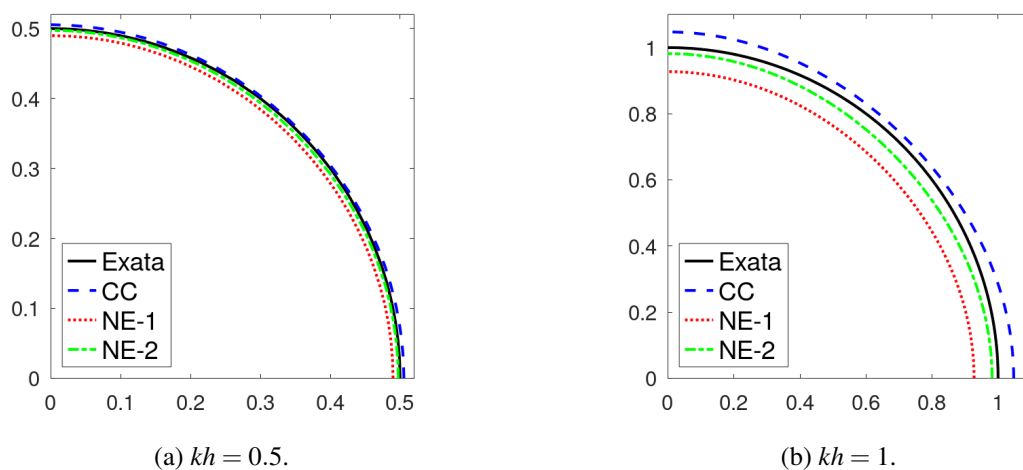


Figura 2: Dispersão 2D para $k = 100$ e $\theta = \pi/4$.

¹<https://www.wolfram.com/mathematica>

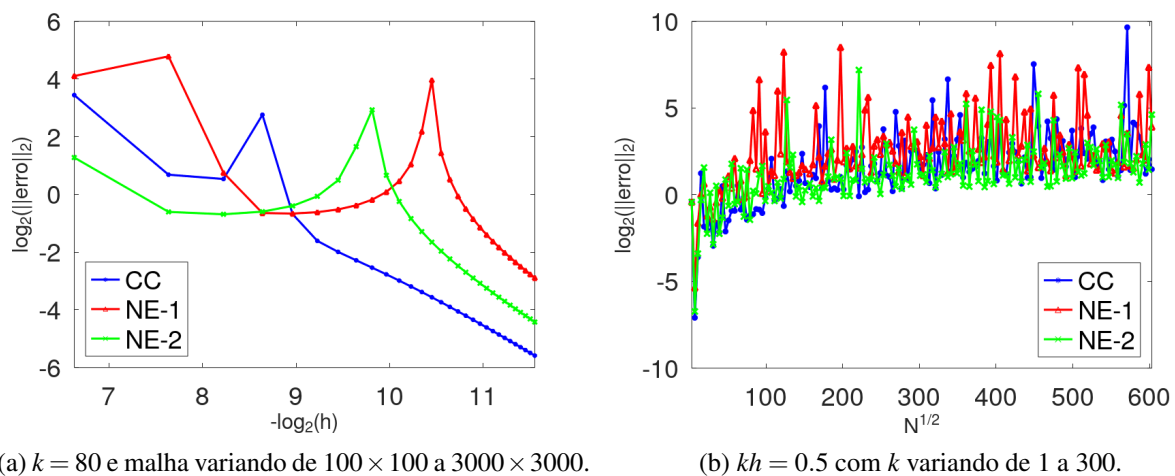

 Figura 3: Relação de dispersão no caso 2D para $0 \leq \theta \leq \pi/2$.

As figuras 1 e 2 mostram que para ambos os casos, 1D e 2D, como esperado, o retardo de fase diminui para malhas mais refinadas. Os esquemas CC e NE-2 apresentam melhores resultados quando comparado ao NE-1, embora todos apresentem dispersão da mesma ordem.

Quando comparamos a relação da dispersão no caso 2D com a direção da onda plana θ nas figuras 3a e 3b, podemos observar que o NE-2 apresentou a menor dispersão em $\theta = 0$ e $\theta = \pi/2$, e a maior dispersão em $\theta = \pi/4$. O método CC apresentou um comportamento contrário ao NE-2, exibindo maior dispersão em $\theta = 0$ e $\theta = \pi/2$, e a menor dispersão em $\theta = \pi/4$. Por outro lado, a dispersão do NE-1 não mostra grandes variações com o ângulo, porém apresenta maior distância da dispersão exata.

Comportamento do Erro

Podemos mostrar o efeito de poluição do erro aumentando o refinamento da malha para um dado número de onda fixo. É esperado que, se não houvesse o efeito de poluição, o erro diminuiria conforme o refinamento da malha. O erro absoluto apresentado aqui $e_{i,j} = U_{i,j} - u_{i,j}$ foi calculado utilizando a equação A.18 para $q = 2$, que encontra-se em LeVeque (2007, p. 252).


 Figura 4: Erro no caso 2D variando-se o tamanho da malha (4a) e mantendo-se kh constante (4b).

Na figura 4a é possível observar que aumentando-se a quantidade de elementos da malha o erro

aumenta em certos pontos (picos) e só após isso temos uma convergência assintótica para a ordem de convergência de cada método. Manter kh constante, como já mencionado, também não é o suficiente para controlar o efeito da poluição do erro, como pode ser visto na figura 4b.

Resultados Numéricos

Códigos computacionais 1D e 2D foram desenvolvidos usando o *software* GNU Octave². São apresentados resultados numéricos da equação homogênea e não-homogênea para um domínio quadrado unitário e malha uniforme composta por 200×200 elementos. As figuras 5a e 5b apresentam as soluções dos três métodos e o interpolante para a equação homogênea com $k = 100$, $kh = 0.5$ e superposição de três ondas planas nas direções $\theta = 0, \pi/4$ e $\pi/8$. Neste caso a solução exata $u(x, y) = \sum_{i=1}^n (\cos(k(x \cos \theta_i + y \sin \theta_i)))$ é a superposição de $n = 3$ ondas planas. Na figuras 6a e 6b considera-se a fonte $f(x, y) = 4 + k^2(x^2 + y^2)$ cuja solução exata é $u(x, y) = x^2 + y^2 + \sin(k(x \cos \theta + y \sin \theta))$. Em ambos casos, homogêneo e não-homogêneo, percebe-se que o método NE-2 exibe resultados mais próximos do interpolante, apresentando menor retardo de fase e diferença de amplitude que os métodos CC e NE-1.

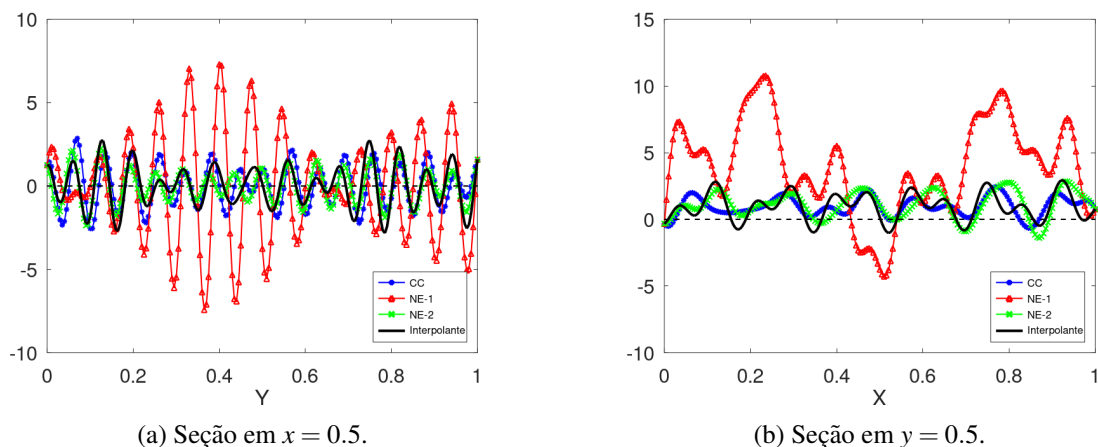


Figura 5: Soluções numéricas no caso 2D homogêneo com $k = 100$, $kh = 0.5$ e $\theta = 0, \pi/4$ e $\pi/8$.

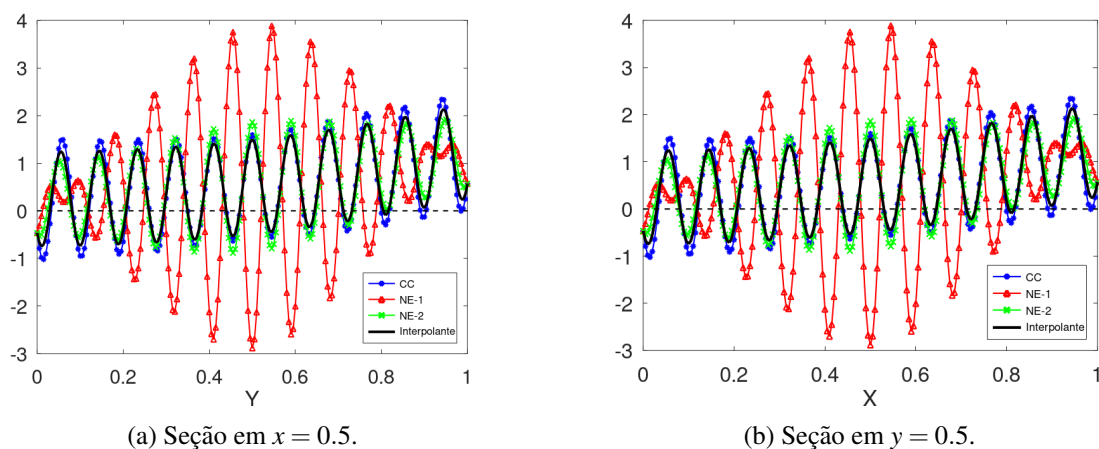


Figura 6: Soluções numéricas no caso 2D não-homogêneo com $k = 100$, $kh = 0.5$ e $\theta = \pi/4$.

²<https://www.octave.org>

Conclusões

Dois novos esquemas de diferenças finitas para a equação de Helmholtz foram apresentados. Estes foram obtidos modificando apenas a aproximação para o segundo termo k^2u , e são de segunda ordem de precisão. As análises de dispersão para os casos 1D e 2D mostraram o bom desempenho do método NE-2. Os gráficos dos erros confirmaram a presença do efeito de poluição do erro para os três métodos, conforme esperado. Por fim, os resultados numéricos corroboram o que foi visto nas análises de dispersão, com o novo método NE-2 apresentando resultados mais próximos aos do interpolante.

Referências

- BABUŠKA, I.; IHLENBURG, F. et al. A Generalized Finite Element Method for solving the Helmholtz equation in two dimensions with minimal pollution. en. **Comput. Methods Appl. Mech. Engrg.**, v. 128, p. 325–359, 1995.
- BABUŠKA, I. M.; SAUTER, S. A. Is the Pollution Effect of the FEM Avoidable for the Helmholtz Equation Considering High Wave Numbers? en. **SIAM Journal on Numerical Analysis**, v. 34, n. 6, p. 2392–2423, dez. 1997.
- FERNANDES, D. T. **Métodos de Diferenças Finitas e Elementos Finitos para o Problema de Helmholtz**. Tese (Doutorado em Modelagem Computacional) – Laboratório Nacional de Computação Científica, Petrópolis, RJ, 2009.
- IHLENBURG, F.; BABUŠKA, I. Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation. en. **International Journal for Numerical Methods in Engineering**, v. 38, n. 22, p. 3745–3774, nov. 1995.
- IHLENBURG, F.; BABUŠKA, I. Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM. en. **Computers & Mathematics with Applications**, v. 30, n. 9, p. 9–37, nov. 1995.
- LEVEQUE, R. J. **Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems**. Seattle, Washington: Society for Industrial e Applied Mathematics, jan. 2007.
- LOULA, A. F. D. et al. A discontinuous finite element method at element level for Helmholtz equation. en. **Computer Methods in Applied Mechanics and Engineering**, v. 196, n. 4-6, p. 867–878, jan. 2007.
- NABAVI, M.; SIDDIQUI, M. K.; DARGAHI, J. A new 9-point sixth-order accurate compact finite-difference method for the Helmholtz equation. en. **Journal of Sound and Vibration**, v. 307, n. 3-5, p. 972–982, nov. 2007.
- SINGER, I.; TURKEL, E. High-order finite difference methods for the Helmholtz equation. en. **Computer Methods in Applied Mechanics and Engineering**, v. 163, n. 1-4, p. 343–358, set. 1998.
- SUTMANN, G. Compact finite difference schemes of sixth order for the Helmholtz equation. **Journal of Computational and Applied Mathematics**, v. 203, n. 1, p. 15–31, jun. 2007.
- WU, T. A dispersion minimizing compact finite difference scheme for the 2D Helmholtz equation. en. **Journal of Computational and Applied Mathematics**, v. 311, p. 497–512, fev. 2017.
- WU, T.; XU, R. An optimal compact sixth-order finite difference scheme for the Helmholtz equation. en. **Computers & Mathematics with Applications**, v. 75, n. 7, p. 2520–2537, abr. 2018.

PROBLEMAS DIFUSIVOS-CONVECTIVOS ABORDADOS PELO MÉTODO DAS DIFERENÇAS FINITAS CENTRADAS E *UPWIND*

Victor de Souza Coelho Freitas
Universidade Federal Fluminense
victorcoelho@id.uff.br

Gustavo Benitez Alvarez
Universidade Federal Fluminense
gustavobenitez@id.uff.br

Cleyton Senior Stampa
Universidade Federal Fluminense
cleytons@id.uff.br

Panthers Rodríguez Bermúdez
Universidade Federal Fluminense
panthersrb@id.uff.br

Resumo: O presente trabalho tem como objetivo abordar numericamente problemas governados pela equação de transporte difusiva-convectiva em regime permanente. Mais especificamente, foram estudados cenários bidimensionais nos quais há o predomínio do fenômeno da convecção, uma vez que desta forma tem-se um problema de perturbação singular. Para a abordagem numérica foi escolhido o método de diferenças finitas centradas e o método de diferenças finitas *upwind*. Foram implementados problemas contendo ou não camadas-limite em suas soluções. Ambos os métodos foram avaliados, com seus resultados comparados para determinar suas vantagens e desvantagens no contexto deste estudo.

Palavras-chave: Diferenças finitas centradas. Diferenças finitas *upwind*. Problema difusivo-convectivo. Oscilações espúrias.

Introdução

Problemas difusivos-convectivos dizem respeito à cenários nos quais tem-se um fenômeno físico no qual partículas, energia, ou outra quantidade física é transportada dentro de um domínio por meio de dois processos acoplados: difusão e convecção. A difusão é o mecanismo de transporte no qual a quantidade do elemento físico de interesse se desloca de um meio mais concentrado para o meio menos concentrado. Já a convecção se dá pela translação desse elemento através do domínio no qual ele está presente.

A Equação 1 abaixo representa a forma geral da equação difusiva-convectiva, na qual U é a quantidade física de interesse. Sendo $\vec{w}$ um campo vetorial de velocidade, k denota o coeficiente difusivo e f representa uma fonte ou sumidouro dentro do domínio estabelecido. Para esta equação, temos que a solução é uma função $U(x, y, t)$ que varia de acordo com o tempo, dentro de um domínio $\Omega \subset \mathbb{R}^2$. Com Ω sendo aberto e delimitado, cujo limite Γ é suave e definido.

$$\frac{\partial U}{\partial t} + \vec{w} \cdot \nabla U = \nabla \cdot (v \nabla) U + f \quad (1)$$

Considerando que U não varia com o tempo, ou seja, se encontra em regime permanente e depende apenas de suas variáveis espaciais, tem-se que $\frac{\partial U}{\partial t} = 0$. Assim, pode-se rescrever a Equação 1 conforme demonstrado pela Equação 2 abaixo. Se caracterizando como uma equação diferencial parcial (EDP) elíptica.

$$-k\Delta U + \vec{w} \cdot \nabla U = f \quad (2)$$

Sabendo que o número de Péclet (P_e) é dado como a razão entre o coeficiente convectivo e o coeficiente difusivo, pode-se definir ε como o inverso de P_e , com $\vec{w} = (a_1, a_2)$, tal que $\varepsilon_n = \frac{k}{a_n}$, para $n = 1, 2$. É esperado que, para casos de convecção dominante, ou seja, $\varepsilon_n \rightarrow 0$, a Equação 2 tende a se reduzir a uma equação diferencial parcial de primeira ordem, que só permite uma condição de contorno.

Entretanto, como $\varepsilon_n > 0$, ainda que muito pequeno, tem-se uma equação de segunda ordem que necessita de duas condições de contorno. Logo, é esperado um panorama de solução anormal próximo às condições de contorno, já que para $\varepsilon_n \rightarrow 0$ o problema está super-especificado.

Conforme discutido por Carmo e Benitez (2003), assim como Volker e Knobloch (2007), este panorama anormal é o que se chamam as oscilações espúrias, uma vez que a solução da equação difusiva-convectiva possui camadas-limite, que são pequenas regiões do domínio que possuem alto gradiente. Haja vista que o tamanho das malhas normalmente utilizadas em soluções numéricas é significativamente maior que a espessura da camada-limite, essa região não é resolvida de forma adequada. Dando assim, origem à tais oscilações que não são físicas, ou seja, não representam a solução do problema posto, mas sim instabilidades no método numérico.

Como discutido por Zhao (2018), problemas convectivos-difusivos, nos quais existe dominância da convecção, apresentam para os métodos clássicos de resolução, como o método das diferenças finitas (MDF) e elementos finitos, instabilidades em seu panorama de solução. Como mostrado por Frerichs e John (2021), pode-se contornar esse problema com artifícios para melhorar a estabilidade, como o método *Upwind* para a metodologia de diferenças finitas. Neste caso, é adicionado uma difusividade numérica na solução, proveniente da formulação matemática do método. Entretanto, esse procedimento tende a prejudicar a precisão dos resultados obtidos, com uma difusão excessiva em alguns casos. Atualmente, não se tem uma formulação que seja ao mesmo tempo estável e precisa, para soluções com a presença de camadas-limite.

Método das diferenças finitas centradas e *upwind*

De acordo com Tanehill, Anderson e Pletcher (2012), a abordagem do método das diferenças finitas se caracteriza como a discretização de um problema inicialmente contínuo, de forma que as variáveis de interesse existam unicamente em pontos discretos. Por essa metodologia, as derivadas presentes são substituídas por diferenças entre tais pontos, logo tem-se uma representação algébrica de uma EDP, na qual tem-se ao final do processo um sistema linear que pode ser facilmente resolvido.

Partindo da Equação 3 abaixo, como mostrado por Maliska (2017) e Pantakar (2018), pode-se substituir as derivadas de segunda ordem (termo difusivo) pela formulação centrada de 2ª ordem. Já as derivadas de primeira ordem (termo convectivo) são substituídas pela formulação centradas de 2ª ordem, obtendo-se assim a discretização pelo método das diferenças centradas, considerando um campo de velocidades $\vec{w} = (a_1, a_2)$, tem-se a Equação 4.

$$-k \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) + \vec{w} \cdot \left(\frac{\partial U}{\partial x} \hat{i} + \frac{\partial U}{\partial y} \hat{j} \right) = f \quad (3)$$

$$\begin{aligned} & -k \left[\left(\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left(\frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \\ & \dots + \left[\left(a_1 \frac{U_{i+1,j} - U_{i-1,j}}{2\Delta x} \right) + \left(a_2 \frac{U_{i,j+1} - U_{i,j-1}}{2\Delta y} \right) \right] = f \end{aligned} \quad (4)$$

Já para o MDF *upwind*, como mostrado por Thomas (2013), a discretização do termo convectivo irá depender da direção do campo de velocidade. Para um valor de velocidade positivo, utiliza-se a

formulação atrasada de 1ª ordem; já para um valor de velocidade negativo, utiliza-se a formulação avançada de 1ª ordem. De maneira que a formulação *upwind* para velocidades positivas e negativas estão representadas, respectivamente, pelas Equações 5 e 6 abaixo.

$$\begin{aligned}
 & -k \left[\left(\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left(\frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \\
 & \dots + \left[\left(a_1 \frac{U_{i,j} - U_{i-1,j}}{\Delta x} \right) + \left(a_2 \frac{U_{i,j} - U_{i,j-1}}{\Delta y} \right) \right] = f
 \end{aligned} \tag{5}$$

$$\begin{aligned}
 & -k \left[\left(\frac{U_{i-1,j} - 2U_{i,j} + U_{i+1,j}}{\Delta x^2} \right) + \left(\frac{U_{i,j-1} - 2U_{i,j} + U_{i,j+1}}{\Delta y^2} \right) \right] + \dots \\
 & \dots + \left[\left(a_1 \frac{U_{i+1,j} - U_{i,j}}{\Delta x} \right) + \left(a_2 \frac{U_{i,j+1} - U_{i,j}}{\Delta y} \right) \right] = f
 \end{aligned} \tag{6}$$

Para ambos métodos, após realizada a discretização apropriada, concatena-se os termos de mesmo índice de maneira que se tenha uma equação matricial $MU = f$, conforme mostrada pela Equação 7 abaixo, sendo esta a equação a partir da qual foram construídos os algoritmos computacionais.

$$AU_{i,j-1} + BU_{i-1,j} + CU_{i,j} + DU_{i+1,j} + EU_{i,j+1} = f \tag{7}$$

Na qual M é a matriz dos coeficientes resultantes da formulação de diferenças finitas utilizada, sendo uma matriz pentadiagonal esparsa de tamanho $n^2 \times n^2$, correspondendo à uma malha com n nós em cada eixo.

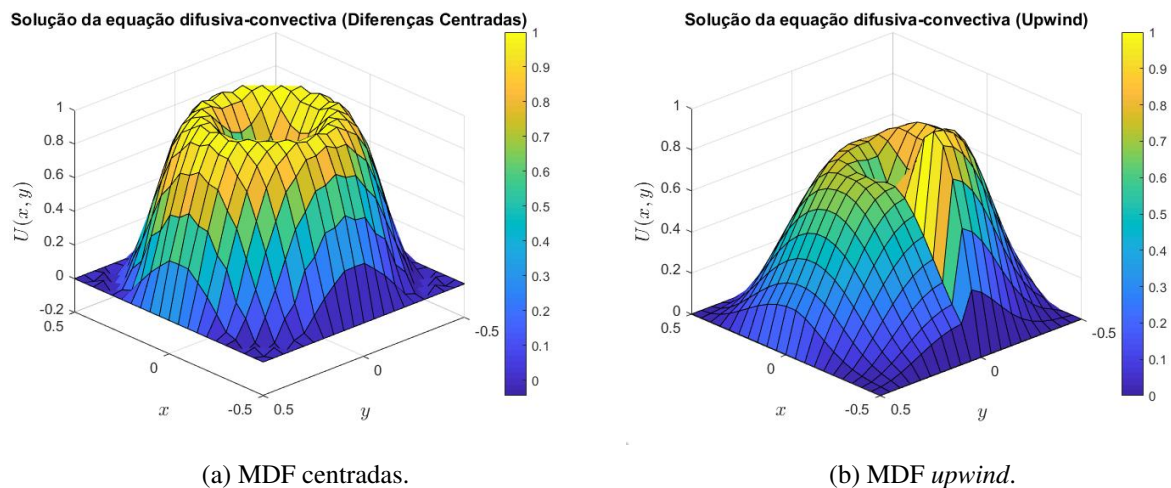
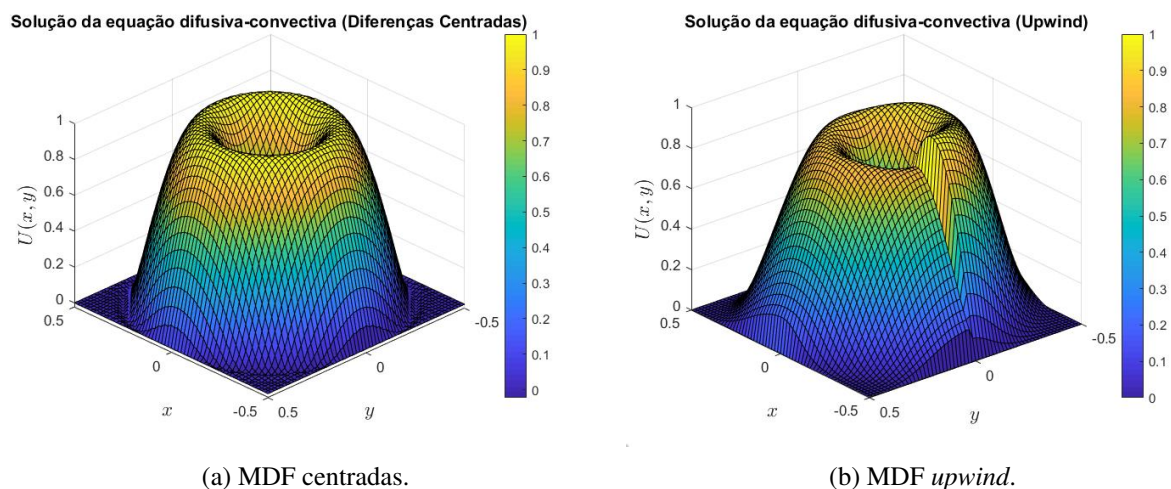
Metodologia

Para o contexto da equação difusiva-convectiva aplicada a um domínio $\Omega \subset \mathbb{R}^2$, não se obteve, para este trabalho, soluções exatas para que se possa validar os códigos criados. Nesse caso optou-se por utilizar casos clássicos da literatura. Assim, os panoramas das soluções são conhecidos de forma que é possível avaliar os resultados obtidos por meio dos códigos criados, para o método das diferenças finitas centradas e *upwind*.

Os casos de estudo implementados tem a finalidade de promover situações de convecção dominante, de forma que se possa constatar a presença, ou não, de oscilações espúrias (instabilidades). Os problemas escolhidos se dividem em dois grupos: os que possuem camada-limite em sua solução e os que possuem solução suave. Com os resultados obtidos, será possível constatar qual método forneceu os melhores resultados para cada um dos cenários propostos.

Resultados numéricos

Para o primeiro problema estudado foi escolhido um cenário no qual não há camadas-limite na solução, de forma que para este foi definido um domínio $(-0.5 < x < 0.5, -0.5 < y < 0.5)$, o campo de velocidades é definido como $\vec{w} = (-y, x)$, a condição de contorno é zero ao redor de todo o domínio, já o coeficiente difusivo é dado por $k = 10^{-10}$. Para este estudo foi também imposta a condição interna dada pela função seno $f(0, y) = \sin(2\pi y) \forall y \in [-0.5, 0]$. Por fim, foram utilizadas as malhas quadradas de tamanhos $n_x = n_y = 20$ e $n_x = n_y = 60$.

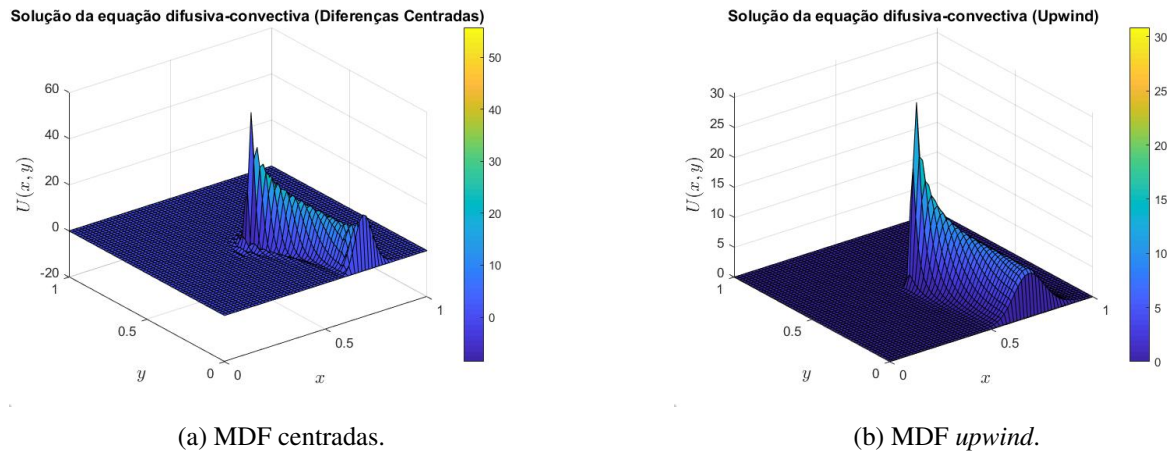
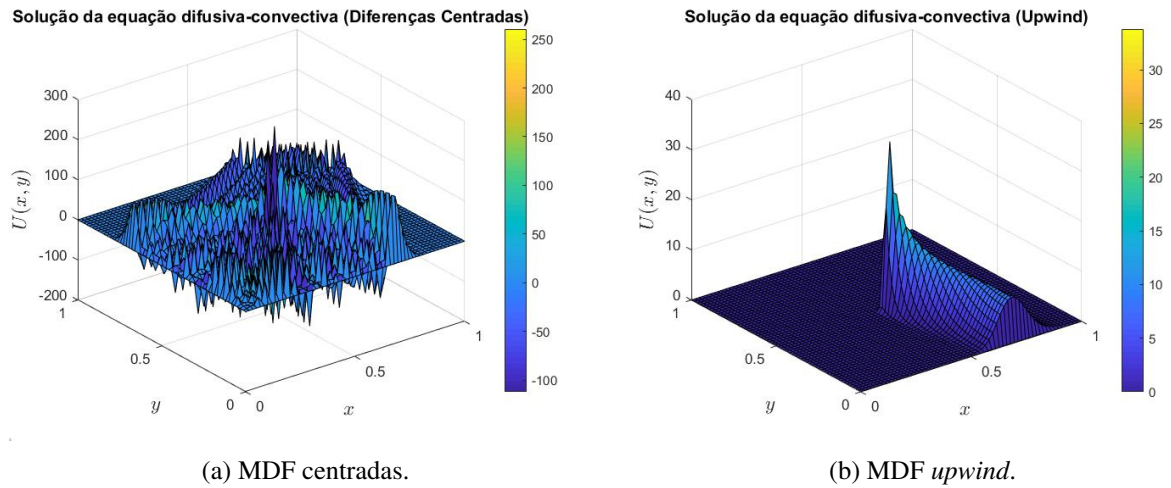

 Figura 1: Solução do primeiro problema para $nx = ny = 20$.

 Figura 2: Solução do primeiro problema para $nx = ny = 60$.

Para este problema pode-se perceber que as soluções apresentadas pelo MDF *upwind* apresentam grande difusão numérica introduzida pelo método, constatada pelo degrau próximo à condição interna imposta. Analisando as Figuras 1 e 2, é visto que conforme a malha é refinada, a diferença entre os valores da solução próximo aos pontos com $y \in [-0.5, 0]$ é reduzida, mas não é eliminada.

Em contraste, as soluções obtidas com o MDF centradas apresentaram um panorama coerente com o esperado, sem a presença de tal degrau. Isso se deve à maior precisão da formulação centrada, e também à tendência do MDF *upwind* de suavizar as soluções, que neste caso resultou em uma diferença significativa em relação ao que era esperado.

Para o segundo problema, as condições de contorno foram consideradas zero, já o termo fonte foi definido como a função Gaussiana representada pela Equação 8 abaixo. O campo de velocidades considerado será variável dentro do domínio, de maneira que $\vec{w} = (y, -x)$. Já o coeficiente difusivo será $k = 10^{-3}$ e $k = 10^{-6}$, com uma malha $nx = ny = 60$.

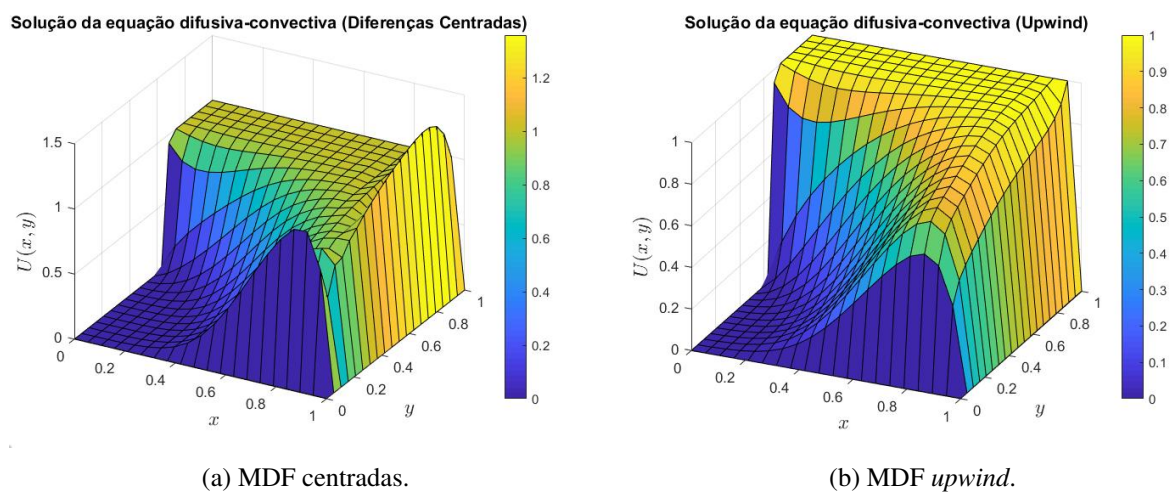
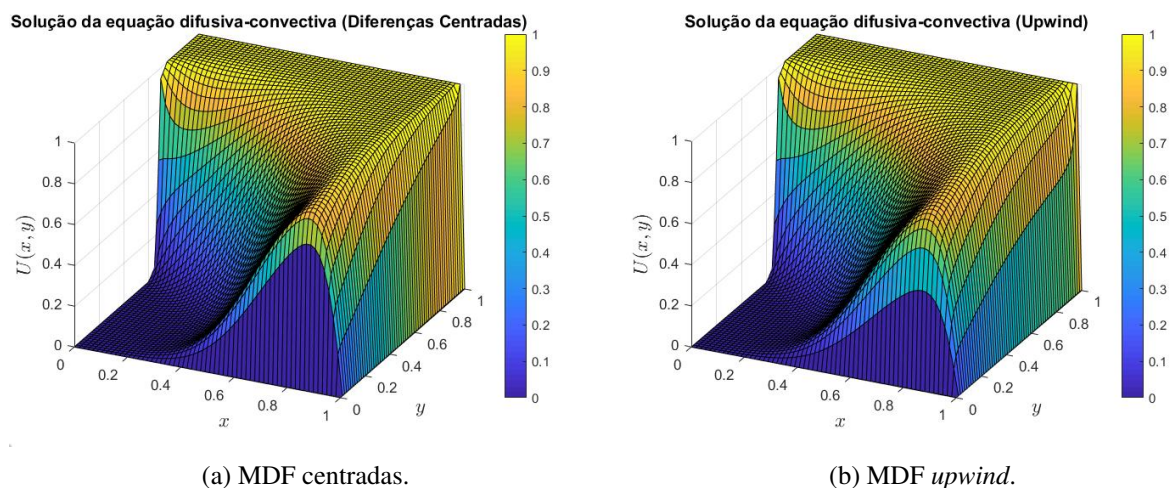
$$f(x, y) = \frac{1}{0.02\pi} e^{-\frac{1}{0.02}[(x-0.5)^2 + (y-0.5)^2]} \quad (8)$$


 Figura 3: Solução do segundo problema para $k = 10^{-3}$.

 Figura 4: Solução do segundo problema para $k = 10^{-6}$.

Observa-se que conforme o coeficiente difusivo é reduzido, o panorama do solução se torna dominado pelo fenômeno convectivo e desta forma tem-se uma concentração maior de U na direção do fluxo de velocidade circular imposto. Nota-se também que para $k = 10^{-3}$ o MDF centradas apresenta instabilidade na forma de oscilações espúrias na base da função Gaussiana e próximo à região de condição de contorno, já para $k = 10^{-6}$ o panorama está povoado de oscilações, sendo coerente com esperado de acordo com a literatura, uma vez que $P_e \gg 1$.

O terceiro e último problema analisado contém as condições de contorno dada pelas Equações 9 a 16. O coeficiente difusivo considerado é $k = 10^{-2}$, já o campo de velocidade adotado é $\vec{w} = (y, -x)$. Por fim, as malhas consideradas são $n_x = n_y = 20$ e $n_x = n_y = 60$, para que se possa observar como a solução se altera com o maior número de pontos discretos. Abaixo podem ser vistos as soluções numéricas obtidas, representas pelas Figuras 5 e 6.

$$\begin{cases}
 U(x, 0) = 0 & \forall x \in [0, 1] & (9) \\
 U(x, 1) = 1 & \forall x \in [0, 1] & (10) \\
 U(1, y) = 0 & \forall y \in [0, 1] & (11) \\
 U(0, y) = 0 & \forall y \in [0, 0.6[& (12) \\
 U(0, y) = y - 0.6 & \forall y \in [0.6, 0.65[& (13) \\
 U(0, y) = 18(y - 0.65) + 0.05 & \forall y \in [0.65, 0.70[& (14) \\
 U(0, y) = (y - 0.70) + 0.95 & \forall y \in [0.70, 0.75[& (15) \\
 U(0, y) = 1 & \forall y \in [0.75, 1] & (16)
 \end{cases}$$


 Figura 5: Solução do segundo problema para $nx = ny = 20$.

 Figura 6: Solução do segundo problema para $nx = ny = 60$.

Considerando as Figuras 5 e 6, percebe-se que conforme a malha se torna mais densa em relação ao número de nós, a solução apresentada pelo MDF centradas se torna mais suave, com a solução apresentada por $nx = ny = 60$ se mostrando livre de instabilidades, com um panorama similar ao apresentado

pelo MDF *upwind*. Ainda, pode se notar que a solução apresentada nesse caso pelo MDF *upwind* indica maior difusão numérica, evidenciada pela suavização da região adjacente aos pontos próximos à condição de contorno $x \in [0.6, 1]$, em comparação ao observado com a formulação centrada.

Conclusões

Para o primeiro problema, com solução sem camadas-limite, a formulação centrada é mais precisa que o *upwind*, se apresentando livre de instabilidades, sendo esta uma constatação coerente com o esperado uma vez que o MDF centradas é uma metodologia de segunda ordem de precisão, e o *upwind* é de primeira ordem de precisão, tendo esta última apresentado excessiva suavização da solução.

Para o segundo e terceiro problema, que possuem camadas-limite em suas soluções, a formulação centrada apresenta oscilações espúrias, resultantes da instabilidade que surge quando o método é submetido à um cenário de convecção dominante. Neste contexto, percebe-se que o refino da malha tende a reduzir o número de oscilações percebidas. Entretanto, o aumento do número de nós na malha, acarreta no aumento do uso de memória e tempo de processamento, recursos estes escassos.

Ainda, para o segundo problema, mesmo nas malhas mais refinadas produzidas neste estudo foram observadas oscilações espúrias quando o coeficiente difusivo utilizado é inferior a 10^{-3} , para o MDF centradas. Neste contexto de convecção dominante, a formulação *upwind* se mostrou estável, uma vez que não apresentou oscilações espúrias para os problemas contendo camada-limite interna ou externa.

Referências

- CARMO, Eduardo Gomes Dutra do; BENITEZ, Gustavo Alvarez. A new stabilized finite element formulation for scalar convection–diffusion problems: the streamline and approximate upwind/Petrov–Galerkin method. **Computer methods in applied mechanics and engineering**, v. 192, n. 31-32, p. 3379-3396, 2003.
- FRERICHS, Derk; JOHN, Volker. On reducing spurious oscillations in discontinuous Galerkin (DG) methods for steady-state convection–diffusion equations. **Journal of Computational and Applied Mathematics**, v. 393, p. 113487, 2021.
- LEVEQUE, Randall J. **Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems**. Society for Industrial and Applied Mathematics, 2007.
- MALISKA, Clovis Raimundo. **Transferência de calor e mecânica dos fluidos computacional**. Grupo Gen-LTC, 2017.
- PATANKAR, Suhas. **Numerical heat transfer and fluid flow**. Taylor & Francis, 2018.
- PLETCHER, Richard H.; TANNEHILL, John C.; ANDERSON, Dale. **Computational fluid mechanics and heat transfer**. CRC press, 2012.
- THOMAS, James William. **Numerical partial differential equations: finite difference methods**. Springer Science & Business Media, 2013.
- VOLKER, John; KNOBLOCH, Petr. On spurious oscillations at layers diminishing (SOLD) methods for convection–diffusion equations: Part I–A review. **Computer methods in applied mechanics and engineering**, v. 196, n. 17-20, p. 2197-2215, 2007.
- ZHAO, Shubo et al. Two types of spurious oscillations at layers diminishing methods for convection–diffusion–reaction equations on surface. **Numerical Heat Transfer, Part A: Applications**, v. 74, n. 7, p. 1387-1404, 2018.

Tool for simulating tumor growth in various regions of the human body in three dimensions.

Carlos Carret Miranda

University of Havana

carlos.carret@estudiantes.matcom.uh.cu

Reinaldo Rodríguez Ramos

Department of Mathematics, Faculty of Mathematics and Computer Science, University of Havana, Cuba, and PPG-MCCT, Federal Fluminense University, Volta Redonda, Rio de Janeiro, Brazil.

reinaldo@matcom.uh.cu and reinaldorr@id.uff.br

Panthers Rodríguez Bermúdez

Departamento de Ciências Exatas, Universidade Federal Fluminense, Volta Redonda, Rio de Janeiro, Brazil

pantersrb@id.uff.br

Summary: Cancer is a disease characterized by the uncontrolled growth of abnormal cells in the body. It is a complex and multifaceted disease that has challenged researchers and doctors for decades. The ability to visualize and understand tumor growth can provide valuable insights into how cancer develops and spreads, leading to significant improvements in cancer diagnosis, treatment, and prevention. The three-dimensional tumor growth simulation tool that is being developed is an important step in this direction. It allows for detailed visualization of tumor growth in different parts of the human body, which can provide valuable insights into how cancer develops and spreads. Additionally, the ability of this tool to simulate tumor growth in different parts of the body means that it can be used to study a wide range of cancer types. This tool utilizes a cellular automaton and a small-world network to create connections between cells, allowing for a more accurate representation of organ and tumor structures. Furthermore, it allows for the loading of configurations and parameters from external files, providing great flexibility to the tool and allowing for customization of the simulation to the specific needs of each case. For 3D rendering, the Marching Cubes technique is used, which enables detailed and accurate three-dimensional representation of tumors.

Keywords: Scientific Computing, Numerical Methods, and Applications. Cellular Automaton. Marching Cubes. 3D. Cancer.

Introduction

The challenge of representing biological phenomena mathematically, physically, and computationally requires interdisciplinary synergy among experts in these fields. This collaboration enriches the traditional experimental method used in biological sciences by implementing mathematical models, which serve as tools to formulate and test hypotheses, guide experimental research, and refine the model based on the obtained results. [7]

Cancer is a disease that affects a large number of living organisms and is characterized by the presence of a group of abnormal cells that grow uncontrollably, disregarding the normal rules of cell division. It particularly affects humans, where its occurrence and development pose a threat to life. The malignancy of cancer varies and depends on factors such as the growth rate of cancer cells, their ability to spread to other tissues, and the possibility of recurrence after surgical removal. [7]

The purpose of this type of research is to achieve a deeper understanding of biological processes through an iterative cycle of theory and experimentation. Additionally, mathematical models can be used to assist in the conception and design of therapeutic strategies, providing a more precise and personalized insight into the treatment of each patient. [7]

In the case of this project, a cellular automaton and a small-world network are used to model the interactions between cells, providing a more accurate representation of tumor growth. The parameters and configurations can be loaded from external files, offering great flexibility in adapting the simulation to the specific needs of each case. The work being conducted is a part of a larger thesis project that is currently in progress. The project aims to simulate tumor growth in small organs and involves loading and utilizing parameters for the simulation. The graph of cells with their connections is visualized and analyzed, along with the visualization of the tumor size throughout the simulation. The focus of the next stages is to improve the existing implementation and incorporate changes based on details found in medical literature. Additionally, efficient algorithms will be implemented to process large amounts of cells and their connections. In the future, the project aims to incorporate Artificial Intelligence and Machine Learning techniques to achieve better approximations to reality.

The technique of Marching Cubes [5] is used for 3D rendering, providing a detailed and precise visualization of tumors. This visualization can provide valuable insight into how cancer develops and spreads, which can be essential for the development of effective therapies and treatments. By visualizing tumor growth in three dimensions, doctors and scientists can gain a better understanding of tumor evolution and how it may affect surrounding tissues. This information can be crucial for the development of effective therapies and treatments for cancer.¹

Submission

Cellular Automaton

In this section, the cellular automaton model presented in this work is conceived. It begins by formally defining a cellular automaton [7].

A cellular automaton is a tuple $(\mathcal{L}; \mathcal{N}; \mathcal{E}; \mathcal{R})$ composed of the following representative elements [2]:

$\mathcal{L}$: $\mathcal{L}$ is a potentially infinite set of cells.

$\mathcal{N}$: $\mathcal{L} \times \mathcal{L} \rightarrow \{0, 1\}$ is a neighborhood function, which can be seen as a relation, usually reflexive and symmetric, between cells. This function shows which pairs of cells are neighbors, that is, the geometry of the cellular organization.

$\mathcal{E}$: It is a set of states. Each cell in the set $\mathcal{L}$ is assigned an associated state at each time step.

$\mathcal{R}$: $\mathcal{E}^{|\mathcal{N}(v)|} \rightarrow \mathcal{E}$ is a locally defined transition function. This function is the core of the dynamics of a cellular automaton and is commonly expressed through rules that define the state of the cell in the next time step based on the state of the neighboring cells. The set containing the state of the neighboring cells is obtained through the function $\mathcal{N}(v)$, which is defined below.

The sets $A^n(G)$ and $A^d(G)$ group the edges of the graph that correspond to immediate and distant connections, respectively. These sets have the following properties:

$$A^n(G) \cup A^d(G) = A(G), \quad (1a)$$

$$A^n(G) \cap A^d(G) = \emptyset. \quad (1b)$$

These properties indicate that the subsets of edges $A^n(G)$ and $A^d(G)$ form a partition of the set of edges $A(G)$

Based on the sets of vertices $V(G)$ and edges $A(G)$, the representative elements $\mathcal{L}$ and $\mathcal{N}$ of the cellular automaton model are defined as follows:

The set of cells $\mathcal{L}$ is defined based on the set of vertices of the graph $V(G)$:

$$\boxed{\mathcal{L} = V(G)}. \quad (2)$$

The neighborhood function $\mathcal{N}$ is defined based on the set of edges of the graph $A(G)$ as shown below:

$$\mathcal{N} : V(G) \times V(G) \rightarrow \{0, 1\}, \quad (3a)$$

$$\mathcal{N}(v, w) = \begin{cases} 0 & \text{si } \{v, w\} \notin A(G) \\ 1 & \text{si } \{v, w\} \in A(G) \end{cases}, \quad (3b)$$

In other words, the vertices $v \in V(G)$ and $w \in V(G)$ are neighbors in the cellular automaton if there exists an edge in G that connects them.

The neighborhood of the vertex $v \in V(G)$ is defined based on the neighborhood function $\mathcal{N}(v, w)$ as the set of vertices $\mathcal{N}(v)$ that have edges with the vertex v :

$$\mathcal{N}(v) = \{w \mid \mathcal{N}(v, w) = 1\}. \quad (4)$$

Set of cells: Watts-Strogatz model

In the presented study, a soft tissue is defined as a set of cells that exhibit two types of connections: between nearby neighboring cells and between distant cells. To represent these types of connections, a cellular automaton model based on a graph network is used. In their work [9], Duncan J. Watts and Steven H. Strogatz showed that there are many biological, technological, and social networks that lie between regular and random networks, which have traditionally been used to model different types of dynamic systems.

Let v be a vertex of the graph that has k_v edges connecting it to k_v vertices. The value between the actual number of edges K_v that exist between these k_v vertices and the maximum number of possible edges¹ $k_v(k_v - 1)/2$ is the clustering coefficient of vertex v and is determined as [7]:

$$C_v = \frac{2K_v}{k_v(k_v - 1)}. \quad (5)$$

The global clustering coefficient of the graph C_G is the average of all individual clustering coefficients C_v , that is [7]:

$$C_G = \frac{1}{|V(G)|} \sum_{v=1}^{|V(G)|} C_v. \quad (6)$$

The average path length is the mean of the distances between every pair of vertices belonging to the graph and is denoted as ℓ_G . Due to the existence of numerous distant connections through the circulatory system, the average path length in the network of cells is relatively small.

Therefore, it is hypothesized that a living tissue possesses a high clustering coefficient and a small average path length. These characteristics are characteristic of small-world networks, and they are used to represent living tissue. To generate small-world networks with these characteristics, the Watts and Strogatz model is used [9]. This model starts with a graph with q vertices, each connected to k immediate neighbors, and then randomly rewires each edge of the graph with a probability p , introducing edges that connect distant vertices.

Marching Cubes

The Marching Cubes technique is a computer graphics algorithm used to extract a polygonal mesh of an isosurface from a three-dimensional discrete scalar field, such as computed tomography (CT) scans and magnetic resonance imaging (MRI) data [5]. In the context of this project, it is used for the three-dimensional representation of tumors, providing detailed and accurate visualization.¹

¹The maximum number of possible edges is reached when the neighbors, denoted as k_v of vertex v belong to a clique. In an undirected graph, a clique is a set of vertices such that every pair of vertices is connected by an edge.

This algorithm works by processing the cells of volume data (also known as voxels), checking the intersection between their respective edges and the isosurface. The values of each vertex of the cells are compared with a given isosurface value, and these vertices are classified as "inside" or "outside" the isosurface. Once the type of intersection is determined, an approximation of the isosurface contained in the cell is constructed by building triangles [1].

The resulting visualization can provide valuable insights into how cancer develops and spreads. By visualizing tumor growth in three dimensions, doctors and scientists can gain a better understanding of tumor evolution and how it may impact surrounding tissues. This information can be essential for the development of effective cancer therapies and treatments.

Due to the high computational cost of representing and applying the Marching Cubes algorithm to realistic models containing millions of cells, this work implements a model scaling technique that reduces the length of the dimensions of the original cellular automaton model. The reduction process is as follows:

- Cells are grouped into quadrants of dimensions provided by the user.
- The states of all cells belonging to the quadrant are examined.
- The quadrant adopts the state that is most repeated among the cells in it.
- After performing this process for several quadrants, each quadrant reduces its size from $(n \times m \times l)$ to $(1 \times 1 \times 1)$, where $n \leq S_x, m \leq S_y, l \leq S_z$.

In summary, the Marching Cubes technique is a powerful tool for the three-dimensional visualization of medical data. In the context of cancer research, it can provide a detailed and accurate representation of tumor growth, which can significantly contribute to our understanding of this disease and the development of effective therapies and treatments.

Configuration and Parameters of the simulation.

Some of the parameters and configurations that can be modified are:

- S_x - Dimension of the space declared on the x-axis.
- S_y - Dimension of the space declared on the y-axis.
- S_z - Dimension of the space declared on the z-axis [7].

The ranges of values for the spatial components of the graph vertices are as follows: $0 \leq x \leq S_x$, $0 \leq y \leq S_y$ y $0 \leq z \leq S_z$.

- p - Probability of reconnection in the Watts-Strogatz model
- P_0^a, P_0^v - Initial populations of the avascular and vascular stages respectively [7].
- Parameters corresponding to the number of states that the automaton cells can have and their descriptions.
- Parameters for possible transitions between the states of the automaton.
- Parameters for the probabilities of the transitions between states.

When including parameters for the calculation of certain probabilities, one can take into account the calculation of the probability of interaction between tumor cells and the immune system [6].

- Parameters corresponding to the shape of the organs where the simulation will take place.

- Parameters to describe the schema of the organs where the simulation will be performed. This allows for the consideration of the characteristics of each organ separately and enables a more realistic simulation.

There are many other configurable parameters.

Conclusions

The growth of a tumor can be visualized in 3D using the Marching Cubes technique. It is widely used for medical visualizations, such as computed tomography (CT) and magnetic resonance imaging (MRI) images. Additionally, the Marching Cubes algorithm can reduce the computational time used for sampling in 3D reconstruction. However, one of the main issues with Marching Cubes is the presence of unused voxels that can be generated during the analysis of the coordinates and intensity values of 2D images. These unused voxels can affect the smoothness of the 3D surface [8].

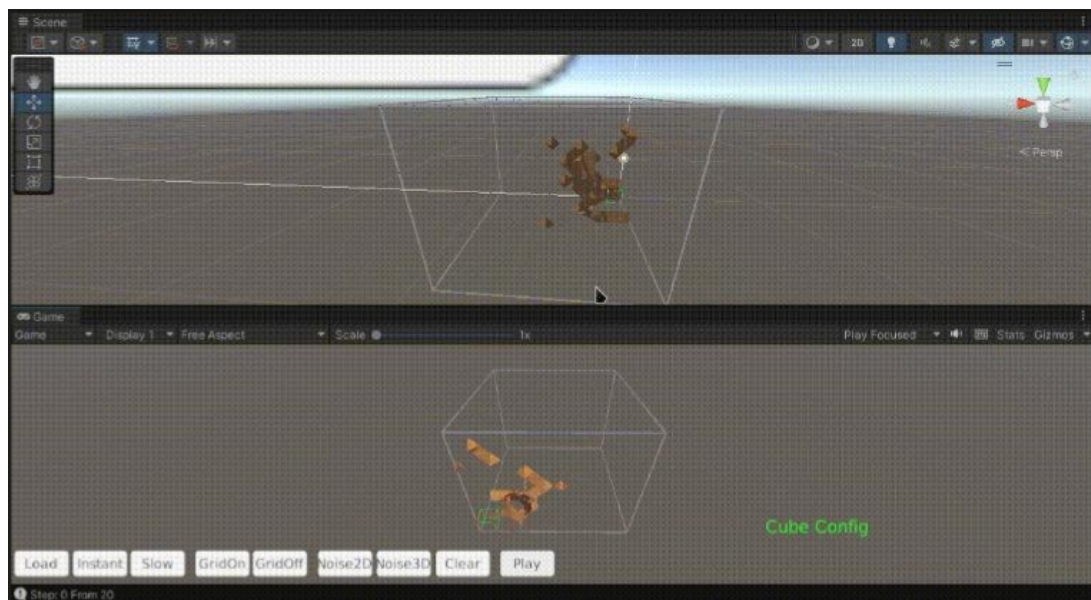


Figure 1: An example of simulating tumor growth in 3D at early stages. The code was developed in Unity

The creation of a tool to simulate tumor growth using cellular automaton in any organ of the human body is a significant advancement in the field of modeling and simulation of biological systems. This tool provides an innovative and flexible approach to studying tumor growth, which has important implications in both basic research and clinical applications.

The ability to load specific configurations and adjust, add, or remove parameters that influence the realism of the simulation allows for the adaptation of the model to different scenarios and conditions. This makes the tool highly versatile and applicable to a wide range of situations and types of tumors.

The use of cellular automaton in tumor growth simulation allows for the accurate representation of the interaction between cells and their environment, as well as the temporal changes that occur. This modeling approach provides a detailed understanding of tumor growth dynamics and can be used to investigate various scenarios and conditions [2].

How much progress has been made so far?

- The loading and usage of parameters for the simulation have been implemented.

- The engine to simulate the growth of a tumor in small organs has been implemented.
- The cell graph with its connections has been visualized and analyzed.
- The size that the tumor will take during the simulation has been visualized.
- The code to render the shape the tumor will have has been implemented.¹

What will be the focus in the next stages?

- Improving all of the above, for example, enhancing the implementation of 3D rendering using Marching Cubes to obtain smoother faces and triangles.
- Implementing changes between states following the details found in the literature.
- Implementing efficient algorithms to process large amounts of cells and their connections, we are talking about millions of both.
- Add Artificial Intelligence and Machine Learning techniques to obtain better approximations to reality.

Finally, the 3D visualization of tumor growth using the Marching Cubes technique can provide a valuable tool for healthcare professionals to better understand the dynamics of tumor growth and develop more effective treatment strategies.

Acknowledgments

R.R.R thanks EDITAL UFF PROPPI No. 05/2022 and PPG-MCCT of Universidade Federal Fluminense, Brazil. P.R.B thanks CAPES, CNPq, FAPERJ, and UFF.

Appendix

The following tools were used to develop this work:

- For the generation of the simulation, C# in its .net7.0 version.
- For the visual part to visualize the cells and their connections in certain regions, Python in its 3.11 version with Streamlit and other dependencies.
- For the visual part to render the tumor, Unity in its 2021.3.28f1 version.

References

- [1] Cirne, M. A.; Pedrini, H. Marching cubes technique for volumetric visualization accelerated with graphics processing units. *Journal of the Brazilian Computer Society*, vol. 19, pages 223-233, 2013. Available at: <https://journal-bcs.springeropen.com/articles/10.1007/s13173-012-0097-z>.
- [2] Deutsch, A.; Maini, P.; Dormann, S. Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis. *Modeling and Simulation in Science, Engineering and Technology*. Birkhauser Boston, 2007.
- [3] Guinot, V. Modelling using stochastic, finite state cellular automata: rule inference from continuum models. *Applied Mathematical Modelling*, 26:701–714, 2002.

- [4] Kansal, A.; Torquato, S. Simulated brain tumor growth dynamics using a three-dimensional cellular automaton. *Journal of Theoretical Biology*, 203:367–382, 2000.
- [5] Lorensen, W. E.; Cline, H. E. Marching cubes: A high-resolution 3D surface construction algorithm. *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 163-169, 1987. Available at: <https://people.eecs.berkeley.edu/~jrs/meshpapers/LorensenCline.pdf>.
- [6] Ruanxiaogang, H. A simple cellular automaton model for tumor-immunity system. In *Robotics, Intelligent Systems and Signal Processing, 2003. Proceedings. 2003 IEEE International Conference on*, 2.
- [7] Viera Barredo, D. Universidad de La Habana, Facultad de Matemática y Computación. Departamento de Matemática. Autómata celular estocástico en redes complejas para el estudio de la invasión, migración y metástasis del cáncer. Asesores: Reinaldo Rodríguez Ramos, Rubén Interián, Ariel Ramírez Torres, Rocío Rodríguez Sánchez. Undergraduate thesis. June, 2019.
- [8] Visutsak, P. Marching Cubes and Histogram Pyramids for 3D Medical Visualization. *Journal of Imaging*, vol. 6, 2020. Available at: <https://ncbi.nlm.nih.gov/pmc/articles/PMC8321043> and <https://api.semanticscholar.org/CorpusID:225446991>.
- [9] Watts, D. J. & Strogatz, S. H. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.

UM TUTORIAL PARA A IMPLEMENTAÇÃO DO ALGORITMO QUÂNTICO DE DEUTSCH-JOZSA

Silas da Silva Caldas Júnior

Universidade Federal Rural do Rio de Janeiro

silas.caldas10@gmail.com

Prof. Dr. Marcelo Ferreira Farias

Universidade Federal Rural do Rio de Janeiro

mfarias@ufrj.com

Me. Bruno Clemente Guingo

Laboratório Nacional de Computação Científica

bruno.guingo@gmail.com

Prof. Dr. Erito Marques de Souza Filho

Universidade Federal Rural do Rio de Janeiro

proferitomarques@gmail.com

Resumo: A computação quântica é um campo em desenvolvimento cujo o objetivo é superar a eficiência da computação clássica na resolução de problemas através de algoritmos mais eficientes. O presente trabalho tem como objetivo apresentar o algoritmo quântico de Deutsch-Jozsa, fornecendo um tutorial para sua implementação em *Python 3*. Para este fim, foi utilizado o *Qiskit*, um kit de desenvolvimento de software de código aberto desenvolvido pela *IBM*, que permite aos usuários programar, simular e executar algoritmos quânticos através de computadores quânticos reais ou de simuladores quânticos. O ambiente de programação utilizado foi o *Jupyter Notebook*, desenvolvido pela organização *Project Jupyter*.

Palavras-chave: Computação Quântica. Algoritmo de Deutsch-Jozsa. Qiskit.

Introdução

Na área da computação quântica, o algoritmo conhecido como Deutsch-Jozsa, que foi proposto no ano de 1992 pelos cientistas David Deutsch e Richard Jozsa, trouxe uma nova abordagem ao evidenciar um cenário exemplar em que a capacidade da computação quântica sobrepuja de maneira significativa sua contraparte clássica. Nesse contexto, considera-se um oráculo que incorpora uma função cujas características são desconhecidas. O propósito consiste em determinar se essa função é constante, ou seja, se ela retorna o mesmo resultado para todas as entradas, ou se é balanceada, isto é, se ela apresenta dois resultados distintos de maneira igualitária. No contexto clássico, no cenário mais desfavorável, seria imperativo avaliar cinquenta por cento das entradas da função para resolver a questão. Porém, o Algoritmo de Deutsch-Jozsa ilustra que, através da utilização da computação quântica, é viável alcançar a solução com apenas uma única consulta, resultando em uma aceleração de natureza exponencial (DEUTSCH e JOZSA, 1992, p.556-557).

Portas Quânticas

Nesta seção, serão introduzidas as portas lógicas quânticas utilizadas ao longo do trabalho. A primeira é a porta Hadamard, representada matematicamente pela matriz unitária H abaixo. A segunda porta quântica utilizada é a porta X , de um qubit, também conhecida como porta NOT . Por fim, há ainda a porta CX (*Controlled-X*), também conhecida como $CNOT$ (*Controlled-NOT*). Trata-se da versão controlada da porta X , sendo uma porta de dois qubits, onde um deles é um qubit de controle e outro é o qubit alvo. Essa porta atua da seguinte maneira: se o qubit de controle estiver no estado $|0\rangle$, nada ocorre; se o qubit de controle estiver no estado $|1\rangle$, o qubit alvo sofrerá uma rotação de π rad no eixo x da Esfera de Bloch. Esta porta não apenas atua sobre os estados da base computacional, como também pode ser aplicada a qubits superpostos, e mesmo emaranhados. A porta CX é representada matematicamente pela matriz unitária CX abaixo (PORTUGAL, 2012, p. 29,33).

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; CX = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Algoritmo de Deutsch-Jozsa

Considerando-se uma função booleana $f : \{0, 1\}^n \rightarrow \{0, 1\}$, na qual cada entrada binária de n bits produz uma saída binária de 1 bit, pode-se observar que, de acordo com esta definição, o conjunto total de possibilidades para as entradas de f é composto por 2^n variantes. Ao mesmo tempo, as opções de saída possíveis para f são limitadas aos valores 0 e 1. A função f é caracterizada como constante caso sua saída seja exclusivamente 0 ou 1, independente do input, e é qualificada como balanceada quando a quantidade de saídas iguais a 0 equivale ao número de saídas iguais a 1. No contexto da problemática sob análise, a função f assume somente dois possíveis estados: constante ou balanceada. Com base nas saídas produzidas por f , a meta reside em determinar qual dessas categorias é aplicável a f . No pior cenário, utilizando-se um computador clássico, a primeira metade dos inputs avaliados resultaria no mesmo output, implicando que a conclusão a respeito do comportamento da função só seria obtida após testar o input subsequente. Nesse contexto, a realização de $2^{n-1} + 1$ testes seria necessária (DEUTSCH e JOZSA, 1992, p.555). Por outro lado, ao se aplicar o algoritmo de Deutsch-Jozsa para resolver essa mesma problemática, utiliza-se o operador unitário U_f , um oráculo quântico que atua da seguinte maneira: dado um conjunto de estados $|x\rangle$ composto por n qubits e um estado $|y\rangle$ constituído por 1 qubit, a transformação realizada por U_f é expressa por

$$|x\rangle|y\rangle \xrightarrow{U_f} |x\rangle|y \oplus f(x)\rangle.$$

Para a implementação do algoritmo, o circuito quântico utiliza dois registradores, o primeiro possui n qubits inicializados no estado $|0\rangle$, enquanto o segundo contém 1 qubit inicializado no estado $|1\rangle$. O procedimento do algoritmo segue as seguintes etapas (IBM 2023):

1. É executada a operação de produto tensorial entre o estado $|0\rangle^{\otimes n}$ do primeiro registrador e o estado $|1\rangle$ do segundo registrador, resultando no estado $|\psi_0\rangle = |0\rangle^{\otimes n}|1\rangle$.
2. A porta de Hadamard é aplicada a cada qubit de $|\psi_0\rangle$, gerando $|\psi_1\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} |x\rangle (|0\rangle - |1\rangle)$.
3. O operador U_f é aplicado a $|\psi_1\rangle$, resultando em $|\psi_2\rangle = \frac{1}{\sqrt{2^{n+1}}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle (|0\rangle - |1\rangle)$.

4. O segundo registrador é ignorado, e a porta Hadamard é aplicada a cada qubit do primeiro registrador, formando $|\psi_3\rangle = \frac{1}{2^n} \sum_{z=0}^{2^n-1} \left[\sum_{x=0}^{2^n-1} (-1)^{f(x)} (-1)^{xz} \right] |z\rangle$, onde $xz = x_0z_0 \oplus x_1z_1 \oplus \dots \oplus x_{n-1}z_{n-1}$, que é a soma do produto bit-a-bit.
5. Por fim, o primeiro registrador é medido.

Deve-se observar que o estado $|\psi_3\rangle$ consiste em uma superposição de 2^n estados de n qubits. Dentro dessa "coleção", encontra-se o estado $|0\rangle^{\otimes n}$, cuja amplitude em $|\psi_3\rangle$ é dada por $\frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)}$. Após se medir o estado do primeiro registrador, a probabilidade de se obter $|0\rangle^{\otimes n}$ é igual a $\left| \frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} \right|^2$, cujo valor será 1 caso $f(x)$ seja constante e 0 caso $f(x)$ seja balanceada. Assim, após a medição do estado do primeiro registrador, a obtenção do estado $|0\rangle^{\otimes n}$ mostrará que f é constante; qualquer outro resultado basta para concluir que a função é balanceada. Com isso, uma única aplicação do algoritmo de Deutsch-Jozsa basta para se saber se f é constante ou balanceada (CLEVE, 1997, p. 4). Na Figura 1, exibe-se o circuito associado ao algoritmo de Deutsch-Jozsa.

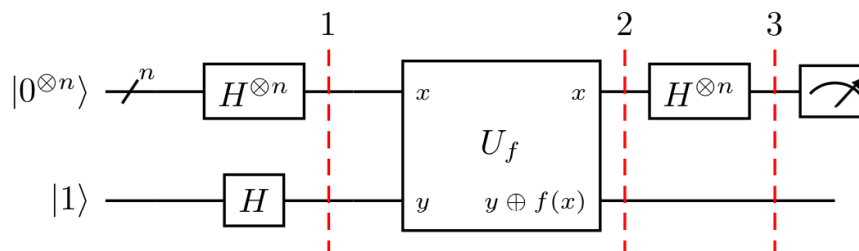


Figura 1: Circuito lógico do algoritmo de Deutsch-Jozsa (IBM, 2023).

Implementação

O primeiro passo para a implementação computacional do algoritmo de Deutsch-Jozsa no Jupyter Notebook é a importação dos recursos que serão necessários. O Jupyter Notebook é uma ferramenta de código livre baseada em browser, que funciona como um laboratório virtual onde se pode programar em linguagens de programação como Python 3 (RANGLES et al, 2017).

```
In [1]: import numpy as np

from qiskit import IBMQ, Aer
from qiskit.providers.ibmq import least_busy
from qiskit import QuantumCircuit, assemble, transpile
from qiskit.visualization import plot_histogram
import matplotlib.pyplot as plt

from qiskit.visualization import plot_histogram
```

Figura 2: Realização das importações necessárias (IBM, 2023).

Após isso, pode-se criar o oráculo U_f , conforme a Figura 3.

```
In [19]: def dj_oracle(case, n):
    oracle_qc = QuantumCircuit(n+1)
    if case == "balanced":
        b = np.random.randint(1,2**n)
        b_str = format(b, '0'+str(n)+'b')
        for qubit in range(len(b_str)):
            if b_str[qubit] == '1':
                oracle_qc.x(qubit)
        for qubit in range(n):
            oracle_qc.cx(qubit, n)
        for qubit in range(len(b_str)):
            if b_str[qubit] == '1':
                oracle_qc.x(qubit)

    if case == "constant":
        output = np.random.randint(2)
        if output == 1:
            oracle_qc.x(n)

    oracle_gate = oracle_qc.to_gate()
    oracle_gate.name = "Oracle"
    return oracle_gate
```

Figura 3: Criação do oráculo (IBM, 2023).

O código então segue com a implementação do algoritmo completo através da função *dj_algorithm*, como mostrado na Figura 4.

```
In [20]: def dj_algorithm(oracle, n):
    dj_circuit = QuantumCircuit(n+1, n)
    dj_circuit.x(n)
    dj_circuit.h(n)
    for qubit in range(n):
        dj_circuit.h(qubit)
    dj_circuit.append(oracle, range(n+1))
    for qubit in range(n):
        dj_circuit.h(qubit)

    for i in range(n):
        dj_circuit.measure(i, i)

    return dj_circuit
```

Figura 4: Aplicação do algoritmo de Deutsch-Jozsa (IBM, 2023).

Testando o algoritmo

Nesta seção, será rodado o algoritmo para cada um dos dois casos (i.e, oráculo balanceado e oráculo constante). Em ambos, serão realizadas sucessivas medições ao final do algoritmo, resultando em uma distribuição de probabilidades, que será analisada e comparada com o resultado esperado. Nos seguintes testes, será considerado $n = 4$.

Caso 1: Oráculo Balanceado

O primeiro caso é o de oráculo balanceado. Como na Figura 5, definido o caso a ser analisado, transpila-se o circuito *dj_circuit* para uma forma que seja adequada para a execução no simulador quântico.

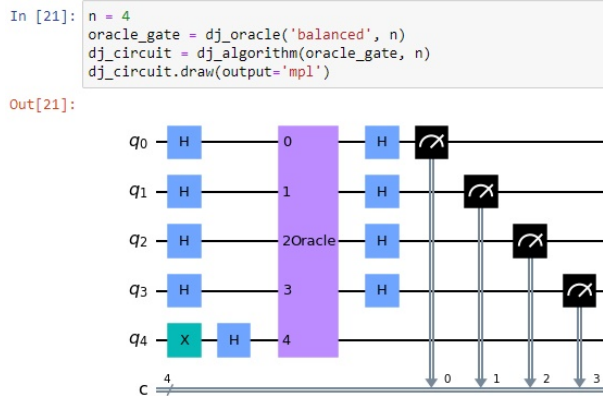


Figura 5: Chamando o oráculo balanceado (IBM, 2023).

O circuito é executado e o histograma dos resultados é plotado usando a função `plot_histogram(answer)`. Esse histograma mostrará a distribuição de probabilidade dos resultados obtidos após a execução do circuito. Neste exemplo, conforme mostrado na Figura 6, o registrador foi medido 1024 vezes, e em todas elas o estado obtido foi o $|1111\rangle$.

```
In [22]: transpiled_dj_circuit = transpile(dj_circuit, aer_sim)
qobj = assemble(transpiled_dj_circuit)
results = aer_sim.run(qobj).result()
answer = results.get_counts()
plot_histogram(answer)
```

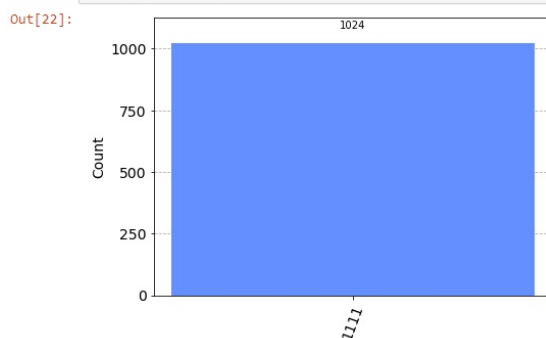


Figura 6: Distribuição de probabilidades para o caso balanceado (IBM, 2023).

Caso 2: Oráculo Constante

No caso em que o oráculo é constante, realiza-se o mesmo procedimento; agora, porém, a variável `oracle_gate` é definida chamando-se a função `dj_oracle` com os argumentos `'constant'` e `n`. O procedimento é conforme mostrado na Figura 7.

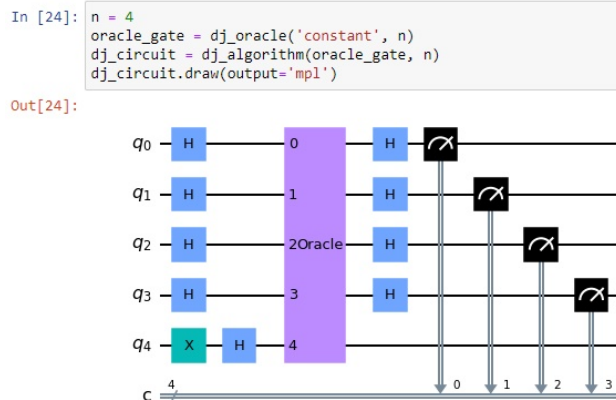


Figura 7: Chamando o oráculo constante (IBM, 2023).

É passado o valor *'constant'* como *case*, indicando que o oráculo para uma função constante é o desejado. O procedimento restante é exatamente como no caso em que o oráculo é balanceado. Resta, então, analisar a distribuição de probabilidades relativas às medições do estado do primeiro registrador, que é mostrado na Figura 8.

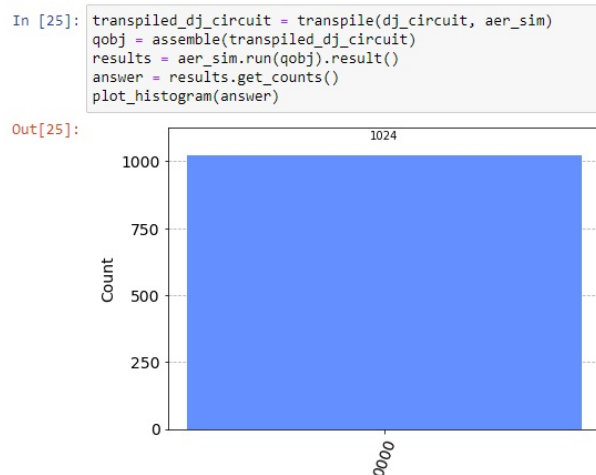


Figura 8: Distribuição de probabilidades para o caso constante (IBM, 2023).

Foram realizadas 1024 medições, e em todas elas o estado $|0000\rangle$ foi obtido.

Conclusões

Em comparação aos algoritmos clássicos empregados na resolução do problema proposto, o algoritmo de Deutsch-Jozsa se destaca por sua eficiência notável. De fato: enquanto múltiplas consultas seriam necessárias a um algoritmo clássico para determinar se uma função desconhecida é constante ou balanceada, o algoritmo de Deutsch-Jozsa se mostra capaz de realizar a tarefa com uma única consulta. Embora o algoritmo não tenha aplicações práticas imediatas, ele desempenhou um papel fundamental na demonstração das capacidades dos computadores quânticos. Contudo, se fosse utilizado um computador clássico probabilístico, onde os inputs são escolhidos aleatoriamente, ao invés da abordagem determinística utilizada, também seria possível determinar o comportamento da função rapidamente (NIELSEN e CHUANG, 2002). Vale também destacar a importância da biblioteca *Qiskit Aer* na simulação do circuito quântico, bem como na representação visual dos resultados. Graças à capacidade de simulação,

pode-se verificar precisamente o funcionamento do algoritmo sem que seja necessário utilizar um computador quântico real.

Referências

CLEVE, R. et al. Quantum Algorithms Revisited. **Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 454, n.1969, 1997.

DEUTSCH, D.; JOZSA, R. Rapid solutions of problems by quantum computation. **Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 439, n.1907, p. 553-558, 1992.

IBM. Deutsch-Jozsa Algorithm. 2023. Disponível em:
<https://learn.qiskit.org/course/ch-algorithms/deutsch-jozsa-algorithm>. Acesso em: 12 ag. 2023.

NIELSEN, M.A., CHUANG, I. **Quantum Computation and Quantum Information**. Cambridge: Cambridge University Press, 2002.

PORTUGAL, R. et al. **UMA INTRODUÇÃO À COMPUTAÇÃO QUÂNTICA**. 2.ed. São Paulo: Sociedade Brasileira de Matemática Aplicada e Computacional, 2012.

RANDLER, B. et al. Using the Jupyter Notebook as a tool for open science: An empirical study. In: ACM/IEEE Joint Conference on Digital Libraries, 2017, Toronto. Toronto: Institute of Electrical and Electronics Engineers (IEEE), 2017. p. 1-2.

UM TUTORIAL PARA A IMPLEMENTAÇÃO DO ALGORITMO QUÂNTICO DE GROVER

Silas da Silva Caldas Júnior

*Universidade Federal Rural do Rio de Janeiro
silas.caldas10@gmail.com*

Prof. Dr. Marcelo Ferreira Farias

*Universidade Federal Rural do Rio de Janeiro
mfarias@ufrj.com*

Me. Bruno Clemente Guingo

*Laboratório Nacional de Computação Científica
bruno.guingo@gmail.com*

Prof. Dr. Erito Marques de Souza Filho

*Universidade Federal Rural do Rio de Janeiro
proferitomarkes@gmail.com*

Resumo: Este trabalho visa a apresentação do algoritmo de Grover na computação quântica, fornecendo um guia para a sua implementação prática. Para alcançar este propósito, utilizou-se o ambiente *Qiskit*, uma plataforma de código aberto criada pela *IBM* que permite aos usuários programar, simular e executar algoritmos quânticos, tanto em computadores quânticos reais quanto simuladores. A programação foi conduzida no ambiente *Jupyter Notebook*, desenvolvido pela equipe do *Project Jupyter*. A linguagem utilizada foi o *Python 3*.

Palavras-chave: Computação Quântica. Algoritmo de Grover. Qiskit.

Introdução

No campo da computação quântica, o algoritmo de Grover, desenvolvido por Lov Grover em 1996, é utilizado para de maneira eficaz recuperar informações de um conjunto de dados não ordenado, proporcionando uma clara vantagem em comparação com os algoritmos convencionais. Esse algoritmo é utilizado principalmente para a identificação de um elemento específico (chamado de item marcado, ou *winner*) em uma lista não ordenada contendo N elementos, na qual somente um item é a resposta desejada. Enquanto a busca um cenário clássico exigiria um tempo proporcional a N para localizar o item procurado, o algoritmo de Grover permite encontrar o item em um tempo proporcional a $\sqrt{N}$ (PORTUGAL, 2012, p.37; GROVER, 1996, p.1).

1 O problema

Toma-se em consideração um conjunto de índices não ordenado $L = \{0, 1, \dots, N-1\}$, em que $N = 2^n$ e n denota a quantidade de bits em cada item do conjunto. Desse modo, L engloba a totalidade dos números inteiros não negativos que podem ser representados por n bits e expressos no sistema decimal. O desafio reside na identificação de um item específico a_w da lista $L' = \{a_0, a_1, \dots, a_{N-1}\}$, onde $w \in L$. Classicamente, uma função $f : \{0, 1, \dots, N-1\} \rightarrow \{0, 1\}$ é definida, atuando do seguinte modo:

$$f(x) = \begin{cases} 1, & \text{se } x = w \\ 0, & \text{se } x \neq w \end{cases}$$

Na situação mais desfavorável, será preciso examinar todos os N elementos presentes na lista.

2 Solução através da computação quântica

O algoritmo de Grover utiliza um registrador quântico com n qubits inicializados no estado $|0\rangle$ cada um. Cada elemento $x \in L$ é associado a um estado quântico $|x\rangle$ de n qubits da base computacional $\mathcal{B} = \{|0\rangle, |1\rangle, \dots, |N-1\rangle\}$. Introdz-se o operador U_w , que opera sobre $|x\rangle$ como delineado a seguir:

$$U_w|x\rangle = \begin{cases} |x\rangle, & \text{se } x \neq w \\ -|x\rangle, & \text{se } x = w \end{cases}$$

Portanto, U_w atua sobre um estado $|x\rangle$ conforme a expressão abaixo (IBM, 2023):

$$|x\rangle \xrightarrow{U_w} (-1)^{f(x)}|x\rangle.$$

Em termos matriciais, o operador U_w pode ser representado como

$$U_w = \begin{bmatrix} (-1)^{f(0)} & 0 & \dots & 0 \\ 0 & (-1)^{f(1)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (-1)^{f(2^n-1)} \end{bmatrix}.$$

A execução do algoritmo de Grover começa com a aplicação da porta Hadamard em cada qubit do registrador quântico, gerando o estado superposto $|s\rangle$ (CHUANG e NIELSEN, 2002, p.250).

$$|s\rangle = H^{\otimes n}|0\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle.$$

É fácil ver que os vetores $|s\rangle$ e $|w\rangle$ são linearmente independentes, gerando um plano bidimensional no espaço $\mathbf{C}^N$. Denotando-se por $|s'\rangle$ um estado de n qubits desse plano, onde $|s'\rangle$ é ortogonal a $|w\rangle$ e θ é o ângulo entre $|s\rangle$ e $|s'\rangle$, tem-se que o estado $|s\rangle$ é uma combinação linear de $|s'\rangle$ e $|w\rangle$ (IBM, 2023).

$$|s\rangle = \sin(\theta)|w\rangle + \cos(\theta)|s'\rangle,$$

onde

$$\theta = \arcsin(\langle s|w\rangle) = \arcsin\left(\frac{1}{\sqrt{N}}\right).$$

A próxima etapa do algoritmo de Grover consiste na aplicação de U_w a $|s\rangle$, gerando $|\psi_1\rangle$, conforme a expressão abaixo.

$$|\psi_1\rangle = U_w|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} (-1)^{f(x)}|x\rangle = |s\rangle - 2|w\rangle$$

Após a aplicação do oráculo U_w ao estado $|s\rangle$ do registrador quântico, o estado desejado $|w\rangle$ foi marcado em $|\psi_1\rangle$ como sendo o único que teve sua amplitude alterada de $\frac{1}{\sqrt{N}}$ para $-\frac{1}{\sqrt{N}}$. Geometricamente,

isso significa que U_w refletiu $|s\rangle$ em relação a $|s'\rangle$. Contudo, esta informação não está disponível classicamente, mas restrita ao nível quântico, uma vez que a probabilidade de obtenção do estado $|w\rangle$ após uma medição do registrador é igual a $\left|-\frac{1}{\sqrt{N}}\right|^2 = \frac{1}{N}$, ou seja: a mesma dos demais estados da superposição. O que se deseja, agora, é refletir $|\psi_1\rangle$ em relação a $|s\rangle$ e, com isso, aumentar a amplitude do estado $|w\rangle$ no registrador quântico, tornando mais provável a obtenção de $|w\rangle$ após uma medição. Essa reflexão é promovida pelo operador $U_s = 2|s\rangle\langle s| - I$, onde I é o operador identidade (PORTUGAL et al, 2012, p.48). Portanto, a próxima etapa do algoritmo é a aplicação de U_s sobre $|\psi_1\rangle$, gerando $|\psi_2\rangle$.

$$|\psi_2\rangle = U_s|\psi_1\rangle = (2\langle s|s\rangle)|s\rangle - \left(\frac{4}{\sqrt{N}}\langle s|w\rangle\right)|s\rangle - |s\rangle + \frac{2}{\sqrt{N}}|w\rangle.$$

Uma vez que $\langle s|w\rangle = \frac{1}{\sqrt{N}}$, segue que

$$|\psi_2\rangle = \frac{N-4}{N}|s\rangle + \frac{2}{\sqrt{N}}|w\rangle.$$

Assim, aplicando-se o operador de Grover $G = U_s U_w$ sobre $|s\rangle$, a amplitude de $|w\rangle$ é dada por

$$\left(\frac{N-4}{N}\right)\left(\frac{1}{\sqrt{N}}\right) + \frac{2}{\sqrt{N}} = \frac{3N-4}{N\sqrt{N}}.$$

Com isso, a nova probabilidade de se obter o estado $|w\rangle$ após uma medição do registrador quântico é igual a

$$\left|\frac{3N-4}{N\sqrt{N}}\right|^2 = \frac{9N^2 + 24N + 16}{N^3}.$$

Um exemplo que demonstra a eficiência do algoritmo de Grover é o caso em que $N = 4$: antes da aplicação do operador de Grover sobre $|s\rangle$, a probabilidade de se obter $|w\rangle$ em uma medição é de 25%, subindo para 100% com uma única aplicação de G . Naturalmente, essa probabilidade ainda é pequena para valores maiores de N ; contudo, o operador G pode ser aplicado mais de uma vez. No caso em que $N = 512$, por exemplo, uma única aplicação do operador de Grover resulta em uma probabilidade de aproximadamente $0,017 = 1,7\%$ de se obter o estado desejado $|w\rangle$, sendo necessária a reaplicação do operador de Grover para ampliar essa probabilidade. A Figura 1 mostra uma visualização do circuito lógico correspondente ao algoritmo de Grover, onde $U_f = U_w$.

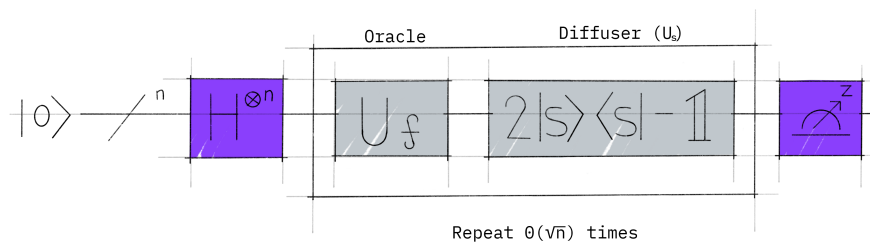


Figura 1: Circuito lógico do algoritmo de Grover (IBM, 2023).

3 Implementação

Uma vez compreendido matematicamente o funcionamento do algoritmo de Grover, pode-se seguir com sua implementação prática. Isso é feito através das linhas de código abaixo, escritas no Jupyter Notebook, uma ferramenta open-source baseada em browser, que cumpre o papel de um laboratório virtual no qual se pode programar em diferentes linguagens de programação, como Python 3 (RANDLES et al, 2017). O exemplo escolhido neste trabalho é o caso em que $N = 4$ e $w = 3$. Portanto, o estado que se deseja obter é $|3\rangle = |11\rangle$. Conforme visto anteriormente, como $N = 4$, espera-se que a probabilidade de obtenção do elemento desejado $|11\rangle$ após uma única aplicação de G seja igual a 100%.

3.1 Definindo os operadores

Uma vez que $N = 4$ e $w = 3$, o oráculo U_w é representado pela seguinte matriz unitária:

$$\begin{aligned} U_w &= \begin{bmatrix} (-1)^{f(0)} & 0 & 0 & 0 \\ 0 & (-1)^{f(1)} & 0 & 0 \\ 0 & 0 & (-1)^{f(2)} & 0 \\ 0 & 0 & 0 & (-1)^{f(3)} \end{bmatrix} \\ &= \begin{bmatrix} (-1)^0 & 0 & 0 & 0 \\ 0 & (-1)^0 & 0 & 0 \\ 0 & 0 & (-1)^0 & 0 \\ 0 & 0 & 0 & (-1)^1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = CZ \end{aligned}$$

Neste caso, U_w coincide com a porta CZ . Assim sendo, a porta controlada Z exercerá a função de oráculo no circuito. Por outro lado, o operador de difusão, que reflete o estado $|\psi_1\rangle$ em relação a $|s'\rangle$, pode ser matematicamente descrito pela expressão $U_s = H^{\otimes 2}CZZ^{\otimes 2}H^{\otimes 2}$ (IBM 2023). Dessa forma, o circuito quântico para a implementação do algoritmo de Grover é dado pela Figura 2.

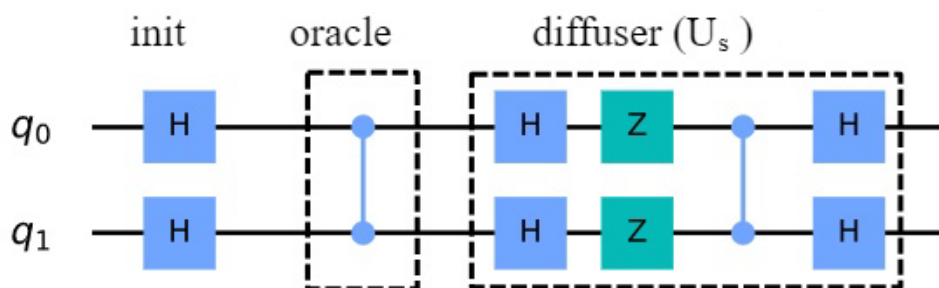


Figura 2: Circuito completo do algoritmo de Grover para o caso em que $|w\rangle = |11\rangle$ (IBM, 2023).

3.2 O código

Nesta etapa, será seguido diretamente o exemplo fornecido pela IBM (IBM, 2023). Antes da implementação do algoritmo, é necessário importar os recursos que serão utilizados no código, conforme a Figura 3.

```
In [15]: import matplotlib.pyplot as plt
import numpy as np
import math

from qiskit import IBMQ, Aer, transpile, execute
from qiskit import QuantumCircuit, ClassicalRegister, QuantumRegister
from qiskit.providers.ibmq import least_busy

from qiskit.visualization import plot_histogram
```

Figura 3: Importações necessárias para a implementação do algoritmo de Grover (IBM, 2023).

Como $N = 4 = 2^2$, o circuito quântico neste caso terá dois qubits. Criado o circuito, a porta Hadamard é aplicada a ambos os qubits, inicializando com isso o estado de superposição $|s\rangle = \frac{1}{2} \sum_{x=0}^3 |x\rangle$. Esse procedimento computacional é realizado segundo as linhas de código da Figura 4.

```
In [16]: n = 2
grover_circuit = QuantumCircuit(n)

In [17]: def initialize_s(qc, qubits):
        """Apply a H-gate to 'qubits' in qc"""
        for q in qubits:
            qc.h(q)
        return qc

In [18]: grover_circuit = initialize_s(grover_circuit, [0,1])
grover_circuit.draw(output='mpl')
```

Out[18]:



Figura 4: Inicialização do estado $|s\rangle$ (IBM, 2023).

A Figura 5 mostra o procedimento para a aplicação do oráculo U_w , representado pela porta controlada Z. Conforme visto anteriormente, sabe-se que ele altera o sinal da amplitude do elemento procurado $|11\rangle$, refletindo $|s\rangle$ em relação a $|s'\rangle$.

```
In [19]: grover_circuit.cz(0,1)
grover_circuit.draw(output='mpl')
```

Out[19]:

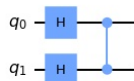


Figura 5: Aplicação do oráculo U_w ao circuito (IBM, 2023).

Por fim, é aplicado o operador de difusão $U_s = H^{\otimes 2} CZZ^{\otimes 2} H^{\otimes 2}$, que reflete $|\psi_1\rangle$ em relação a $|s\rangle$, conforme a Figura 6.

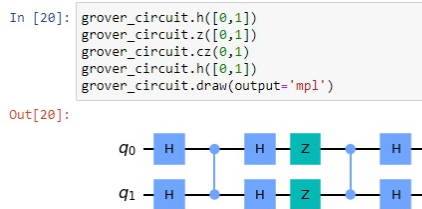


Figura 6: Aplicação do operador de difusão U_s , finalizando o circuito (IBM, 2023).

4 Testando o algoritmo

O algoritmo implementado será agora simulado. Primeiramente, é utilizado o *statevector_simulator* para se obter o estado final após a execução do algoritmo, que é igual a $|\psi\rangle = |11\rangle$, conforme a Figura 7.

```

In [21]: sv_sim = Aer.get_backend('statevector_simulator')
          result = sv_sim.run(grover_circuit).result()
          statevec = result.get_statevector()
          from qiskit.visualization import array_to_latex
          array_to_latex(statevec, prefix="|\psi\rangle = ")

Out[21]: |\psi\rangle = [0  0  0  1]

```

Figura 7: O vetor $|\psi\rangle$ é impresso (IBM, 2023).

A Figura 8 mostra a realização de 1024 medições dos qubits do registrador, resultando no histograma abaixo. Em todas as medições, o único estado obtido foi $|11\rangle$, confirmando a predição do algoritmo.

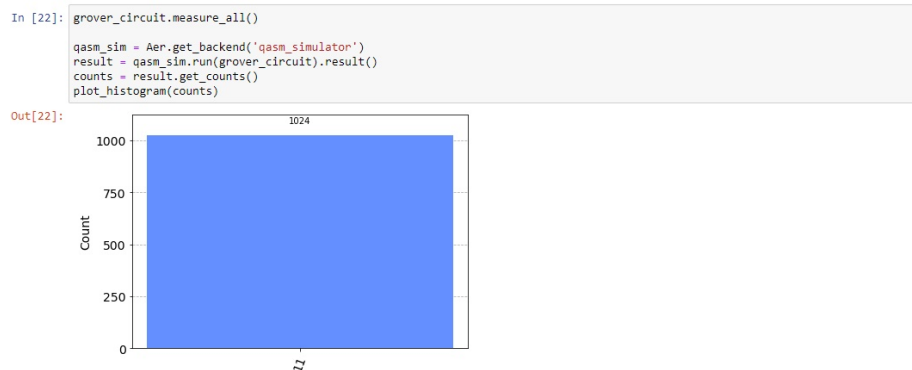


Figura 8: Simulação do algoritmo de Grover (IBM, 2023).

5 Conclusões

O potencial dos computadores quânticos para resolver certos problemas de forma mais eficiente do que as abordagens clássicas é evidenciado pelo algoritmo de Grover, que permite a busca em uma lista não ordenada de N elementos com uma complexidade de tempo proporcional a $\sqrt{N}$, ao contrário da busca clássica, que requer um tempo proporcional a N para a mesma tarefa. No entanto, o êxito do algoritmo de Grover está intrinsecamente ligado à sua habilidade de implementar um oráculo quântico eficaz, o qual desempenha a função de refletir o estado em relação ao valor alvo. No caso abordado por este trabalho,

a porta *CZ* desempenhou o papel de oráculo, alterando o sinal do ultimo estado da superposição. O escopo de aplicação do algoritmo de Grover é diversificado e abarca campos como otimização (GILLIAN, 2020), criptografia (SAKHI, 2012) e simulação quântica (CHUANG e NIELSEN, 2002, p.255-261). Entretanto, é imperativo compreender que nem todos os problemas experimentam uma melhoria substancial por meio da utilização do algoritmo de Grover, e a busca por algoritmos quânticos eficazes voltados a problemas específicos permanece como um domínio ativamente pesquisado (FLUHRER, 2017). Adicionalmente, vale destacar que, apesar de seu poder em teoria, o algoritmo de Grover em seu estágio atual ainda não supera de forma definitiva suas contrapartes clássicas em muitos problemas. Não obstante, implementações experimentais bem-sucedidas do algoritmo em computadores quânticos reais já foram realizadas, proporcionando indícios promissores sobre sua viabilidade prática (WANG, 2021).

Referências

- FLUHRER, S. Reassessing Grover's Algorithm. (2017). <https://eprint.iacr.org/2017/811>.
- GILLIAM, A., PISTOIA, M., GONCIULEA, C. Optimizing Quantum Search Using a Generalized Version of Grover's Algorithm. (2020). ArXiv. <https://arxiv.org/pdf/2005.06468.pdf>.
- GROVER, L. A fast quantum mechanical algorithm for database search. **IN: Proceedings of the 28th Annual ACM Symposium on the Theory of Computing**, 28., 1996. Philadelphia, 1996. p. 212-219.
- IBM. Grover's Algorithm. Disponível em: <https://learn.qiskit.org/course/ch-algorithms/grovers-algorithm>. Acesso em: 12 ag. 2023.
- NIELSEN, M.A., CHUANG, I. **Quantum Computation and Quantum Information**. Cambridge: Cambridge University Press, 2002.
- PORTUGAL, R. et al. **UMA INTRODUÇÃO À COMPUTAÇÃO QUÂNTICA**. 2.ed. São Paulo: Sociedade Brasileira de Matemática Aplicada e Computacional, 2012.
- RANDLER, B. et al. Using the Jupyter Notebook as a tool for open science: An empirical study. In: ACM/IEEE Joint Conference on Digital Libraries, 2017, Toronto. Toronto: Institute of Electrical and Electronics Engineers (IEEE), 2017. p. 1-2.
- SAKHI, Z. et al. Quantum cryptography based on Grover's algorithm. **Second International Conference on the Innovative Computing Technology (INTECH 2012)**. Casablanca, Marrocos. 2012. p. 33-37.
- WANG, R. Comparing Grover's Quantum Search Algorithm with Classical Algorithm on Solving Satisfiability Problem. 2021 IEEE Integrated STEM Education Conference (ISEC). Conferência realizada de forma virtual devido às restrições da pandemia de COVID-19. 2021. p. 204.

UMA FORMULAÇÃO MULTIESCALA NÃO LINEAR E DESCONTÍNUA PARA PROBLEMAS DE CONVECÇÃO-DIFUSÃO-REAÇÃO

Enéas Mendes de Jesus

Universidade Federal do Espírito Santo (UFES) - PPGI

Instituto Federal do Espírito Santo (IFES) - Campus Piúma

eneas.jesus@ufes.edu.br

Isaac Pinheiro dos Santos

Universidade Federal do Espírito Santo (UFES)

isaac.santos@ufes.br

Resumo: Este trabalho apresenta uma formulação multiescala não linear e descontínua para resolver problemas de convecção-difusão-reação. Considerando uma formulação de Galerkin descontínua, o método introduz um termo de difusão artificial não linear em ambas as escalas de discretização, enquanto utiliza a abordagem descontínua somente na macro escala. A micro escala é modelada através de funções bolha, permitindo a aplicação do processo de condensação estática em cada elemento. A discretização resulta em um sistema global de equações associado aos pontos nodais apenas da macro escala. Para avaliar a estabilidade das soluções numéricas obtidas pelo esquema proposto, assim como sua convergência, alguns experimentos numéricos são realizados.

Palavras-chave: Galerkin descontínuo. Difusão artificial. Convecção-difusão-reação. Funções bolha.

Introdução

O método Galerkin descontínuo (GD) foi introduzido por Redd e Hill em 1973 (REED; HILL, 1973). Inicialmente, a utilização do método esbarrou no seu alto custo computacional, em comparação com o método Galerkin contínuo, decorrente do número elevado de graus de liberdade presentes na formulação. Dessa forma, muitas metodologias foram propostas na literatura para reduzir o custo computacional do GD, como por exemplo, o “multiscale discontinuous Galerkin” (MDG) (HUGHES, et al, 2006) e o “hybridizable discontinuous Galerkin” (HDG) (NGUYEN; PERAIRE; COCKBURN, 2009). Diante da possibilidade de se obter um método com custo computacional atrativo e com propriedades de estabilidade superiores as do método Galerkin contínuo para problemas de convecção dominante, muitas pesquisas foram direcionadas aos métodos GD.

Dentre as várias abordagens desenvolvidas no contexto do método GD para problemas com convecção dominante, destacamos o método de Difusão Dinâmica (DD) (ARRUDA; ALMEIDA; CARMO, 2010a e 2010b). Essa metodologia considera uma decomposição do espaço de aproximação e do campo de velocidade em escalas resolvidas (macro) e não-resolvidas (micro). Um operador não linear de difusão artificial é adicionado à formulação numérica. A quantidade de difusão artificial é determinada pela solução da equação na escala resolvida à nível de elemento, conduzindo a um método auto-adaptativo e livre de parâmetros de estabilização. Além disso, são necessárias duas malhas do domínio do problema: uma malha representando o espaço macro e uma malha mais refinada, onde o problema é resolvido. Isso resulta em uma formulação com um custo computacional elevado, em comparação com o método GD clássico. Mais recentemente, o método DD foi proposto no contexto contínuo usando funções bolhas na micro escala (VALLI, et al, 2018, SANTOS, et al. 2021), resultando em um método com boas propriedades de convergência e estabilidade, além de ter um custo computacional da mesma ordem da formulação clássica de Galerkin contínuo.

A formulação apresentada neste trabalho busca solucionar problemas de convecção-difusão-reação a partir de uma abordagem descontínua, multiescala e não linear. Esta é uma versão descontínua do

método DD apresentado em (VALLI, et al., 2018). Além disso, essa formulação é diferente da abordagem apresentada por (ARRUDA; ALMEIDA; CARMO, 2010a e 2010b), pois utiliza funções bolhas na micro escala e requer apenas uma malha, associada à macro escala, para resolver o problema. O método tem como ponto de partida uma formulação GD apresentada em (HOUSTON; SCHWAB; SULI, 2002), que é a mesma utilizada por (ARRUDA; ALMEIDA; CARMO, 2010b).

O presente trabalho está organizado da seguinte forma: apresentamos a equação de convecção-difusão-reação, os conceitos e notações utilizados ao longo do texto, assim como a formulação GD. Em seguida, o novo método é proposto. Por fim, alguns experimentos numéricos são apresentados.

Método Galerkin Descontínuo

Seja $\Omega \in \mathbb{R}^2$ um domínio aberto limitado com fronteira de Lipschitz, Γ . O problema de convecção-difusão-reação com condição de contorno de Dirichlet é dado por

$$\begin{aligned} -\kappa \Delta u + \beta \cdot \nabla u + \sigma u &= f \text{ em } \Omega, \\ u &= g \text{ em } \Gamma, \end{aligned} \quad (1)$$

onde $\kappa > 0$, $\beta \in [W^{1,\infty}(\Omega)]^2$ com $\nabla \cdot \beta = 0$, $\sigma \in L^\infty(\Omega)$, $f \in L^2(\Omega)$ e $g \in H^{1/2}(\Gamma)$. A fronteira de entrada é dada por

$$\Gamma_- = \{x \in \Gamma; \beta(x) \cdot \mathbf{n}(x) < 0\}, \quad (2)$$

onde $\mathbf{n}(x)$ representa o vetor unitário normal na direção exterior a Γ no ponto $x \in \Gamma$.

Considerando uma partição triangular regular do domínio Ω , denotada por $\mathcal{T}_h = \{K\}$, onde h é o parâmetro de malha, define-se o espaço quebrado de elementos finitos descontínuos da seguinte forma,

$$V_h = \{v \in L^2(\Omega) : v|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}, \quad (3)$$

onde $\mathbb{P}_1(K)$ denota o espaço de polinômios lineares sobre K .

A fim de obter a formulação Galerkin Descontínua (GD) é necessário introduzir alguns conceitos. Denotamos por $\mathcal{E}_h$ o conjunto de todas as arestas em $\mathcal{T}_h$. Os conjuntos $\mathcal{E}_h^0$ e $\mathcal{E}_h^\Gamma$ contém as arestas do interior e da fronteira, respectivamente. Dada uma função escalar φ (vetorial τ), suave por partes em $\mathcal{T}_h$, definimos o salto e a média, respectivamente, sobre uma aresta $e \in \mathcal{E}_h^0$ da seguinte forma,

$$[\![\varphi]\!] = \varphi|_{K_1} \mathbf{n}^1 + \varphi|_{K_2} \mathbf{n}^2 \quad ([\![\tau]\!] = \tau|_{K_1} \cdot \mathbf{n}^1 + \tau|_{K_2} \cdot \mathbf{n}^2), \quad \{\varphi\} = \frac{\varphi|_{K_1} + \varphi|_{K_2}}{2} \quad \left(\{\tau\} = \frac{\tau|_{K_1} + \tau|_{K_2}}{2} \right), \quad (4)$$

onde K_i são os elementos que compartilham a aresta e e $\mathbf{n}^i$ é o vetor unitário normal a aresta e apontando para o exterior do elemento K_i . Além disso, se $e \in \mathcal{E}_h^\Gamma$ as definições de salto e média coincidem com os traços sobre e das respectivas funções.

O método GD para o problema descrito na Equação (1), definido em (HOUSTON; SCHWAB; SULI, 2002), consiste em achar $u_h \in V_h$ tal que

$$B_{GD}(u_h, v_h) = F_{GD}(v_h), \quad \forall v_h \in V_h, \quad (5)$$

onde

$$\begin{aligned} B_{GD}(u_h, v_h) &= \sum_{K \in \mathcal{T}_h} \int_K (\kappa \nabla u_h \cdot \nabla v_h + \beta \cdot \nabla u_h + \sigma u_h v_h) dx \\ &\quad - \kappa \sum_{e \in \mathcal{E}_h^0} \int_e \{\nabla u_h\} [\![v_h]\!] ds + \varepsilon_0 \kappa \sum_{e \in \mathcal{E}_h^0} \int_e [\![u_h]\!] \{\nabla v_h\} ds + \kappa \eta_0 h^{-1} \sum_{e \in \mathcal{E}_h^0} \int_e [\![u_h]\!] [\![v_h]\!] ds \\ &\quad - \sum_{e \in \mathcal{E}_h^{0-}} \int_e (\beta \cdot [\![u_h]\!]) v_h ds - \sum_{e \in \mathcal{E}_h^{\Gamma-}} \int_e (\beta \cdot \mathbf{n}) u_h v_h ds \\ &\quad - \kappa \sum_{e \in \mathcal{E}_h^{\Gamma}} \int_e (\nabla u_h \cdot \mathbf{n}) v_h ds + \varepsilon_0 \kappa \sum_{e \in \mathcal{E}_h^{\Gamma}} \int_e u_h (\nabla v_h \cdot \mathbf{n}) ds + \kappa \eta_\Gamma h^{-1} \sum_{e \in \mathcal{E}_h^{\Gamma}} \int_e u_h v_h ds, \end{aligned} \quad (6)$$

$$F_{GD}(v_h) = \sum_{K \in \mathcal{T}_h} \int_K f v_h dx + \varepsilon_0 \kappa \sum_{e \in \mathcal{E}_h^\Gamma} \int_e g(\nabla v_h \cdot \mathbf{n}) ds + \kappa \eta_\Gamma h^{-1} \sum_{e \in \mathcal{E}_h^\Gamma} \int_e g v_h ds - \sum_{e \in \mathcal{E}_h^{\Gamma^-}} \int_e (\beta \cdot \mathbf{n}) g v_h ds, \quad (7)$$

onde $\varepsilon_0 = \{-1, 0, 1\}$, η_0 e η_Γ são os parâmetros de penalização das arestas do interior e da fronteira, respectivamente.

Método Difusão Dinâmica Descontínua com Funções Bolhas

Esta formulação multiescala considera a metodologia descontínua aplicada apenas na escala macro da discretização. A escala micro ou fina é discretizada usando funções bolha. Além disso, um operador de difusão artificial não linear é adicionado a ambas as escalas. O método tem origem nos trabalhos apresentados em (ARRUDA et al., 2010) e (VALLI et al., 2018).

O espaço definido na Equação (3) é enriquecido com o espaço de funções bolhas, definido por

$$V_b = \{w \in H^1(\Omega) : w|_K \in H_0^1(K), \forall K \in \mathcal{T}_h\}. \quad (8)$$

Um função bolha $\varphi_b \in V_b$ possui suporte contido em um único elemento, é nula na fronteira e assume valor 1 no baricentro do elemento. Neste trabalho utilizamos a função bolha da forma $\varphi_b = 27N_1^K N_2^K N_3^K$, onde $N_j^K = N_j^K(x, y)$ são as funções bases do espaço V_h . O espaço enriquecido é representado pela soma direta

$$V_E = V_h \oplus V_b. \quad (9)$$

Portanto, a formulação Difusão Dinâmica Descontínua com funções Bolhas (DDB) consiste em achar $u_E = u_h + u_b \in V_E$, com $u_h \in V_h$ e $u_b \in V_b$, tal que

$$B_{GD}(u_E, v_E) + D(u_h; u_E, v_E) = F_{GD}(v_E), \quad \forall v_E \in V_E, \quad (10)$$

onde $v_E = v_h + v_b$, com $v_h \in V_h$ e $v_b \in V_b$. O operador não linear de difusão artificial $D(\cdot; \cdot, \cdot)$ é definido como

$$D(u_h; u_E, v_E) = \sum_{K \in \mathcal{T}_h} \int_K \xi(u_h) \nabla u_E \cdot \nabla v_E \, dx, \quad (11)$$

onde

$$\xi(u_h) = \begin{cases} \frac{\hbar}{2} \frac{|R(u_h)|}{\|\nabla u_h\|}, & \text{se } \|\nabla u_h\| > \text{tol}_\xi, \\ 0, & \text{caso contrário.} \end{cases} \quad (12)$$

Na Equação (12), $\hbar = c_0 \sqrt{|K|}$ onde c_0 é uma constante, $|K|$ é a área do elemento K , $R(u_h) = -\kappa \Delta u_h + \beta \cdot \nabla u_h + \sigma u_h - f$ e tol_ξ é um número positivo pequeno suficiente para evitar divisão por zero. Neste trabalho utilizamos $\text{tol}_\xi = 10^{-5}$.

Dada a não-linearidade do operador (11), utilizamos um processo iterativo para solução do método DDB, definido por: dada solução u_h^n , achar u_h^{n+1} satisfazendo

$$B_{GD}(u_E^{n+1}, v_E) + D(u_h^n; u_E^{n+1}, v_E) = F_{GD}(v_E), \quad \forall v_E \in V_E. \quad (13)$$

Substituindo $u_E^{n+1} = u_h^{n+1} + u_b^{n+1}$ na Equação (13) e considerando separadamente $v_E^{n+1} = v_h^{n+1}$ e $v_E^{n+1} = v_b^{n+1}$, obtemos dois subproblemas, um macro dado por

$$B_{GD}(u_h^{n+1}, v_h) + B_{GD}(u_b^{n+1}, v_h) + \sum_{K \in \mathcal{T}_h} \int_K \xi(u_h^n) \nabla u_h^{n+1} \cdot \nabla v_h \, dx = F_{GD}(v_h), \quad \forall v_h \in V_h \quad (14)$$

e um problema micro, descrito pela equação,

$$B_{GD}(u_h^{n+1}, v_b) + B_{GD}(u_b^{n+1}, v_b) + \sum_{K \in \mathcal{T}_h} \int_K \xi(u_h^n) \nabla u_b^{n+1} \cdot \nabla v_b \, dx = F_{GD}(v_b), \quad \forall v_b \in V_b, \quad (15)$$

uma vez que

$$\int_K \xi(u_h^n) \nabla u_h^{n+1} \cdot \nabla v_b \, dx = \int_K \xi(u_h^n) \nabla u_b^{n+1} \cdot \nabla v_h \, dx = 0.$$

Os problemas (14) e (15) resultam em cada elemento $K \in \mathcal{T}_h$ no seguinte sistema local,

$$\begin{bmatrix} A_{hh} & A_{hb} \\ A_{bh} & A_{bb} \end{bmatrix} \begin{bmatrix} U_h \\ U_b \end{bmatrix} = \begin{bmatrix} F_{h,K} \\ F_{b,K} \end{bmatrix}, \quad (16)$$

onde A_{hh} é a matriz local associada ao termo $B_{GD}(u_h^{n+1}, v_h)|_K + \int_K \xi(u_h^n) \nabla u_h^{n+1} \cdot \nabla v_h \, dx$; A_{hb} é a matriz local associada ao termo $B_{GD}(u_b^{n+1}, v_h)|_K$; A_{bh} é a matriz local associada ao termo $B_{GD}(u_h^{n+1}, v_b)|_K$; A_{bb} é a matriz local associada ao termo $B_{GD}(u_b^{n+1}, v_b)|_K + \int_K \xi(u_h^n) \nabla u_b^{n+1} \cdot \nabla v_b \, dx$; $F_{h,K}$ é o vetor local associada ao termo $F_{GD}(v_h)|_K$ e $F_{b,K}$ é o vetor local associada ao termo $F_{GD}(v_b)|_K$. Os vetores U_h e U_b representam os valores das soluções u_h e u_b nos pontos nodais da macro e da micro escala, respectivamente, em cada elemento K .

Como o suporte da função bolha está contido no elemento K , a solução U_b não depende de outros elementos, sendo assim, utilizamos o processo de condensação estática para eliminar a variável U_b na Equação (16), resultando no sistema local reduzido,

$$A_K U_h = F_K, \quad (17)$$

onde $A_K = A_{hh} - A_{hb}(A_{bb})^{-1}A_{bh}$ e $F_K = F_{h,K} - A_{hb}(A_{bb})^{-1}F_{b,K}$. As matrizes e os vetores locais são utilizados para obter o sistema global que fornecerá uma solução aproximada para o problema (13).

Para melhorar a convergência do processo iterativo não linear, utilizamos o esquema de relaxação apresentado em (SANTOS; ALMEIDA, 2007):

$$\begin{aligned} \xi^* &\leftarrow \omega \xi(u_h^n) + (1 - \omega) \xi(u_h^{n-1}), \\ \xi(u_h^n) &\leftarrow \xi^*, \end{aligned}$$

onde o parâmetro ω , neste trabalho, é definido por

$$\omega = \begin{cases} 0, & \text{se } ||R(u_h^{n+1})| - |R(u_h^n)|| < C_R, \\ 1/2, & \text{caso contrário,} \end{cases}$$

com $C_R = 0, 2$.

Experimentos Numéricos

Nesta seção apresentamos os experimentos numéricos considerando um problema de convecção-difusão com camada limite, um problema de difusão-reação e um problema com solução suave para avaliarmos as taxas de convergência do método proposto nas normas $L^2(\Omega)$ e $H^1(\Omega)$. Os resultados numéricos são comparados com o método GD descrito na Equação (5). Vamos adotar $c_0 = \sqrt{2}$, de modo que $h = \sqrt{2|K|}$. Conforme discutido em (ARNOLD, et al., 2002), o parâmetro de penalização pode ser ajustado a fim de minimizar os saltos nas fronteiras dos elementos na solução, produzindo assim uma continuidade entre os elementos. Mais ainda, tal parâmetro pode fazer com que a condição de contorno de Dirichlet seja satisfeita fraca ou fortemente, a depender do valor assumido. Em geral, não é possível garantir simultaneamente a continuidade entre os elementos e a satisfação da condição de contorno Dirichlet fortemente. Diante disso, neste trabalho, consideramos separadamente parâmetros de penalização no interior, buscando a continuidade inter-elementar, e na fronteira, buscando a condição de contorno de Dirichlet satisfeita fortemente.

Problema de Convecção-Difusão

Nesse experimento vamos avaliar o desempenho do método DDB para um problema bidimensional de convecção-difusão com convecção dominante. Consideramos $\Omega = (0, 1) \times (0, 1)$ e estabelecemos os parâmetros $\kappa = 10^{-6}$ e $\beta = (1, 1)$. O termo de fonte é definido de modo que a solução exata seja

$$u(x, y) = \sin(\pi x/2) \sin(\pi y/2) (1 - e^{(x-1)/\kappa}) (1 - e^{(y-1)/\kappa}). \quad (18)$$

Neste caso, a condição de contorno é de Dirichlet homogênea. A Figura 1 apresenta os resultados

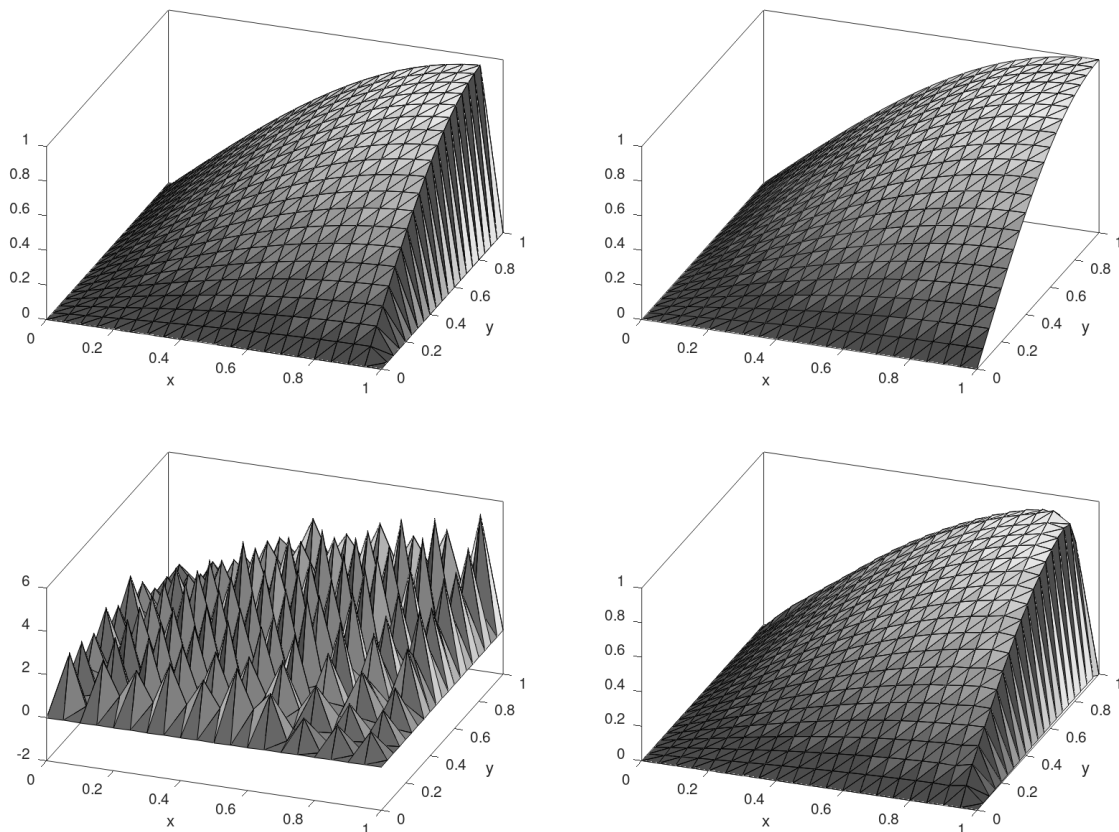


Figura 1: Solução analítica (superior esquerdo), GD com condição de Dirichlet imposta fracamente (superior direito), GD com condição de Dirichlet imposta fortemente (inferior esquerdo) e DDB (inferior direito).

dos métodos GD e DDB. Observa-se que ao impor fracamente a condição de contorno, a formulação GD apresenta uma boa aproximação para a solução¹. No entanto, quando as condições de contorno de Dirichlet são impostas fortemente, o método GD apresenta oscilações espúrias por todo o domínio. Neste caso, o método DDB apresenta uma solução livre de oscilações numéricas.

Problema de Difusão-Reação

Nesse experimento vamos avaliar o desempenho do método DDB para um problema bidimensional reação-difusão, onde $\Omega = (0, 1) \times (0, 1)$, $\kappa = 3 \times 10^{-5}$, $\sigma = 1$ e $f = 1$. Além disso, a condição de contorno é do tipo Dirichlet homogênea.

A Figura 2 apresenta os resultados dos métodos GD e DDB. Neste caso, devido o problema ser reação dominante, as oscilações restringem-se às regiões próximas às camadas limites. O método GD

¹O método DDB apresenta uma solução semelhante a solução GD.

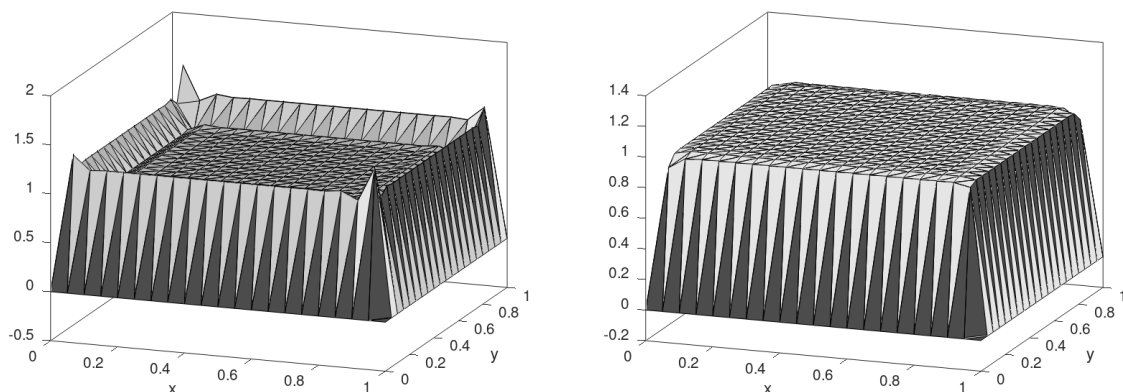


Figura 2: Solução obtida com o método GD (à esquerda) e com o método DDB (à direita).

não é capaz de eliminar essas instabilidades numéricas, enquanto que o método DDB apresenta uma solução sem oscilações espúrias.

Taxas de convergência

Neste exemplo vamos avaliar as taxas de convergência do método DDB nas normas $L^2(\Omega)$ e $H^1(\Omega)$. Para isso, consideramos um problema bidimensional de convecção-difusão-reação no domínio $\Omega = (0, 1) \times (0, 1)$, onde $\kappa = 10^{-6}$, $\beta = (1, 0)^T$ e $\sigma = 1$. O termo de fonte f e as condições de contorno de Dirichlet são definidas de modo que a solução exata (suave) do problema descrito na Equação (1) seja

$$u(x, y) = \sin(\pi x) \cos(\pi x).$$

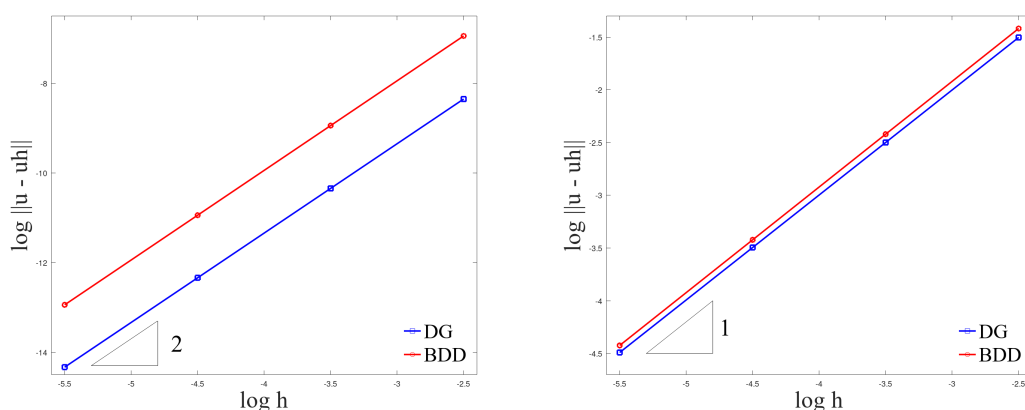


Figura 3: Taxas de convergência dos métodos DG e BDD nas norma de $L^2(\Omega)$ à esquerda e de $H^1(\Omega)$ à direita.

Taxas ótimas de convergência são obtidos para ambos os métodos, GD e DDB, isto é, taxa quadrática na norma de $L^2(\Omega)$ e linear na norma de $H^1(\Omega)$, conforme Figura 3. O método DDB apresenta erros na norma de $L^2(\Omega)$ levemente maiores do que os erros obtidos pelo método GD.

Conclusões

Apresentamos um método multiescala descontínuo e não linear para solução numérica de problemas de reação-difusão-convecção. Um operador de difusão artificial não linear é adicionado a uma formulação GD descrita em um cenário de discretização de duas escalas. O método proposto é eficiente em eliminar as oscilações espúrias que aparecem próximas às camadas limites de problemas com convecção/reação dominantes, conforme apresentado nos experimentos numéricos. Além disso, o novo esquema numérico apresentou taxas de convergência ótimas nas normas $L^2(\Omega)$ e $H^1(\Omega)$.

O método desenvolvido oferece uma abordagem promissora para resolver problemas de transporte com convecção ou reação dominante. Como trabalhos futuros, destacamos o ajuste correto dos parâmetros h e C_R , relacionadas com a difusão artificial e o processo iterativo, respectivamente, a inserção de algum tipo de estabilização nas arestas dos elementos, desenvolvimento de estratégias para redução do custo computacional inerente de formulações descontínuas, descrição de estratégias para aceleração de convergência do processo iterativo não linear, análise numérica do método e extensão para problemas transientes.

Referências

- ARNOLD, D. N. et al. Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. **SIAM Journal on Numerical Analysis**, vol. 39, n. 5, pp. 1749–1779, 2002.
- ARRUDA, N. C. B.; ALMEIDA, R. C.; DO CARMO, E. G. D. Discontinuous subgrid formulations for transport problems. **Computer Methods in Applied Mechanics and Engineering**, vol. 199, n. 49, pp. 3227–3236, 2010a.
- ARRUDA, N. C. B.; ALMEIDA, R. C.; DO CARMO, E. G. D. Dynamic diffusion formulation for advection dominated transport problems. **Mecânica computacional**, vol. 29, n. 20, p. 2011–2025, 2010b.
- HOUSTON, P.; C. SCHWAB, C.; SULI, E. Discontinuous hp-finite element methods for advection-diffusion-reaction problems. **SIAM Journal on Numerical Analysis**, v. 39, n. 6, p. 2133–2163, 2002.
- HUGHES, T. J. R. et al. A multiscale discontinuous Galerkin method with the computational structure of a continuous Galerkin method. **Comput. Methods Appl. Mech. Engrg.** 195 (2006) 2761–2787
- JOHN, V.; KNOBLOCH, P. On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations: Part II - Analysis for P1 and Q1 finite elements, **Comput. Methods Appl. Mech. Engrg.** 197 (21–24) (2008) 1997–2014.
- NGUYEN, N.C.; PERAIRE, J.; COCKBURN B. An implicit high-order hybridizable discontinuous Galerkin method for linear convection–diffusion equations, **J. Comput. Phys.** 228 (2009) 3232–3254.
- REED, W.H.; HILL, T.R. Triangular mesh methods for the neutron transport equation, **Tech. Report LA-UR-73-479**, Los Alamos Scientific Laboratory, 1973.
- SANTOS, I.P.; ALMEIDA, R.C. A nonlinear subgrid method for advection-diffusion problems, **Comput. Methods Appl. Mech. Engrg.** 196 (2007) 4771–4778.
- SANTOS, I. P. et al. Convergence analysis of a new dynamic diffusion method. **Computers and Mathematics with Applications**, vol. 98, pp. 1–9, 2021.
- VALLI, A. M. et al. A parameter-free dynamic diffusion method for advection–diffusion–reaction problems. **Computers and Mathematics with Applications**, vol. 75, n. 1, p. 307–321, 2018.

DISCRETIZAÇÃO DE UM MODELO PRESA-PREDADOR COM COMPETIÇÃO: DISPERSÃO-MIGRAÇÃO

Vitor Hugo Muniz Oliveira
IMECC - UNICAMP
v264980@dac.unicamp.br

Daniela Ribeiro Monteiro
Centro Universitário IBMEC - RJ
danielarmonteiro@gmail.com

João Frederico da Costa Azevedo Meyer
IMECC - UNICAMP
jmeyer@unicamp.br

Resumo: O trabalho desenvolvido tem como objetivo principal apresentar um sistema de Equações Diferenciais Parciais de Dispersão-Migração das densidades populacionais de P_1 e P_2 que interagem entre si. O sistema é discretizado espacialmente através do Método de Galerkin via Elementos Finitos e temporalmente através do Método de Crank-Nicolson no mesmo meio. São aqui apresentadas as equações não-lineares do sistema empregado na modelagem e suas formulações variacionais e a transformação num sistema não-linear discreto com o qual serão realizadas simulações numéricas que não constam do presente trabalho.

Palavras-chave: Equações Diferenciais Parciais. Método de Galerkin. Método de Crank-Nicolson. Formulação Variacional. Discretização Espacial. Discretização Temporal.

Modelo Presa-Predador com competição

Há um longo período em que os modelos que se concentram em uma única espécie têm sido contestados, com base na ideia de que nenhuma espécie opera de maneira isolada, sem interagir com outras. Desde os estudos realizados por LOTKA (1925) e VOLTERRA (1926), sistemas de Equações Diferenciais Ordinárias que abarcam as interações entre diferentes espécies começaram a ser empregados de várias maneiras para investigar e compreender diversos fenômenos ligados à sua coexistência. No entanto, esses modelos somente levavam em consideração as mudanças ao longo do tempo, assumindo uma distribuição uniforme das populações no espaço. Os processos de dispersão e migração, presentes em praticamente todas as populações naturais, levaram à necessidade de incorporar elementos espaciais nos modelos, sendo SKELLAM (1951) um dos pioneiros nesse aspecto.

Neste trabalho apresentaremos o estudo de um sistema de Equações Diferenciais Parciais de Dispersão-Migração envolvendo termos do tipo Lotka-Volterra e dinâmicas vitais do tipo Verhulst (MURRAY, 2003). Pode-se encontrar os primeiros estudos sobre esta ideia em SOSSAE (2003), que envolve um sistema que descreve um convívio com relações do tipo presa-predador com competição. Para a modelagem do sistema aqui apresentado, consideraremos a dispersão populacional de cada espécie, os processos migratórios de cada espécie, o decaimento das espécies devido à presença de um poluente, as dinâmicas vitais e as relações inter e intra-específicas.

A partir dos fenômenos descritos, iremos levar em conta duas populações P_1 e P_2 que interagem entre si. O sistema não-linear que descreve os fenômenos considerados para as densidades populacionais $P_1(x, y, t)$ e $P_2(x, y, t)$, com $(x, y) \in \Omega \subset \mathbb{R}^2$ e $t \in (0, T]$ é dado por:

$$\begin{cases} \frac{\partial P_1}{\partial t} - \operatorname{div}(\alpha_1 \nabla P_1) + \operatorname{div}(\vec{U} P_1) + \rho_1 \sigma P_1 = \beta_1 P_1 \left(1 - \frac{P_1 + P_2}{K}\right) - \delta_1 P_1 P_2, \\ \frac{\partial P_2}{\partial t} - \operatorname{div}(\alpha_2 \nabla P_2) + \operatorname{div}(\vec{W} P_2) + \rho_2 \sigma P_2 = \beta_2 P_2 \left(1 - \frac{P_1 + P_2}{K}\right) - \delta_2 P_1 P_2, \end{cases} \quad (1)$$

com os significados biológicos dos parâmetros descritos por, $\alpha_1 = \alpha_1(x, y, t)$ e $\alpha_2 = \alpha_2(x, y, t)$ sendo os coeficientes de dispersão populacional de cada espécie; $\vec{U} = (U_1(x, y, t), U_2(x, y, t))$ e $\vec{W} = (W_1(x, y, t), W_2(x, y, t))$ sendo os vetores velocidade de migração populacional ou advectiva de cada espécie, com $\nabla \cdot \vec{U} = \nabla \cdot \vec{W} = 0$; $\sigma = \sigma(x, y, t)$ sendo a função decaimento populacional das duas espécies no meio $\Omega \in (0, T]$; ρ_1 e ρ_2 sendo as taxas da função decaimento populacional de cada espécie devido à mortalidade causada pela presença do poluente σ ; β_1 e β_2 sendo as taxas de crescimento intrínsecos de cada espécie; K sendo a capacidade de suporte do meio para as duas espécies; δ_1 e δ_2 sendo as taxas de relação interespecífica.

Considera-se que $\vec{U}$, $\vec{W}$, α_1 e α_2 sejam constantes e, então, podemos reescrever o sistema 1 com as condições iniciais e condições de contorno de Dirichlet homogêneas em Γ_0 e Von-Neumann homogêneas em Γ_1 da seguinte forma:

$$\begin{cases} \begin{cases} \frac{\partial P_1}{\partial t} - \alpha_1 \Delta P_1 + \nabla \cdot (\vec{U} P_1) + \rho_1 \sigma P_1 = \beta_1 P_1 - \gamma_1 P_1^2 - \theta_1 P_1 P_2, \\ \frac{\partial P_2}{\partial t} - \alpha_2 \Delta P_2 + \nabla \cdot (\vec{W} P_2) + \rho_2 \sigma P_2 = \beta_2 P_2 - \gamma_2 P_2^2 - \theta_2 P_1 P_2, \end{cases} & \forall (x, y) \in \Omega \subset \mathbb{R}^2, \quad \forall t \in (0, T] \\ P_1(x, y, 0) = P_{1_0}(x, y), \quad P_2(x, y, 0) = P_{2_0}(x, y), & \forall (x, y) \in \Omega, \\ P_1|_{\Gamma_0} = P_2|_{\Gamma_0} = 0, \quad \frac{\partial P_1}{\partial \eta} \Big|_{\Gamma_1} = \frac{\partial P_2}{\partial \eta} \Big|_{\Gamma_1} = 0, & \forall t \in (0, T], \quad \text{com } \Gamma_0 \text{ e } \Gamma_1 \text{ tais que } \Gamma_0 \cup \Gamma_1 = \partial \Omega, \end{cases} \quad (2)$$

onde $\gamma_1 = \frac{\beta_1}{K}$, $\gamma_2 = \frac{\beta_2}{K}$, $\theta_1 = \gamma_1 + \delta_1$ e $\theta_2 = \gamma_2 + \delta_2$.

Os parâmetros θ_1 e θ_2 definidos acima são as taxas de relação interespecífica entre as espécies P_1 e P_2 . Assumindo θ_1 e θ_2 não nulos para que existam as interações, podemos considerar as seguintes situações: (i) se $\theta_1 < 0$ e $\theta_2 > 0$, o sistema é chamado presa-predador com o parâmetro com índice 1 representando a densidade populacional da presa e o índice 2 a do predador; (ii) se $\theta_1 < 0$ e $\theta_2 < 0$ há uma competição entre as populações P_1 e P_2 ; (iii) se $\theta_1 > 0$ e $\theta_2 > 0$ têm-se mutualismo ou cooperação ou simbiose; (iv) se $\theta_1 < 0$ e $\theta_2 = 0$ (ou $\theta_1 = 0$ e $\theta_2 < 0$) há comensalismo; (v) se $\theta_1 > 0$ e $\theta_2 = 0$ (ou $\theta_1 = 0$ e $\theta_2 > 0$) têm-se amensalismo. Pode-se encontrar a fundamentação das relações ecológicas mencionadas de (i)-(v) em (SANTOS, 2023).

Formulação Variacional

Nesta seção, desenvolveremos a formulação variacional do sistema de Equações Diferenciais Parciais (EDPs) em questão. Nota-se que, no decorrer do desenvolvimento do raciocínio é conveniente adotar a formulação fraca ou variacional para resolver esse sistema em vez de utilizar a formulação clássica, pelo fato daquela aceitar exigências menos restritas nos parâmetros nas condições iniciais, nas derivações, na solução, e presta-se excepcionalmente à obtenção construtiva de soluções aproximadas e seus métodos. A seguir, apresentaremos o sistema em estudo que descreve as equações das populações $P_1(x, y, t)$, $P_2(x, y, t) \in \mathcal{V}_P$, onde $\mathcal{V}_P = \{P_1, P_2 \in L_2((0, T], H_1(\Omega)) : \frac{\partial P_1}{\partial t}, \frac{\partial P_2}{\partial t} \in L_2(\Omega), \forall t \in (0, T]\}$. Então:

$$\begin{cases} \int_{\Omega} \frac{\partial P_1}{\partial t} \phi \, d\mu - \underbrace{\int_{\Omega} (\alpha_1 \Delta P_1) \phi \, d\mu}_{(1)} + \int_{\Omega} \nabla \cdot (\vec{U} P_1) \phi \, d\mu + \rho_1 \int_{\Omega} \sigma P_1 \phi \, d\mu - \beta_1 \int_{\Omega} P_1 \phi \, d\mu \\ + \gamma_1 \int_{\Omega} P_1^2 \phi \, d\mu + \int_{\Omega} \theta_1 P_1 P_2 \phi \, d\mu = 0, \\ \int_{\Omega} \frac{\partial P_2}{\partial t} \phi \, d\mu - \underbrace{\int_{\Omega} (\alpha_2 \Delta P_2) \phi \, d\mu}_{(2)} + \int_{\Omega} \nabla \cdot (\vec{W} P_2) \phi \, d\mu + \rho_2 \int_{\Omega} \sigma P_2 \phi \, d\mu - \beta_2 \int_{\Omega} P_2 \phi \, d\mu \\ + \gamma_2 \int_{\Omega} P_2^2 \phi \, d\mu + \theta_2 \int_{\Omega} P_1 P_2 \phi \, d\mu = 0, \quad \forall \phi \in \mathcal{V}_P. \end{cases} \quad (3)$$

Pode-se observar que o espaço $\mathcal{V}_P$ restringe o desenvolvimento da formulação fraca, pelo fato de exigir a pertinência das derivadas primeiras em $L_2(\Omega)$, o objetivo é recorrer ao **Teorema de Green** (IÓRIO, 2018) nos termos (1) e (2) do sistema 3, obtendo:

$$\begin{cases} \int_{\Omega} \frac{\partial P_1}{\partial t} \phi \, d\mu + \int_{\Omega} \alpha_1 \nabla P_1 \cdot \nabla \phi \, d\mu - \alpha_1 \int_{\partial\Omega} \frac{\partial P_1}{\partial \eta} \phi \, d\partial\Omega + \int_{\Omega} \nabla \cdot (\vec{U} P_1) \phi \, d\mu \\ + \rho_1 \int_{\Omega} \sigma P_1 \phi \, d\mu - \beta_1 \int_{\Omega} P_1 \phi \, d\mu + \gamma_1 \int_{\Omega} P_1^2 \phi \, d\mu + \theta_1 \int_{\Omega} P_1 P_2 \phi \, d\mu = 0, \\ \int_{\Omega} \frac{\partial P_2}{\partial t} \phi \, d\mu + \int_{\Omega} \alpha_2 \nabla P_2 \cdot \nabla \phi \, d\mu - \alpha_2 \int_{\partial\Omega} \frac{\partial P_2}{\partial \eta} \phi \, d\partial\Omega + \int_{\Omega} \nabla \cdot (\vec{W} P_2) \phi \, d\mu \\ + \rho_2 \int_{\Omega} \sigma P_2 \phi \, d\mu - \beta_2 \int_{\Omega} P_2 \phi \, d\mu + \gamma_2 \int_{\Omega} P_2^2 \phi \, d\mu + \theta_2 \int_{\Omega} P_1 P_2 \phi \, d\mu = 0. \end{cases} \quad (4)$$

Considerando α_1, α_2 os coeficientes de dispersão e $\vec{U}, \vec{W}$ os campos vetoriais de migração independentes do tempo, espaço e das próprias populações e aplicando as condições de contorno de Dirichlet homogênea e Von-Neumann homogênea no sistema 4, obtemos:

$$\begin{cases} \int_{\Omega} \frac{\partial P_1}{\partial t} \phi \, d\mu + \alpha_1 \int_{\Omega} \nabla P_1 \cdot \nabla \phi \, d\mu + U_1 \int_{\Omega} \frac{\partial P_1}{\partial x} \phi \, d\mu + U_2 \int_{\Omega} \frac{\partial P_1}{\partial y} \phi \, d\mu + \rho_1 \int_{\Omega} \sigma P_1 \phi \, d\mu \\ - \beta_1 \int_{\Omega} P_1 \phi \, d\mu + \gamma_1 \int_{\Omega} P_1^2 \phi \, d\mu + \theta_1 \int_{\Omega} P_1 P_2 \phi \, d\mu = 0, \\ \int_{\Omega} \frac{\partial P_2}{\partial t} \phi \, d\mu + \alpha_2 \int_{\Omega} \nabla P_2 \cdot \nabla \phi \, d\mu + W_1 \int_{\Omega} \frac{\partial P_2}{\partial x} \phi \, d\mu + W_2 \int_{\Omega} \frac{\partial P_2}{\partial y} \phi \, d\mu + \rho_2 \int_{\Omega} \sigma P_2 \phi \, d\mu \\ - \beta_2 \int_{\Omega} P_2 \phi \, d\mu + \gamma_2 \int_{\Omega} P_2^2 \phi \, d\mu + \theta_2 \int_{\Omega} P_1 P_2 \phi \, d\mu = 0. \end{cases} \quad (5)$$

Escrevendo o sistema 5 em notação de produto interno no espaço $L_2(\Omega)$, obtemos:

$$\begin{cases} \left(\frac{\partial P_1}{\partial t} \mid \phi \right) + \alpha_1 (\nabla P_1 \mid \nabla \phi) + U_1 \left(\frac{\partial P_1}{\partial x} \mid \phi \right) + U_2 \left(\frac{\partial P_1}{\partial y} \mid \phi \right) + \rho_1 (\sigma P_1 \mid \phi) - \beta_1 (P_1 \mid \phi) \\ + \gamma_1 (P_1^2 \mid \phi) + \theta_1 (P_1 P_2 \mid \phi) = 0, \\ \left(\frac{\partial P_2}{\partial t} \mid \phi \right) + \alpha_2 (\nabla P_2 \mid \nabla \phi) + W_1 \left(\frac{\partial P_2}{\partial x} \mid \phi \right) + W_2 \left(\frac{\partial P_2}{\partial y} \mid \phi \right) + \rho_2 (\sigma P_2 \mid \phi) - \beta_2 (P_2 \mid \phi) \\ + \gamma_2 (P_2^2 \mid \phi) + \theta_2 (P_1 P_2 \mid \phi) = 0. \end{cases} \quad (6)$$

Discretização Espacial pelo Método de Galerkin via MEF

Nesta seção faremos a discretização espacial através do Método de Galerkin via MEF (Método de Elementos Finitos) (KARDESTUNCER, 1987) da formulação variacional do sistema 6 para obtermos uma solução numérica aproximada do modelo em desenvolvimento.

Consideremos o subespaço $\mathcal{V}_{P_h} \subset H_1(\Omega)$ de dimensão finita gerado pela base $\mathfrak{B}_P = \{\phi_1, \phi_2, \dots, \phi_N\}$ sendo a solução para o sistema 6 definida como a combinação linear de elementos dessa base $\mathfrak{B}_P$. Construiremos uma aproximação da solução nesse espaço realizando uma separação entre as variáveis espaciais e a variável temporal dadas da seguinte forma:

$$\begin{aligned} P_1(x, y, t) \cong P_{1h}(x, y, t) &= \sum_{j=1}^N P_{1j}(t) \phi_j(x, y), & \frac{\partial P_{1h}}{\partial t} &= \sum_{j=1}^N \frac{dP_{1j}(t)}{dt} \phi_j, & \frac{\partial P_{1h}}{\partial x} &= \sum_{j=1}^N P_{1j}(t) \frac{\partial \phi_j}{\partial x}, & \frac{\partial P_{1h}}{\partial y} &= \sum_{j=1}^N P_{1j}(t) \frac{\partial \phi_j}{\partial y}, \\ P_2(x, y, t) \cong P_{2h}(x, y, t) &= \sum_{j=1}^N P_{2j}(t) \phi_j(x, y), & \frac{\partial P_{2h}}{\partial t} &= \sum_{j=1}^N \frac{dP_{2j}(t)}{dt} \phi_j, & \frac{\partial P_{2h}}{\partial x} &= \sum_{j=1}^N P_{2j}(t) \frac{\partial \phi_j}{\partial x}, & \frac{\partial P_{2h}}{\partial y} &= \sum_{j=1}^N P_{2j}(t) \frac{\partial \phi_j}{\partial y} \end{aligned} \quad (7)$$

e, para a concentração do poluente $\sigma = \sigma(x, y, t) \cong \sigma_h(x, y, t) = \sum_{k=1}^N u_k(t) \phi_k(x, y)$.

Substituindo os termos das densidades populacionais de P_1 e P_2 por P_{1h} e P_{2h} respectivamente em sua formulação variacional dada pelo sistema 6, tomando $\psi_h \in \mathcal{V}_{P_h}$ e aplicando os termos definidos em 7 e na função da concentração de poluente σ , segue que:

$$\left\{ \begin{array}{l} \sum_{j=1}^N \frac{dP_{1j}(t)}{dt} (\phi_j | \psi_h) + \alpha_1 \sum_{j=1}^N P_{1j}(t) (\nabla \phi_j | \nabla \psi_h) + U_1 \sum_{j=1}^N P_{1j}(t) \left(\frac{\partial \phi_j}{\partial x} | \psi_h \right) + U_2 \sum_{j=1}^N P_{1j}(t) \left(\frac{\partial \phi_j}{\partial y} | \psi_h \right) \\ - \beta_1 \sum_{j=1}^N P_{1j}(t) (\phi_j | \psi_h) + \rho_1 \underbrace{\left(\sum_{k=1}^N u_k(t) \phi_k \right) \left(\sum_{j=1}^N P_{1j}(t) \phi_j | \psi_h \right)}_{(1)} \\ + \gamma_1 \underbrace{\left(\sum_{k=1}^N P_{1k}(t) \phi_k \right) \left(\sum_{j=1}^N P_{1j}(t) \phi_j | \psi_h \right)}_{(2)} + \theta_1 \underbrace{\left(\sum_{k=1}^N P_{2k}(t) \phi_k \right) \left(\sum_{j=1}^N P_{1j}(t) \phi_j | \psi_h \right)}_{(3)} = 0, \\ \sum_{j=1}^N \frac{dP_{2j}(t)}{dt} (\phi_j | \psi_h) + \alpha_2 \sum_{j=1}^N P_{2j}(t) (\nabla \phi_j | \nabla \psi_h) + W_1 \sum_{j=1}^N P_{2j}(t) \left(\frac{\partial \phi_j}{\partial x} | \psi_h \right) + W_2 \sum_{j=1}^N P_{2j}(t) \left(\frac{\partial \phi_j}{\partial y} | \psi_h \right) \\ - \beta_2 \sum_{j=1}^N P_{2j}(t) (\phi_j | \psi_h) + \rho_2 \underbrace{\left(\sum_{k=1}^N u_k(t) \phi_k \right) \left(\sum_{j=1}^N P_{2j}(t) \phi_j | \psi_h \right)}_{(4)} \\ + \gamma_2 \underbrace{\left(\sum_{k=1}^N P_{2k}(t) \phi_k \right) \left(\sum_{j=1}^N P_{2j}(t) \phi_j | \psi_h \right)}_{(5)} + \theta_2 \underbrace{\left(\sum_{k=1}^N P_{1k}(t) \phi_k \right) \left(\sum_{j=1}^N P_{2j}(t) \phi_j | \psi_h \right)}_{(6)} = 0. \end{array} \right. \quad (8)$$

Rearranjando convenientemente os termos não-lineares (1)-(6) do sistema 8, utilizando o **Método de Galerkin**, pelo fato, de garantir que as funções testes $\psi_h \in \mathcal{V}_{P_h}$ serem da mesma classe das funções de base ϕ_i e fazendo algumas manipulações algébricas, obtemos:

$$\left\{ \begin{array}{l} \sum_{j=1}^N \frac{dP_{1j}(t)}{dt} (\phi_j | \phi_i) + \sum_{j=1}^N P_{1j}(t) \left\{ \alpha_1 (\nabla \phi_j | \nabla \phi_i) + U_1 \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + U_2 \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_1 (\phi_j | \phi_i) \right. \\ \left. + \rho_1 \sum_{k=1}^N u_k(t) (\phi_k \phi_j | \phi_i) + \gamma_1 \sum_{k=1}^N P_{1k}(t) (\phi_k \phi_j | \phi_i) + \theta_1 \sum_{k=1}^N P_{2k}(t) (\phi_k \phi_j | \phi_i) \right\} = 0, \\ \sum_{j=1}^N \frac{dP_{2j}(t)}{dt} (\phi_j | \phi_i) + \sum_{j=1}^N P_{2j}(t) \left\{ \alpha_2 (\nabla \phi_j | \nabla \phi_i) + W_1 \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + W_2 \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_2 (\phi_j | \phi_i) \right. \\ \left. + \rho_2 \sum_{k=1}^N u_k(t) (\phi_k \phi_j | \phi_i) + \gamma_2 \sum_{k=1}^N P_{2k}(t) (\phi_k \phi_j | \phi_i) + \theta_2 \sum_{k=1}^N P_{1k}(t) (\phi_k \phi_j | \phi_i) \right\} = 0. \end{array} \right. \quad (9)$$

Os campos vetoriais de migração relacionados aos termos advectivos $U(x, y) = (U_1(x, y), U_2(x, y))$ e $W(x, y) = (W_1(x, y), W_2(x, y))$, respectivamente, descritos em suas equações no sistema 9 possuem valores pontuais aproximados e são obtidos através da resolução numérica da **Equação de Stokes** dado pelo sistema a seguir:

$$\begin{cases} -\Delta \vec{U} + \nabla \mathcal{S}_1 = r_1, \\ -\Delta \vec{W} + \nabla \mathcal{S}_2 = r_2, \end{cases} \quad (10)$$

com $\mathcal{S}_1, \mathcal{S}_2$ sendo os termos que representam a pressão e r_1, r_2 que descrevem as perturbações (que podem ser nulas). Interpolando esses valores através das funções de base $\mathfrak{B}_P$ temos:

$$U_1(x, y) = \sum_{m=1}^N U_{1m} \phi_m(x, y), \quad U_2(x, y) = \sum_{p=1}^N U_{2p} \phi_p(x, y), \quad W_1(x, y) = \sum_{l=1}^N W_{1l} \phi_l(x, y) \quad \text{e} \quad W_2(x, y) = \sum_{q=1}^N W_{2q} \phi_q(x, y). \quad (11)$$

Substituindo os valores dos campos de velocidade de dispersão populacional $U_1(x, y), U_2(x, y), W_1(x, y)$ e $W_2(x, y)$ dados em 11 e fazendo algumas manipulações algébricas no sistema 9, obtemos:

$$\left\{ \begin{aligned} & \sum_{j=1}^N \frac{dP_{1j}(t)}{dt} (\phi_j | \phi_i) + \sum_{j=1}^N P_{1j}(t) \left\{ \alpha_1 (\nabla \phi_j | \nabla \phi_i) + \sum_{m=1}^N U_{1m} \phi_m \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + \sum_{p=1}^N U_{2p} \phi_p \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_1 (\phi_j | \phi_i) \right. \\ & \left. + \sum_{k=1}^N [\rho_1 u_k(t) + \gamma_1 P_{1k}(t) + \theta_1 P_{2k}(t)] (\phi_k \phi_j | \phi_i) \right\} = 0, \\ & \sum_{j=1}^N \frac{dP_{2j}(t)}{dt} (\phi_j | \phi_i) + \sum_{j=1}^N P_{2j}(t) \left\{ \alpha_2 (\nabla \phi_j | \nabla \phi_i) + \sum_{l=1}^N W_{1l} \phi_l \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + \sum_{q=1}^N W_{2q} \phi_q \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_2 (\phi_j | \phi_i) \right. \\ & \left. + \sum_{k=1}^N [\rho_2 u_k(t) + \gamma_2 P_{2k}(t) + \theta_2 P_{1k}(t)] (\phi_k \phi_j | \phi_i) \right\} = 0. \end{aligned} \right. \quad (12)$$

O sistema 12 pode ser escrito na forma matricial, $\begin{cases} \mathbb{C} P_1(t)^{(n)} + \mathbb{D} (u_k^{(n)}, P_{1k}^{(n)}) P_1(t)^{(n)} = 0 \\ \mathbb{E} P_2(t)^{(n)} + \mathbb{F} (u_k^{(n)}, P_{2k}^{(n)}) P_2(t)^{(n)} = 0 \end{cases}$, onde:

$$\begin{aligned} \mathbb{C} = c_{ji} &= (\phi_j | \phi_i)_{L_2(\Omega)} & P_1(t)^{(n)} &= \sum_{j=1}^N \frac{dP_{1j}(t)}{dt}, & P_1(t)^{(n)} &= \sum_{j=1}^N P_{1j}(t), \\ \mathbb{E} = e_{ji} &= (\phi_j | \phi_i)_{L_2(\Omega)}, & P_2(t)^{(n)} &= \sum_{j=1}^N \frac{dP_{2j}(t)}{dt}, & P_2(t)^{(n)} &= \sum_{j=1}^N P_{2j}(t), \\ \mathbb{D} = d_{ji} &= \alpha_1 (\nabla \phi_j | \nabla \phi_i) + \sum_{m=1}^N U_{1m} \phi_m \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + \sum_{p=1}^N U_{2p} \phi_p \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_1 (\phi_j | \phi_i) \\ &+ \sum_{k=1}^N [\rho_1 u_k(t) + \gamma_1 P_{1k}(t) + \theta_1 P_{2k}(t)] (\phi_k \phi_j | \phi_i), \\ \mathbb{F} = f_{ji} &= \alpha_2 (\nabla \phi_j | \nabla \phi_i) + \sum_{l=1}^N W_{1l} \phi_l \left(\frac{\partial \phi_j}{\partial x} | \phi_i \right) + \sum_{q=1}^N W_{2q} \phi_q \left(\frac{\partial \phi_j}{\partial y} | \phi_i \right) - \beta_2 (\phi_j | \phi_i) \\ &+ \sum_{k=1}^N [\rho_2 u_k(t) + \gamma_2 P_{2k}(t) + \theta_2 P_{1k}(t)] (\phi_k \phi_j | \phi_i), \end{aligned}$$

caracterizando-se como um **Sistema de Equações Diferenciais Ordinárias Não-Lineares** na variável t com as condições iniciais já discretizadas e dadas implicitamente por meio da formulação variacional:

$$\sum_{j=1}^N P_{1j}^{(0)} (\phi_j | \phi_i) = (P_{1j_0} | \phi_i) \quad \text{e} \quad \sum_{j=1}^N P_{2j}^{(0)} (\phi_j | \phi_i) = (P_{2j_0} | \phi_i). \quad (13)$$

Discretização Temporal pelo Método de Crank-Nicolson

Nesta seção desenvolveremos a discretização temporal para o modelo de Dispersão-Migração dado pelo sistema 2. A partir do método de Crank-Nicolson (MARTINELLI, 2020), tomaremos $t_n + \frac{\Delta t}{2}$ onde n representa o passo no tempo e o termo dado é definido como diferença centrada no tempo. Assim, para as derivadas temporais, temos as aproximações dadas por:

$$\frac{dP_{1j}^{n+\frac{1}{2}}}{dt} \cong \frac{P_{1j}^{n+1} - P_{1j}^n}{\Delta t} \quad \text{e} \quad \frac{dP_{2j}^{n+\frac{1}{2}}}{dt} \cong \frac{P_{2j}^{n+1} - P_{2j}^n}{\Delta t}, \quad (14)$$

onde o índice n representa o passo no tempo.

Para os valores de P_{1j} , P_{1k} , P_{2j} , P_{2k} e u_k , têm-se as aproximações definidas por:

$$P_{1j}^{n+\frac{1}{2}} = \frac{P_{1j}^{n+1} + P_{1j}^n}{2}, \quad P_{1k}^{n+\frac{1}{2}} = \frac{P_{1k}^{n+1} + P_{1k}^n}{2}, \quad P_{2j}^{n+\frac{1}{2}} = \frac{P_{2j}^{n+1} + P_{2j}^n}{2}, \quad P_{2k}^{n+\frac{1}{2}} = \frac{P_{2k}^{n+1} + P_{2k}^n}{2} \quad \text{e} \quad u_k^{n+\frac{1}{2}} = \frac{u_k^{n+1} + u_k^n}{2}. \quad (15)$$

Utilizando as aproximações dadas em 14 e 15, e multiplicando as equações do sistema 12 por Δt , obtemos:

$$\begin{aligned}
 & \left\{ \sum_{j=1}^N P_{1_j}^{n+1} \left\{ (\phi_j | \phi_i) + \frac{\Delta t}{2} \left[\alpha_1 (\nabla \phi_j \parallel \nabla \phi_i) + \sum_{m=1}^N U_{1_m} \phi_m \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + \sum_{p=1}^N U_{2_p} \phi_p \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) - \beta_1 (\phi_j | \phi_i) \right. \right. \right. \\
 & \quad \left. \left. + \sum_{k=1}^N \left[\rho_1 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_1 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} + \theta_1 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\} \right. \\
 & = \sum_{j=1}^N P_{1_j}^n \left\{ (\phi_j | \phi_i) - \frac{\Delta t}{2} \left[\alpha_1 (\nabla \phi_j \parallel \nabla \phi_i) + \sum_{m=1}^N U_{1_m} \phi_m \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + \sum_{p=1}^N U_{2_p} \phi_p \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) - \beta_1 (\phi_j | \phi_i) \right. \right. \\
 & \quad \left. \left. + \sum_{k=1}^N \left[\rho_1 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_1 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} + \theta_1 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}, \right. \\
 & \left. \sum_{j=1}^N P_{2_j}^{n+1} \left\{ (\phi_j | \phi_i) + \frac{\Delta t}{2} \left[\alpha_2 (\nabla \phi_j \parallel \nabla \phi_i) + \sum_{l=1}^N W_{1_l} \phi_l \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + \sum_{q=1}^N W_{2_q} \phi_q \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) - \beta_2 (\phi_j | \phi_i) \right. \right. \right. \\
 & \quad \left. \left. + \sum_{k=1}^N \left[\rho_2 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_2 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} + \theta_2 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\} \right. \\
 & = \sum_{j=1}^N P_{2_j}^n \left\{ (\phi_j | \phi_i) - \frac{\Delta t}{2} \left[\alpha_2 (\nabla \phi_j \parallel \nabla \phi_i) + \sum_{l=1}^N W_{1_l} \phi_l \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + \sum_{q=1}^N W_{2_q} \phi_q \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) - \beta_2 (\phi_j | \phi_i) \right. \right. \\
 & \quad \left. \left. + \sum_{k=1}^N \left[\rho_2 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_2 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} + \theta_2 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}. \right.
 \end{aligned} \tag{16}$$

O sistema não-linear 16 caracteriza-se pela relação $P_{1_k}^{(n)} = P_1(x_k, y_k, t_n)$, $P_{2_k}^{(n)} = P_2(x_k, y_k, t_n)$ e $u_k^{(n)} = u_k(x_k, y_k, t_n)$ e pode ser reescrito por:

$$\begin{cases} \mathbb{C} \left(P_{1_k}^{(n+1)}, P_{1_k}^{(n)}, P_{2_k}^{(n+1)}, P_{2_k}^{(n)}, u_k^{(n+1)}, u_k^{(n)} \right) P_1^{(n+1)} = \mathbb{D} \left(P_{1_k}^{(n+1)}, P_{1_k}^{(n)}, P_{2_k}^{(n+1)}, P_{2_k}^{(n)}, u_k^{(n+1)}, u_k^{(n)} \right) P_1^{(n)}, \\ \mathbb{E} \left(P_{1_k}^{(n+1)}, P_{1_k}^{(n)}, P_{2_k}^{(n+1)}, P_{2_k}^{(n)}, u_k^{(n+1)}, u_k^{(n)} \right) P_2^{(n+1)} = \mathbb{F} \left(P_{1_k}^{(n+1)}, P_{1_k}^{(n)}, P_{2_k}^{(n+1)}, P_{2_k}^{(n)}, u_k^{(n+1)}, u_k^{(n)} \right) P_2^{(n)}, \end{cases} \tag{17}$$

e as matrizes $\mathbb{C}, \mathbb{D}, \mathbb{E}$ e $\mathbb{F}$, tornam-se:

$$\begin{aligned}
 \mathbb{C} = c_{ji} &= \left[1 - \beta_1 \frac{\Delta t}{2} \right] (\phi_j | \phi_i) + \frac{\Delta t}{2} \left\{ \alpha_1 (\nabla \phi_j \parallel \nabla \phi_i) + U_1 \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + U_2 \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) \right. \\
 & \quad \left. + \sum_{k=1}^N \left[\rho_1 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_1 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} + \theta_1 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}, \\
 \mathbb{D} = d_{ji} &= \left[1 + \beta_1 \frac{\Delta t}{2} \right] (\phi_j | \phi_i) - \frac{\Delta t}{2} \left\{ \alpha_1 (\nabla \phi_j \parallel \nabla \phi_i) + U_1 \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + U_2 \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) \right. \\
 & \quad \left. + \sum_{k=1}^N \left[\rho_1 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_1 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} + \theta_1 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}, \\
 \mathbb{E} = e_{ji} &= \left[1 - \beta_2 \frac{\Delta t}{2} \right] (\phi_j | \phi_i) + \frac{\Delta t}{2} \left\{ \alpha_2 (\nabla \phi_j \parallel \nabla \phi_i) + W_1 \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + W_2 \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) \right. \\
 & \quad \left. + \sum_{k=1}^N \left[\rho_2 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_2 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} + \theta_2 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}, \\
 \mathbb{F} = f_{ji} &= \left[1 + \beta_2 \frac{\Delta t}{2} \right] (\phi_j | \phi_i) - \frac{\Delta t}{2} \left\{ \alpha_2 (\nabla \phi_j \parallel \nabla \phi_i) + W_1 \left(\frac{\partial \phi_j}{\partial x} \mid \phi_i \right) + W_2 \left(\frac{\partial \phi_j}{\partial y} \mid \phi_i \right) \right. \\
 & \quad \left. + \sum_{k=1}^N \left[\rho_2 \frac{u_k^{n+1} + u_k^n}{2} + \gamma_2 \frac{P_{2_k}^{n+1} + P_{2_k}^n}{2} + \theta_2 \frac{P_{1_k}^{n+1} + P_{1_k}^n}{2} \right] (\phi_k \phi_j | \phi_i) \right\}.
 \end{aligned}$$

Esse sistema será resolvido iterativamente a cada passo no tempo, a partir das condições iniciais P_{1_0} e P_{2_0} já apresentadas pelos termos em 14. Com isso, podemos considerar as condições iniciais dadas por $P_{1_k}^{(0)} = P_1(x_k, y_k, t_n)$ e $P_{2_k}^{(0)} = P_2(x_k, y_k, t_n)$ ou implicitamente por meio da formulação variacional na forma:

$$\sum_{k=1}^N P_{1_k}^{(0)}(\phi_k | \phi_i) = (P_{1_{k_0}} | \phi_i) \quad \text{e} \quad \sum_{k=1}^N P_{2_k}^{(0)}(\phi_k | \phi_i) = (P_{2_{k_0}} | \phi_i).$$

Considerações Finais

Até o momento o estudo contribuiu para o desenvolvimento científico do pesquisador, possibilitando o aprendizado de técnicas e conceitos matemáticos e computacionais mais elaborados numericamente nas áreas de Ecologia Matemática e Análise Numérica.

Encontra-se em desenvolvimento a discretização espacial por MEF da região em estudo do problema. A região é conhecida como Bacia Sedimentar da Foz do Amazonas; foi desenvolvido a malha da região através das coordenadas geográficas disponibilizadas pelo software Google Earth e implementadas no software GMSH através de um arquivo de texto com entrada desses dados; será desenvolvido um algoritmo no software Matlab com as EDPs de Dispersão-Migração, juntamente com a malha gerada no software GMSH e carregada no algoritmo com finalidade de apresentar resultados pertinentes sobre o comportamento das densidades populacionais de P_1 e P_2 e os impactos do poluente σ em meio aquático na região descrita.

Os resultados finais serão publicados na versão final da dissertação de mestrado do autor apresentador do trabalho. Pode-se destacar também, que haverá resultados para outros tipos de dinâmicas populacionais, como por exemplo, espécies consideradas generalistas ou invasoras.

Referências

- LOTKA, A. J. **Elements of Physical Biology**. Science Progress in the Twentieth Century (1919-1933), vol. 21, no. 82, 1926, p. 341-343. DOI: 10.1038/116461b0.
- VOLTERRA, V. **Fluctuations in The Abundance of A Species Considered Mathematically**. Nature, no. 118, 1926, p. 558-560. DOI: 10.1038/118558a0.
- SKELLAM, J. G. **Random dispersal in theoretical populations**. Bulletin of Mathematical Biology, vol. 53, 1991, p. 135-165, ISSN: 0092-8240.
- MURRAY, J. D. **Mathematical Biology I: An Introduction**. 3rd. ed. Springer-Verlag, 2003. (Interdisciplinary Applied Mathematics, v. 17). ISBN: 9780387952239.
- SOSSAE, R.C. **A presença evolutiva de um material impactante e seu efeito no transiente populacional de espécies interativas: modelagem e aproximação**. 2003. 98 f. Dissertação (Mestrado em Matemática Aplicada) - Unicamp, Campinas-SP, 2003.
- SANTOS, H. S. **Relações Ecológicas**. BiologiaNet. 2023. Disponível em: <https://www.biologianet.com/ecologia/relacoes-ecologicas.htm>. Acesso em: 17 ago. 2023.
- KARDESTUNCER, H.; NORRIE, D. H.; BREZZI, F. **Finite element handbook**. New York: McGraw-Hill, Inc., 1987. ISBN: 007033305X.
- IÓRIO, V. M. **EDP: um curso de graduação**. 4. ed. Rio de Janeiro: Instituto de Matemática Pura e Aplicada - IMPA, 2018. (Coleção Matemática Universitária). ISBN: 9788524404221.
- MARTINELLI, G. L. **Determinação do fluxo de calor em uma região plana com condições de contorno mistas, utilizando o método de Crank-Nicolson**. 57 f. Monografia (Trabalho de Conclusão de Curso em Licenciatura em Matemática) - UFFS, Chapecó-SC, 2020.

Estabilidade e Número Básico de Reprodução - R_0 para o Modelo SIRS-B da Cólera em Moçambique

José Manuel Samo

Departamento de Matemática Aplicada, IMECC, UNICAMP

email. jmsamo@ime.unicamp.br

João Frederico C. A. Meyer

Departamento de Matemática Aplicada, IMECC, UNICAMP

email. jmeyer@unicamp.br

Resumo: Para este trabalho, é considerado o modelo matemático clássico *SIR* para estudar o comportamento da doença, e ao modelo incorporamos um compartimento para representar as bactérias responsáveis pela propagação da mesma. Para estudar a estabilidade e o número básico de reprodução R_0 para a doença da cólera, utilizamos as equações diferenciais descritas no modelo, tendo em consideração os parâmetros apresentados para determinar o número de infecções secundárias produzidas por um único indivíduo infectado em uma população completamente suscetível. Foi possível estabelecer uma relação entre o número reprodutivo básico parcial induzido pela transmissão do meio ambiente para a população humana (R_{0B}) e o número reprodutivo básico parcial induzido pela transmissão de população humana para humana (R_{0I}). A estabilidade linear é determinada pelo número básico de reprodução R_0 , matematicamente definido como o raio espectral da matriz de próxima geração.

Palavras-chave: Cólera. Equações Diferenciais Ordinárias. Estabilidade. Número Básico de Reprodução

Introdução

A sétima pandemia de cólera começou nas Ilhas Celebes, na Indonésia, e atingiu o continente africano em 1970. Depois, em 1991, a doença regressou a África. O agente causador desta doença é uma bactéria que apresenta alta viabilidade no meio ambiente, denominada de biótipo *Vibrio Cholerae*, também conhecida como Vibrião Colérico, descoberta por Robert Koch em 1883 (ARAGÓN *et al.*, 1994). Esta bactéria para além de viver no meio ambiente, principalmente em ambientes aquáticos, ela utiliza os humanos como hospedeiros, multiplica-se rapidamente no intestino humano, causando dor abdominal e diarreia aquosa intensa.

A epidemia de cólera tornou-se um dos principais problemas de saúde pública em Moçambique. Ela apareceu pela primeira vez na cidade de Maputo em 1973 e, posteriormente, espalhou-se por todo o país em menor escala, persistindo por longos períodos. Isso levou a uma mudança no caráter epidemiológico da doença, passando de 'epidêmica' para 'endêmica', com casos ocorrendo ao longo do ano. Isso tornou o tratamento da cólera um fardo significativo para o Serviço Nacional de Saúde. Essa variação pode ser atribuída à deterioração das condições de saneamento, como o abastecimento de água precário, a venda de alimentos sem proteção adequada, a falta de um controle alimentar eficaz e práticas deficientes de higiene, tanto individual quanto coletiva (MISAU, 2010).

O Ministério da Saúde realiza inquéritos sobre os indicadores de cólera em Moçambique para permitir uma avaliação de base dos indicadores-chave antes de intensificar as intervenções para desenvolver estratégias para implementar medidas para melhorar o estado de saúde da população, de acordo com Ministério da Saúde da República de Moçambique (MISAU, 2016).

Para ajustar as taxas de transmissão, recuperação e mortalidade é necessário um estudo da estabilidade e dos pontos de equilíbrio nos modelos epidemiológicos. Enquanto isso, o número básico de reprodução R_0 fornece informações sobre a potencial propagação de uma doença numa população sem medidas imunológicas ou de controle prévias (DIEKMANN; HEESTERBEEK; METZ, 1990). A compreensão destes conceitos é essencial para prever o comportamento da doença nas populações, desenvolver estratégias de controle e intervenção e tomar decisões informadas em saúde pública.

Este trabalho tem como objetivo, com base no modelo, analisar a estabilidade dos pontos de equilíbrio e número básico de reprodução de forma a avaliar o comportamento da doença e criar ferramentas de apoio em políticas públicas para implementar as medidas recomendadas pela Organização Mundial da Saúde (OMS) para a prevenção e erradicação da doença.

Estabilidade e Número básico de reprodução R_0

Pontos de equilíbrio - Um ponto de equilíbrio é um estado em que as variáveis do modelo permanecem constantes ao longo do tempo. Em modelos epidemiológicos, os pontos de equilíbrio representam estados em que não ocorre variação na prevalência da doença. Existem dois tipos principais de pontos de equilíbrio em modelos epidemiológicos, conforme descrito por Mwasa e Tchuente (2011):

Ponto de equilíbrio livre de doenças - Neste ponto, não há casos da doença na população.

Ponto de equilíbrio endêmico - Neste ponto, a doença está presente na população permanentemente.

Estabilidade - A estabilidade de um ponto de equilíbrio num modelo epidemiológico refere-se à capacidade do sistema retornar a esse ponto. Existem dois tipos principais de estabilidade:

Estabilidade do Ponto de Equilíbrio Livre de Doença - Se, após uma pequena perturbação, o sistema retornar ao ponto de equilíbrio livre de doença, esse ponto é considerado estável.

Estabilidade do Ponto de Equilíbrio Endêmico - Se, após uma perturbação, o sistema retornar ao ponto de equilíbrio endêmico, ele é considerado estável.

O estado de equilíbrio mais importante para o controle da doença é o equilíbrio livre de doença *Disease Free Equilibrium* - DFE), e sua estabilidade linear é determinada pelo número básico de reprodução R_0 , matematicamente definido como o raio espectral da matriz de próxima geração. É o número de infecções secundárias produzidas por um único indivíduo infetado em uma população completamente suscetível (HETHCOTE, 2000).

Segundo Diekmann, Heesterbeek e Metz, (1990), número básico de reprodução, R_0 , é definido como sendo o número esperado de infecções secundárias produzidas por índice de casos em uma população completamente suscetível. Este número é uma medida que mede capacidade de propagação da doença na população.

Se, $R_0 \leq 1$, então alguns indivíduos infectados introduzidos na população suscetível à doença não conseguirão se reinfectar, e a doença não se espalhará.

Por outro lado, se $R_0 > 1$, então o número de indivíduos infetados aumentará a cada geração e a doença se espalhará.

Teorema 1. *Considere uma Equação Diferencial Ordinária, os seguintes resultados são válidos: (1) Se $R_0 \leq 1$, então o DFE é global assintoticamente estável em Ω_c e (2) Se $R_0 > 1$, então o DFE é instável.*

Observe que o (R_0) é um parâmetro limite para a invasão de um organismo doente em uma população totalmente suscetível. Uma vez que a doença começou a se propagar, as condições que favorecem a propagação mudarão e R_0 pode não ser mais uma boa medida de transmissão da doença (DIEKMANN; HEESTERBEEK; METZ, 1990). O R_0 não pode ser determinado apenas a partir da estrutura do modelo matemático, mas depende da definição de compartimentos infetados e não infetados, como aponta Mwasa e Tchuente (2011). Para o seu cálculo é importante distinguir novas infecções de todas as outras quando se observa mudanças na população.

Segundo Van Den Driessche e Watmough (2002), considerando $\mathcal{F}_i(x)$ como a taxa de aparecimento de novas infecções no compartimento i , seja $\mathcal{V}_i^+(x)$ a taxa de transferência de indivíduos para o compartimento i por todos os outros meios, e seja $\mathcal{V}_i^-(x)$ a taxa de transferência de indivíduos para fora do compartimento i . Assume-se que cada função é continuamente diferenciável pelo menos duas vezes em cada variável. O modelo de transmissão da doença consiste em condições iniciais não negativas juntamente com o seguinte sistema de equações:

$$\dot{x}_i = f_i(x) = \mathcal{F}_i(x) - \mathcal{V}_i(x), \quad i = 1, 2, \dots, n. \quad (1)$$

onde $\mathcal{V}_i(x) = \mathcal{V}_i^-(x) - \mathcal{V}_i^+(x)$ e as funções, que representam uma transferência direta de indivíduos e todas elas são não negativas, satisfazem as premissas descritas abaixo.

Lema 1. Se x_0 é um DFE de (1) e $f_i(x)$ satisfaz as condições anteriores, então as derivadas $D\mathcal{F}(x_0)$ e $D\mathcal{V}(x_0)$ são particionadas como sendo $D\mathcal{F}(x_0) = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix}$, $D\mathcal{V}(x_0) = \begin{bmatrix} V & 0 \\ J_3 & J_4 \end{bmatrix}$, onde F e V são as matrizes quadradas definidas por $F = \left[\frac{\partial \mathcal{F}_i(x_0)}{\partial x_j} \right]$ e $V = \left[\frac{\partial \mathcal{V}_i(x_0)}{\partial x_j} \right]$ com $1 \leq i, j \leq m$. Além disso, F é não negativo, V é uma M -matriz não singular e todos os autovalores de J_4 têm parte real positiva.

Para interpretar as entradas de FV^{-1} e desenvolver uma definição significativa de R_0 , considere o destino de um indivíduo infectado introduzido no compartimento k de uma população livre de doenças. O $(j; k)$ entrada de V^{-1} é o período médio de tempo que este indivíduo passa no compartimento j durante a sua vida útil, assumindo que a população permaneça próxima ao DFE e barrando a reinfecção. O $(i; j)$ entrada de F é a taxa a que os indivíduos infectados no compartimento j produzem novas comunidades no compartimento i . Assim, o $(i; k)$ entrada do produto FV^{-1} é o número esperado de novas comportadas em compartimento i produzido pelo indivíduo infectado originalmente apresentado no compartimento k (VAN DEN DRIESSCHE; WATMOUGH, 2002). Assim, FV^{-1} será designada como a matriz de próxima geração para o modelo e definida por:

$$R_0 = \rho(FV^{-1}) \quad (2)$$

onde $\rho(FV^{-1})$ denota o raio espectral de uma matriz (FV^{-1}) .

Assim sendo, um número básico de reprodução mais geral pode ser definido como o número de novas infecções produzidas por um indivíduo infectante típico em uma população em um DFE (VAN DEN DRIESSCHE; WATMOUGH 2002).

A DFE, x_0 , é localmente assintoticamente estável se todos os autovalores da matriz $Df(x_0)$ tiverem partes reais negativas e instável se algum autovalor de $Df(x_0)$ tiver parte real positiva. Pelo Lema 1, os autovalores de $Df(x_0)$ podem ser particionados em dois conjuntos correspondentes aos compartimentos infectados e não infectados. Esses dois conjuntos são os autovalores de $F - V$ e os de $-J_4$. Novamente pelo Lema 1, todos os autovalores de $-J_4$ têm parte real negativa, portanto a estabilidade da DFE é determinada pelos autovalores de $F - V$. O teorema a seguir afirma que R_0 é um parâmetro limite para a estabilidade do DFE.

Teorema 2. Considere o modelo de transmissão de doenças dado por (1) com $f(x)$ satisfazendo as condições do Lema 1. Se x_0 é um DFE do modelo, então x_0 é (1) localmente assintoticamente estável se $R_0 < 1$ e (2) instável se $R_0 > 1$, onde R_0 é definido por (2).

Equilíbrio Endêmico e Estabilidade Global para o Modelo SIRS – B

Um dos modelos clássicos de epidemiologia é o modelo SIR, desenvolvido por Kermack e McKendrick em 1927, que divide a população em segmentos: suscetível, infectado e removido, que pode ser morto ou resistente (recuperado) (KERMACK; MCKENDRICK, 1927). Neste trabalho, formulamos uma variante deste modelo denominada pelo modelo SIRS – B que descreve a propagação da infecção em uma população ao longo do tempo para estudar a dinâmica da cólera, cujas hipóteses do modelo são as seguintes: (1) O sistema envolve uma população total de humanos que varia com o tempo, denotada por $N(t) = (S(t) + I(t) + R(t))$ e dividida em 3 compartimentos de subpopulações: Suscetível ($S = S(t)$), Infectada ($I = I(t)$), Resistente ($R = R(t)$), incluímos um compartimento no modelo para representar a população das bactérias ($B = B(t)$) responsáveis pela transmissão de doenças. (2) O modelo assume um recrutamento da população humana suscetível com um crescimento logístico de Verhulst com uma taxa que depende do tempo t , por ser usada especialmente em contextos biológicos e demográficos, denotada por $\Gamma = \Lambda \left(1 - \frac{S(t)+I(t)+R(t)}{k_s} \right)$ de indivíduos suscetíveis, onde Λ é a taxa de crescimento intrínseco da

população e k_S é a capacidade de suporte da população humana (limite máximo que a população pode alcançar). (3) Assumimos que a população total das bactérias que entram no reservatório aquático de Vibrião Colérico no modelo com um crescimento logístico de Verhulst cuja taxa de $n \left(1 - \frac{B(t)}{k_B}\right)$ proporcional à densidade de bactérias, onde $n > 0$ é a taxa de crescimento per capita para Vibrião Colérico, $k_B > 0$ é a capacidade de suporte do ambiente ou meio (limite máximo que a população das bactérias pode alcançar no meio). (4) A contribuição média (células/L por dia) de cada pessoa infectada para a população de Vibrião Colérico no ambiente ou meio aquática tem uma taxa fixa ψ (CODEÇO, 2001). Geralmente, de acordo com Joh *et al.* (2009), os indivíduos infectados com cólera liberam a bactéria de volta para o ambiente aquático, aumentando assim a transmissibilidade da cólera. (5) Consideramos também que $\beta_B > 0$ é a taxa de ingestão da bactéria através de fontes de água e alimentos contaminados, h é a meia constante de saturação da população de bactérias e $\frac{\beta_B B}{h+B}$ é a possibilidade de um indivíduo infectado ter a doença com sintomas, dado o contato com fontes contaminadas. Os demais parâmetros do modelo são μ_S , μ_I e μ_R representam as taxas de mortalidade natural da população humana suscetível, infectada e recuperada, respectivamente. d representa uma taxa adicional à taxa de mortalidade natural, denominada taxa de letalidade, que representa a morte ocasionada por doenças infecciosas, θ é a taxa com a qual a população humana recuperada da doença torna-se suscetível, γ é a taxa com a qual a população humana infectada torna-se recuperada da doença, β_I é a taxa de contaminação das pessoas suscetíveis ao entrar em contato com as contaminadas, m é a taxa de mortalidade bacteriana no reservatório. Vamos considerar o Sistema de Equações Diferenciais Ordinárias abaixo e estudar estabilidade local dos pontos de equilíbrio da dinâmica da doença.

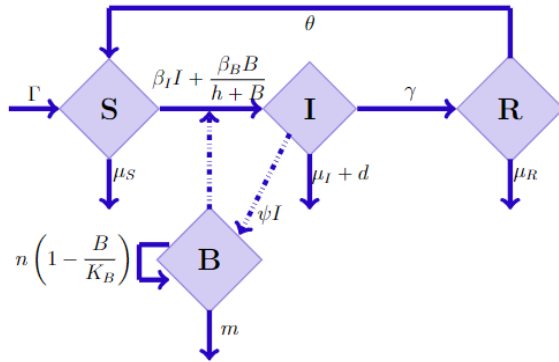


Figura 1: Esquema do Modelo SIRS-B

$$\begin{cases} \frac{dS}{dt} = \Lambda(S+I+R) \left(1 - \frac{S+I+R}{K_S}\right) + \theta R - \mu_S S - \beta_I I S - \frac{\beta_B B}{h+B} S; \\ \frac{dI}{dt} = \beta_I I S + \frac{\beta_B B}{h+B} S - \gamma I - (\mu_I + d) I; \\ \frac{dR}{dt} = \gamma I - \mu_R R - \theta R; \\ \frac{dB}{dt} = nB \left(1 - \frac{B}{K_B}\right) - mB + \psi I; \end{cases} \quad (3)$$

Figura 2: Sistema de EDO-SIRS-B

Todas as variáveis e parâmetros do modelo são não negativos. Portanto, o Sistema (3) é biologicamente (epidemiologicamente) e matematicamente bem definido no conjunto fechado $\Psi_B = \{B \in \mathbb{R}_0^+ : 0 \leq B < \infty\}$ e $\Psi_H = \{(S, I, R) \in \mathbb{R}_+^3 : 0 \leq N \leq \frac{\Lambda}{\mu_S}\}$ com $N(t) = S(t) + I(t) + R(t)$, definimos $\Psi_V = \Psi_H \times \Psi_B$. Para estudar a estabilidade local consideramos o Sistema (3) em equilíbrio estacionário, teremos:

$$\begin{cases} 0 = \Lambda(S+I+R) \left(1 - \frac{S+I+R}{K_S}\right) + \theta R - \mu_S S - \beta_I I S - \frac{\beta_B B}{h+B} S; \\ 0 = \beta_I I S + \frac{\beta_B B}{h+B} S - \gamma I - (\mu_I + d) I; \\ 0 = \gamma I - \mu_R R - \theta R; \\ 0 = \left(n \left(1 - \frac{B}{K_B}\right) - m\right) B + \psi I; \end{cases} \Leftrightarrow \begin{cases} \Lambda(S+I+R) \left(1 - \frac{S+I+R}{K_S}\right) + \theta R - \mu_S S - \beta_I I S - \frac{\beta_B B}{h+B} S = 0; \\ S = \frac{\gamma + \mu_I + d}{\psi} \left(m - n + \frac{n}{K_B} B\right) \left(\frac{h+B}{h\beta_I I + (\beta_B + \beta_I I) B}\right) B; \\ R = \frac{\gamma}{(\mu_R + \theta)\psi} \left(m - n + \frac{n}{K_B} B\right) B; \\ I = \frac{1}{\psi} \left(m - n + \frac{n}{K_B} B\right) B. \end{cases} \quad (4)$$

O equilíbrio endêmico do Sistema Modelo (4) denotado ε^B é dado por: $\varepsilon^B = (S^B, I^B, R^B, B) \neq (0, 0, 0, 0)$, com $B = B(\Lambda, \mu_S, \mu_I, \mu_R, \beta_I, \beta_B, K_S, K_B, \gamma, \theta, \psi, m, n, h)$ e:

$$S^B = \frac{\gamma + \mu_I + d}{\psi} \left(m - n + \frac{n}{K_B} B\right) \left(\frac{h+B}{h\beta_I I + (\beta_B + \beta_I I) B}\right) B; \quad (5)$$

$$I^B = \frac{1}{\psi} \left(m - n + \frac{n}{K_B} B \right) B; \quad e \quad R^B = \frac{\gamma}{(\mu_R + \theta)\psi} \left(m - n + \frac{n}{K_B} B \right) B; \quad (6)$$

Lema 2. Para $R_0 < 1$, o único equilíbrio endêmico ε^B existe em Ψ_V e é globalmente assintoticamente estável (Globally Asymptotically Stable - GAS) em $\Psi_V \setminus \Psi_{V_0}$ fornecido $R_0 = R_0(S^B, I^B, R^B, B) \approx 1$

Equilíbrio e Estabilidade Livres da Doença

Consideramos este como sendo um ponto em que não ocorre a doença na população. Naturalmente, isso pode ocorrer para algumas doenças ou como resultado de medidas de controle e prevenção bem-sucedidas.

O equilíbrio livre da doença (DFE) do submodelo para a cólera do Sistema (4), isto é, se $B = 0$, então $I = 0$, $R = 0$ e $\Lambda S - \frac{\Lambda}{K_S} S^2 - \mu_S S = 0 \Leftrightarrow S = 0 \vee S = \frac{(\Lambda - \mu_S)K_S}{\Lambda}$, assim teremos a jacobiana:

$$J(S, I, R, B) = \begin{pmatrix} \Lambda - \frac{2\Lambda(S+I+R)}{K_S} - \mu_S - \beta_I I - \frac{\beta_B B}{h+B} & \Lambda - \beta_I S - \frac{2\Lambda(S+I+R)}{K_S} & \Lambda - \theta - \frac{2\Lambda(S+I+R)}{K_S} & -\frac{h\beta_B S}{(h+B)^2} \\ \beta_I I + \frac{\beta_B B}{h+B} & -\gamma - \mu_I - d & 0 & \frac{h\beta_B S}{(h+B)^2} \\ 0 & \gamma & -\mu_R - \theta & 0 \\ 0 & \psi & 0 & -\frac{2nB}{K_B} + (n-m) \end{pmatrix}$$

O equilíbrio trivial, cuja matriz jacobiana calculada no ponto $\varepsilon^0 = (S, I, R, B) = (0, 0, 0, 0)$, é dado por:

$$J(\varepsilon^0) = \begin{pmatrix} \Lambda - \mu_S & \Lambda & \Lambda - \theta & 0 \\ 0 & -\gamma - \mu_I - d & 0 & 0 \\ 0 & \gamma & -\mu_R - \theta & 0 \\ 0 & \psi & 0 & n - m \end{pmatrix}$$

Os autovalores da matriz $J(\varepsilon^0)$ são $\lambda_1 = \Lambda - \mu_S$, $\lambda_2 = -\gamma - \mu_I - d$, $\lambda_3 = -\mu_R - \theta$ e $\lambda_4 = n - m$. O equilíbrio livre das bactérias, consequentemente a população dos infectados, cuja matriz jacobiana calculada no ponto $\varepsilon^* = (S, I, R, B) = \left(\frac{(\Lambda - \mu_S)K_S}{\Lambda}, 0, 0, 0 \right)$ é dada por:

$$J(\varepsilon^*) = \begin{pmatrix} \mu_S - \Lambda & 2\mu_S - \Lambda - \frac{\beta_I(\Lambda - \mu_S)K_S}{\Lambda} & 2\mu_S - \Lambda - \theta & -\frac{\beta_B(\Lambda - \mu_S)K_S}{h\Lambda} \\ 0 & -\gamma - \mu_I - d & 0 & \frac{\beta_B(\Lambda - \mu_S)K_S}{h\Lambda} \\ 0 & \gamma & -\mu_R - \theta & 0 \\ 0 & \psi & 0 & n - m \end{pmatrix}$$

Os autovalores da matriz $J(\varepsilon^*)$ são $\lambda_1 = \mu_S - \Lambda$, $\lambda_2 = \lambda_3 = \frac{(n-m) - (\gamma + \mu_I + d)}{2}$ e $\lambda_4 = -\mu_R - \theta$, então o equilíbrio é instável, se $\mu_S - \Lambda < 0$ e $n - m < 0$.

Vamos considerar F - Taxa na qual novos infectados entram no compartimento i e V - Denota a taxa de transferência de indivíduos para dentro e para fora do compartimento i , assim, considerando que $\left(\frac{dS}{dt}, \frac{dI}{dt}, \frac{dR}{dt}, \frac{dB}{dt} \right)^T = \mathcal{F} - \mathcal{V}$, com

$$\mathcal{F} = \begin{pmatrix} 0 \\ \beta_I I S + \left(\frac{\beta_B B}{h+B} \right) S \\ 0 \\ 0 \end{pmatrix} \quad e \quad \mathcal{V} = \mathcal{V}^- - \mathcal{V}^+ = \begin{pmatrix} \left(\mu_S + \frac{\beta_B B}{h+B} \right) S - \Lambda(S+I+R) \left(1 - \frac{S+I+R}{K_S} \right) - \theta R \\ (\gamma + \mu_I + d) I \\ (\mu_R + \theta) R - \gamma I \\ \frac{n}{K_B} B^2 + (m-n)B - \psi I \end{pmatrix}.$$

As matrizes jacobianas de $\mathcal{F}$ e de $\mathcal{V}$ são, respectivamente, dadas por:

$$\begin{aligned}
 F = J_{\mathcal{F}}(S, I, R, B) &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \beta_I I + \frac{\beta_B B}{h+B} & \beta_I S & 0 & \frac{\beta_B h}{(h+B)^2} S \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad V = J_{\mathcal{V}}(S, I, R, B) = \\
 &= \begin{pmatrix} \mu_S + \frac{\beta_B B}{h+B} - \Lambda + \beta_I I + \frac{2\Lambda}{k_S}(S+I+R) & -\Lambda + \beta_I S + \frac{2\Lambda}{k_S}(S+I+R) & -\Lambda + \frac{2\Lambda}{k_S}(S+I+R) - \theta & \frac{\beta_B h}{(h+B)^2} S \\ 0 & \gamma + \mu_I + d & 0 & 0 \\ 0 & -\gamma & \mu_R + \theta & 0 \\ 0 & -\psi & 0 & 2\frac{n}{k_B} B + m - n \end{pmatrix} \quad (7)
 \end{aligned}$$

Se $S = 0$ temos que $\varepsilon^0 = (S^0, I^0, R^0, B^0) = (0, 0, 0, 0)$ é o equilíbrio trivial que implica na extinção das populações,

$$\begin{aligned}
 F(\varepsilon^0) &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ \beta_I \cdot 0 + \frac{\beta_B \cdot 0}{h+0} & \beta_I \cdot 0 & 0 & \frac{\beta_B h}{(h+0)^2} \cdot 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{e} \\
 V(\varepsilon^0) &= \begin{pmatrix} \mu_S - \Lambda & -\Lambda & -\Lambda - \theta & 0 \\ 0 & \gamma + \mu_I + d & 0 & 0 \\ 0 & -\gamma & \mu_R + \theta & 0 \\ 0 & -\psi & 0 & m - n \end{pmatrix} \Rightarrow F(\varepsilon^0)V^{-1}(\varepsilon^0) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

O número básico de reprodução para o modelo $SIRS - B$ de cólera no ponto de equilíbrio trivial, $\varepsilon^0 = (S^0, I^0, R^0, B^0) = (0, 0, 0, 0)$, é dado por: $R_0 = 0$. Em modelos epidemiológicos, $R_0 = 0$ indica que a doença não tem capacidade de se espalhar na população. Se $S \neq 0$ temos que $\varepsilon^* = (S^*, I^*, R^*, B^*) = \left(\frac{k_S(\Lambda - \mu_S)}{\Lambda}, 0, 0, 0\right)$ é o equilíbrio livre da doença para $(\Lambda - \mu_S) > 0$. Então,

$$F(\varepsilon^*) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \beta_I \cdot 0 + \frac{\beta_B 0}{h+0} & \beta_I \frac{k_S(\Lambda - \mu_S)}{\Lambda} & 0 & \frac{\beta_B h}{(h+0)^2} \frac{k_S(\Lambda - \mu_S)}{\Lambda} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \beta_I \frac{k_S(\Lambda - \mu_S)}{\Lambda} & 0 & \frac{\beta_B k_S(\Lambda - \mu_S)}{h\Lambda} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

e

$$\begin{aligned}
 V(\varepsilon^*) &= \begin{pmatrix} \mu_S - \Lambda & \frac{\beta_I k_S(\Lambda - \mu_S)}{\Lambda} + \Lambda - 2\mu_S & +\Lambda - 2\mu_S - \theta & \frac{\beta_B k_S(\Lambda - \mu_S)}{h\Lambda} \\ 0 & \gamma + \mu_I + d & 0 & 0 \\ 0 & -\gamma & \mu_R + \theta & 0 \\ 0 & -\psi & 0 & m - n \end{pmatrix} \\
 \Rightarrow F(\varepsilon^*)V^{-1}(\varepsilon^*) &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\beta_I k_S(\Lambda - \mu_S)}{\Lambda(\gamma + \mu_I + d)} + \frac{\psi \beta_B k_S(\Lambda - \mu_S)}{h\Lambda(\gamma + \mu_I + d)(m - n)} & 0 & \frac{\beta_B k_S(\Lambda - \mu_S)}{h\Lambda(m - n)} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
 \end{aligned}$$

O número básico de reprodução (R_0) para o modelo $SIRS - B$ de cólera é dado por:

$$R_0 = \rho [F(\varepsilon^*)V^{-1}(\varepsilon^*)] = \frac{k_S(\Lambda - \mu_S)}{\Lambda(\gamma + \mu_I + d)} \beta_I + \frac{\psi k_S(\Lambda - \mu_S)}{h\Lambda(\gamma + \mu_I + d)(m - n)} \beta_B = R_{0_I} + R_{0_B} \quad (8)$$

onde R_{0_B} é o número reprodutivo básico parcial induzido pela transmissão do meio ambiente para os humanos e R_{0_I} é o número reprodutivo básico parcial induzido pela transmissão entre os humanos.

Considerações Finais

Uma contribuição deste trabalho é a proposta de uso de método matriz da próxima geração básica para determinar a estabilidade da doença de cólera e o cálculo do número básico de reprodução (R_0) de uma forma mais simples e direta usando seus autovalores. Do ponto de vista da saúde pública, diversas situações podem ser analisadas com base no valor de R_0 . Por exemplo, podemos determinar se a doença causará uma epidemia ou pandemia, ou se certas pessoas infectadas introduzidas numa população suscetível à doença não serão reinfectadas e a doença não se espalhará. Outra possibilidade é que o número de infectados aumente a cada geração e a doença se espalhe. Além disso, este método pode ser utilizado para identificar outras doenças infecciosas que podem ser modeladas em estudos epidemiológicos através de sistemas autônomos de equações diferenciais. Assim, R_0 é um parâmetro fundamental em epidemiologia e desempenha um papel crucial na compreensão da propagação de doenças infecciosas numa população. É uma ferramenta valiosa para compreender, prever e controlar a propagação de doenças infecciosas e contribui fortemente na epidemiologia e tomada de decisões em saúde pública.

Referências

- ARAGÓN, M, *et al.* Epidemiologia da cólera em Moçambique no período de 1973-1992. **Revista de Saúde Pública**, v. 28, p. 332-336, 1994;
- CODEÇO, C. T. Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. **BMC Infectious diseases**, v. 1, n. 1, p. 1-14, 2001;
- DIEKMANN, O.; HEESTERBEEK, J. A. P.; METZ, J. A. J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. **Journal of mathematical biology**, v. 28, p. 365-382, 1990;
- HETHCOTE, H. W. The mathematics of infectious diseases. **SIAM review**, v. 42, n. 4, p. 599-653, 2000;
- JOH, R. I, *et al.* Dynamics of indirectly transmitted infectious diseases with immunological threshold. **Bulletin of mathematical biology**, v. 71, p. 845-862, 2009;
- KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. **Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character**, v. 115, n. 772, p. 700-721, 1927;
- MISAU. Ministério da Saúde da República de Moçambique. Direção Nacional de Saúde Pública. Situação da epidemia de Cólera em Moçambique em 2009. Maputo, 2010;
- MISAU. Ministério da Saúde da República de Moçambique. Direção Nacional de Saúde Pública. Manual de Prevenção e Controle da Cólera e de outras doenças Diarreicas agudas. Maputo, 2016;
- MWASA, A.; TCHUENCHE, J. M. Análise matemática de um modelo de cólera com intervenções de saúde pública. **Biosistemas**, v. 105, n. 3, p. 190-200, 2011 ;
- VAN DEN DRIESSCHE, P.; WATMOUGH, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. **Mathematical biosciences**, v. 180, n. 1-2, p. 29-48, 2002.

ESTUDO DE ALGORITMOS DE AGRUPAMENTO APLICADOS A TAXONOMIA E FILOGENIA DE *LEISHMANIA SPP.*

Antonio Revail Alves Pereira

Universidade Federal do Rio de Janeiro

antonio.revail@ppgi.ufrj.br

Marcello Goulart Teixeira

Universidade Federal do Rio de Janeiro

marcellogt@ic.ufrj.br

Flávio Luis de Mello

Universidade Federal do Rio de Janeiro

fmello@poli.ufrj.br

Resumo: O presente trabalho tem como objetivo apresentar resultados de uma proposta alternativa para agrupamento de espécies de *Leishmania* segundo complexos, discutindo as vetorizações utilizadas nos algoritmos de clusterização *k-means*, OPTICS e Fuzzy c-means. Este último algoritmo apresentou resultados mais promissores que os demais, o que sugere seu emprego em processos classificatórios similares.

Palavras-chave: Taxonomia. Filogenia. Aprendizado de Máquina. Clusterização. Lógica Fuzzy

Introdução

A taxonomia biológica é parte da biologia sistemática, que se dedica a análise das relações de parentesco entre organismos. Aristóteles foi o primeiro pensador a classificar seres vivos, mas seu critério não é bem detalhado (RODRIGUES; HIDALGO, 2022). Somente em 1735, Carl Linnaeus propôs uma classificação mais refinada com base em características anatômicas, acreditando que as espécies eram fixas (KLEPKA; CORAZZA, 2018).

No entanto, esse sistema rígido não considerava a evolução. A observação do mundo natural precisa admitir incerteza, instabilidade e mutabilidade porque os fenômenos naturais dificilmente apresentam contornos bem definidos. Tomamos como exemplo o gênero *Leishmania* que ainda é objeto de debate em relação à sua taxonomia.

Os morfotipos foram descritos por Ronald Ross, sendo o nome uma homenagem ao patologista William Boog Leishman (ROSS, 1903) que descreveu pequenos corpos ovais encontrados no baço de um soldado morto após apresentar sintomas como febre alta, disenteria crônica e caquexia.

Várias classificações foram propostas ao longo dos anos, e atualmente (SCHÖNIAN et al., 2018), baseados em dados moleculares, propuseram uma revisão na taxonomia. As espécies analisadas neste trabalho serão as do gênero *Leishmania* subgêneros *Leishmania*, *Viannia* e *Mundinia*.

O desenvolvimento de tecnologias de sequenciamento genético permitiu estudos de filogenia molecular, revelando relações evolutivas entre organismos. Essas revisões de taxonomia são facilitadas pelo uso de algoritmos de aprendizado de máquina, que processam grandes volumes de dados e fornecem resultados de alta qualidade.

Aprendizado de máquina é uma área da Inteligência Artificial que busca desenvolver técnicas computacionais para adquirir conhecimento automaticamente. Algoritmos não supervisionados, como *k-means*, OPTICS e Fuzzy c-means, são usados para resolver problemas de clusterização, incluindo a taxonomia. O Fuzzy c-means, em particular, fornece informações sobre a pertinência dos objetos aos grupos, incorporando imprecisão ao conceito de espécies.

Em resumo, a taxonomia biológica, historicamente baseada em critérios morfológicos, tem evoluído com o uso de dados moleculares e algoritmos de aprendizado de máquina, tornando-se mais flexível e adaptável às mudanças na compreensão da evolução das espécies.

Preparação dos dados

Os dados foram coletados no site do Instituto Nacional de Saúde dos EUA, incluindo sequências de DNA dos cromossomos 10, 28 e 31 de várias espécies de *Leishmania* em formato *.fasta*. Escolheram-se os cromossomos 28 e 31 devido à presença dos genes *hsp70* e *6pg*, característicos da espécie, enquanto o cromossomo 10 foi selecionado aleatoriamente.

O alinhamento foi obtido pelo programa Mauve (DARLING et al., 2004). Para ampliar o conjunto de dados, dado o limitado número de sequências (uma por espécie), foram geradas sequências artificiais. Essas mutações genéticas incluíram substituições de pares de bases, utilizando o modelo de substituição Jukes-Cantor, que pressupõe frequências de equilíbrio iguais para todas as bases e taxas uniformes de mudança nucleotídica (SCHNEIDER, 2003). A taxa de mutação natural foi estimada em até 1% do genoma.

O algoritmo de substituição foi projetado para selecionar aleatoriamente sítios nas sequências de entrada e substituir as bases correspondentes. No total, foram geradas 40 sequências com base nesse modelo de mutação. Esse procedimento ampliado permitiu uma análise mais abrangente das sequências de *Leishmania*.

Vetorização

Após o alinhamento dos cromossomos e a geração de um conjunto expandido de sequências, o próximo passo foi a vetorização, que consiste em representar as bases *A, C, G, T* e *N* por números. Várias abordagens de vetorização foram exploradas. A primeira envolveu a vetorização em cruz $A = (0, 1)$; $C = (1, 0)$; $G = (-1, 0)$; $T = (0, -1)$; $N = (0, 0)$, onde as bases foram associadas a coordenadas no plano cartesiano, mas não trouxe melhorias significativas nos resultados de clusterização.

Outra abordagem foi a vetorização em grau, que representou cada base em um arco de um círculo $A = 73^\circ$, $C = 145^\circ$, $T = 217^\circ$, $G = 289^\circ$, $N = 1^\circ$. No entanto, essa técnica aumentou consideravelmente o tamanho do arquivo de sequências, tornando-o impraticável em termos de tempo computacional e resultando em agrupamentos não representativos.

A vetorização *one-hot* transformou cada base em um grupo de bits $A = (1, 0, 0, 0, 0)$, $T = (0, 1, 0, 0, 0)$, $C = (0, 0, 1, 0, 0)$, $G = (0, 0, 0, 1, 0)$, $N = (0, 0, 0, 0, 1)$, mas o aumento drástico no tamanho do arquivo que inviabilizou seu uso em algoritmos.

O método iCGR (YIN, 2019) representa sequências de DNA por meio de três inteiros. Uma vez que nossas sequências contêm a letra *N*, além das letras representativas das bases, foi necessário adotar um vetor de quatro posições. No entanto, a implementação do iCGR é limitada a sequências com até mil bases. Considerando que nossas sequências consistem de mais de nove milhões de bases, a estratégia de vetorização precisou ser descartada.

A vetorização pelo número de Gödel, que codifica sequências de DNA como números inteiros, foi descartada devido ao tamanho das sequências, que tornava a abordagem impraticável.

O EIIP (NAIR; SREENADHAN, 2006) representa a distribuição das energias dos elétrons livres ao longo de uma sequência de DNA $A = 0,1260$, $C = 0,1340$, $G = 0,0806$, $T = 0,1335$ e, neste caso, consideramos $N = 0$. O tempo de processamento com esta vetorização ficou inviável e não separou bem os grupos.

Finalmente, a vetorização por número atômico, que associou as bases a seus números de prótons (HOLDEN et al., 2007): $A = 70$, $T = 66$, $C = 58$, $G = 78$, $N = 0$, foi a escolhida para a obtenção dos resultados apresentados no artigo. Essa abordagem gerou um arquivo de vetores com tamanho gerenciável

(1.1GB) e proporcionou resultados satisfatórios nos algoritmos de clusterização.

Algoritmos de agrupamento

A clusterização é um processo não supervisionado que agrupa dados com base em suas similaridades, visando alta homogeneidade dentro dos grupos e alta heterogeneidade entre eles. O *k-means* é um algoritmo amplamente utilizado para clusterização (JAIN; MURTY; FLYNN, 1999), mas sua principal desvantagem é a escolha do número de clusters. Para contornar esse problema, foram executados experimentos com diferentes valores de *k* e validados pelo método do cotovelo. No entanto, o *k-means* não produziu resultados satisfatórios, agrupando erroneamente espécies de subgêneros diferentes.

O algoritmo OPTICS (Ordering Points to Identify the Clustering Structure) é baseado em densidade (ANKERST et al., 1999) e tem a vantagem de determinar automaticamente o número de *cluster*. No entanto, seu tempo de processamento e os resultados obtidos levaram à desistência desse algoritmo, uma vez que também agrupou espécies de subgêneros distintos.

O algoritmo *Fuzzy c-means* (FCM) é um algoritmo baseado em lógica *fuzzy* que considera o grau de pertinência de cada ponto do conjunto de dados em relação aos diversos agrupamentos.

Este algoritmo é definido como um programa que gera partições difusas e protótipos para qualquer conjunto de dados numéricos (BEZDEK; EHRICH; FULL, 1984). De forma geral, FCM tenta dividir os dados em conjuntos minimizando a função objetivo abaixo.

$$J = \sum_{i=1}^n \sum_{j=1}^p \mu_{ij}^m d(x_i : c_j)^2 \quad (1)$$

onde μ_{ij} é o grau de pertinência da amostra x_i ao j -ésimo *cluster*, n é o número de instâncias, p é o número de *clusters* considerados no FCM (que deve ser decidido antes da execução), $m > 1$ é o parâmetro Fuzzyficador, x_i um vetor de dados de treinamento onde $i = 1, 2, \dots, n$ representa um atributo do dado, c_j , onde $j = 1, 2, \dots, p$, é o centro de um agrupamento *fuzzy* e $d(x_i : c_j)$ é a distância entre x_i e c_j .

Resultados e discussões

A taxonomia dos subgêneros (*Leishmania*), (*Viannia*) e (*Mundinia*) apresentada por (Akhoundi et al., 2017) é a referência a que foram baseados os resultados. Neste trabalho, os resultados do *k-means* e OPTICS não foram incorporados, pois esses métodos agruparam espécies de subgêneros diferentes em um mesmo grupo ou não trouxeram informações novas em relação à taxonomia existente.

O algoritmo *Fuzzy c-means* foi empregado para alcançar o resultado principal, aplicando-o a 50 sequências vetorizadas por número atômico. A associação de objetos aos *clusters* ocorre por meio da função de pertinência, enquanto a validação desses *clusters* é realizada por meio do coeficiente de partição *fuzzy*, que varia de 0 a 1, com 1 representando o melhor resultado. Durante a execução do programa, uma variação de 2 a 10 *clusters* foi considerada.

A Tabela 1 mostra coeficientes de partição *fuzzy* variando de 0,40785 a 0,87481, usados para avaliar a qualidade da clusterização pelo *Fuzzy c-means*. O agrupamento com 10 *clusters* alcançou o valor máximo, mas não é ideal, pois separaria espécies como *L.donovani* e *L.infantum*, que na literatura estão agrupadas.

O agrupamento com sete *clusters*, além de apresentar o segundo maior valor de coeficiente de partição *fuzzy*, é a clusterização mais próxima da representação taxonômica clássica. O resultado com sete *clusters fuzzy* também manteve a divisão em sete complexos, porém com modificações no nível de espécies que formam os complexos.

Tabela 1: Tabela com valores dos coeficientes de partição fuzzy.

Quantidade de clusters	Coeficiente de Partição
2	0,52483
3	0,49749
4	0,43940
5	0,56039
6	0,61419
7	0,68551
8	0,44459
9	0,40785
10	0,87481

Na primeira tabela da Tabela 2 temos as espécies distribuídas nos sete *clusters*. As espécies dos *clusters* 1, 2 e 4 aparecem agrupadas com outras espécies, diferente do que ocorre na taxonomia convencional e mantendo a coerência em relação aos subgêneros.

A sexta coluna da Tabela 2 apresenta o *cluster* onde as espécies de *L.aethiopica* estão agrupadas, isto é verificado pelo maior valor das linhas. Os valores menores são os de pertinência da espécie em relação aos outros *clusters*.

Tabela 2: Tabelas com agrupamento e graus de pertinências.

Agrupamento com 7 clusters							
Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	
<i>L.braziliensis</i>	<i>L.major</i>	<i>L.peruviana</i>	<i>L.donovani</i>	<i>L.enriettii</i>	<i>L.aethiopica</i>	<i>L.arabica</i>	
<i>L.panamensis</i>	<i>L.mexicana</i>		<i>L.infantum</i>				

Espécies	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7
<i>L.aethiopica base</i>	0,002823538	0,001894672	0,002104146	0,003542764	0,003195455	0,982933219	0,003506206
<i>L.aethiopica 1</i>	0,015929514	0,010782328	0,011964421	0,019897718	0,017907856	0,903854268	0,019663894
<i>L.aethiopica 2</i>	0,003665983	0,002461349	0,002733374	0,00459825	0,004147011	0,977843394	0,00455064
<i>L.aethiopica 3</i>	0,008519494	0,005738579	0,006370907	0,010668743	0,009614474	0,948536541	0,010551261
<i>L.aethiopica 4</i>	0,010558704	0,007121699	0,007904724	0,01321117	0,011903164	0,936235029	0,013065509

A escolha do algoritmo FCM permite expressar situações em que um objeto compartilha similaridade com múltiplos grupos, refletindo complexas relações. Os graus de pertinência indicam que as espécies têm algum grau de pertencimento aos diversos complexos, revelando uma origem comum entre as espécies. Isso sugere que a definição de uma espécie não deve ser estritamente rígida, mas sim considerar graus de identidade entre diferentes espécies, uma representação mais fiel da evolução contínua. Esses valores também permitem quantificar a proximidade ou afastamento entre as espécies e em relação a cada complexo.

Conclusões

Este estudo aplicou o algoritmo Fuzzy c-means (FCM) para agrupar espécies do gênero *Leishmania*, subgêneros *Leishmania*, *Viannia* e *Mundinia*, usando a vetorização por número atômico das sequências de DNA. Os resultados flexíveis do FCM permitiram atribuir graus de pertinência a diferentes complexos, refletindo a complexidade das relações evolutivas entre as espécies. Isso destaca a importância de considerar a variabilidade genética e a evolução contínua na taxonomia, em contraste com abordagens tradicionais mais rígidas.

Além disso, o FCM oferece uma perspectiva promissora para aprimorar a compreensão da diversidade biológica e evolutiva em estudos de taxonomia molecular, possibilitando a quantificação mais precisa da proximidade entre espécies e complexos. Sendo assim, sugere-se o aprofundamento dos estudos de utilização do FCM para a classificação de espécies.

Referências

- AKHOUNDI, M.; DOWNING, T.; VOT ‘YPKA, J.; KUHLS, K.; LUKEŠ, J.; CANNET, A.; RAVEL, C.; MARTY, P.; DELAUNAY, P.; KASBARI, M. et al. Leishmania infections: Molecular targets and diagnosis. **Elsevier**, 2017.1–29 p.
- ANKERST, M.; BREUNIG, M. M.; KRIEGEL, H.-P.; SANDER, J. OPTICS: Ordering points to identify the clustering structure. **ACM New York**, NY, USA, 1999. 49–60 p.
- BEZDEK, J. C.; EHRLICH, R.; FULL, W. FCM: The fuzzy c-means clustering algorithm. **Elsevier**, 1984. 191–203 p.
- DARLING, A. C.; MAU BOBAND BLATTNER, F. R.; PERNA, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **ColdSpring Harbor Lab**, 2004. 1394–1403 p.
- HOLDEN, T.; SUBRAMANIAM, R.; SULLIVAN, R.; CHEUNG, E.; SCHNEIDER, C.; JR, G. T.; FLAMHOLZ, A.; LIEBERMAN, D.; CHEUNG, T. ATCG nucleotide fluctuation of *Deinococcus radiodurans* radiation genes. **In: Instruments, Methods, and Missions for Astrobiology X. SPIE**, 2007. 669417 p.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **Acm New York**, NY, USA, 1999. 264–323 p.
- KLEPKA, V.; CORAZZA, M. J. **O essencialismo na classificação de Lineu e a repercussão dessa controvérsia na Biologia**. História da Ciência e Ensino: construindo interfaces, 2018. 73–110 p.
- NAIR, A. S.; SREENADHAN, S. P. A coding measure scheme employing electron-ion interaction pseudopotential (EIIP). **Biomedical Informatics Publishing Group**, 2006. 197 p.
- NUNES, D. H. F. **Um breve estudo sobre o algoritmo K-means**. 2016. – Universidade de Coimbra, Dissertação (Mestrado), 2016.
- RODRIGUES, M. de P.; HIDALGO, J. M. A classificação dos animais segundo Aristóteles: recorte histórico e inserção didática. **Filosofia e História da Biologia**, v. 17, n. 2, p. 195-218, 2022.
- ROSS, R. Further notes on Leishman’s bodies. **BMJ Publishing Group**, 1903. 1401 p.
- SCHNEIDER, H. Métodos de análise filogenética: um guia prático. **Holos**, 2003.
- SCHÖNIAN, G.; CUPOLILLOO, E.; MAURICIO, I. **Molecular Evolution and Phylogeny of Leishmania**. Springer International Publishing, 2018.
- YILMAZ, A. Assessment of mutation susceptibility in DNA sequences with word vectors. **Journal of Intelligent Systems: Theory and Applications**, v. 3, n. 1, p. 1-6, 2020.
- YIN, C. Encoding and decoding DNA sequences by integer chaos game representation. **Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New**, 2019. 143–151 p.

MODELOS GENOTÍPICOS DE POLIMORFISMO BASEADOS NO JOGO PEDRA-PAPEL-TESOURA

Marcelo Cargnelutti Rossato
Universidade Estadual de Campinas
m264880@dac.unicamp.br

Marcus Aloizio Martinez de Aguiar
Universidade Estadual de Campinas
aguiar@ifi.unicamp.br

Resumo: Nas populações de lagartos *Uta stansburiana* ocorre um notável polimorfismo de coloração nas gargantas, com tons de laranja, amarelo e azul. Esses fenótipos são determinados por dois alelos que correspondem, cada um, a uma das possíveis cores e seguem uma relação de dominância. Dois modelos, um não espacial e outro espacial, foram desenvolvidos para analisar as interações comportamentais desses fenótipos por meio de um jogo pedra-papel-tesoura e levando em conta também os possíveis genótipos. Com o modelo não espacial foram encontrados cenários em que a coexistência dos três fenótipos era estável e em que era instável, enquanto os resultados do modelo espacial foram próximos desses equilíbrios de coexistência. Notou-se ainda uma forte influência do tamanho da vizinhança nos resultados do modelo espacial, no qual vizinhanças menores favoreceram os lagartos laranjas, cujo alelo é dominante em relação aos demais, e vizinhanças maiores geraram resultados menos oscilatórios e mais próximos do equilíbrio do modelo não espacial.

Palavras-chave: Polimorfismo. Lagarto-de-mancha-lateral. Teoria dos jogos. Dinâmica de populações. Modelos discretos.

Introdução

Os lagartos da espécie *Uta stansburiana*, popularmente conhecidos como lagartos-de-mancha-lateral, apresentam um intrigante exemplo de polimorfismo, que representa a ocorrência de múltiplos fenótipos dentro de uma mesma espécie (Ford, 1945), de coloração em sua população, que é objeto de estudo e pesquisa em ecologia evolutiva. Nesse caso específico, os lagartos apresentam variações de cor na região da garganta em tons de laranja, amarelo e azul.

A coloração das gargantas dos lagartos *Uta stansburiana* é determinada por dois alelos, que podem ser *o* (laranja), *y* (amarelo) ou *b* (azul), que interagem de acordo com princípios de dominância. Para os lagartos machos, o alelo *o* é dominante em relação aos demais e o alelo *y* é dominante em relação ao alelo *b*, portanto os indivíduos com genótipo *oo*, *oy* e *ob* apresentam um fenótipo de garganta laranja, enquanto *yy* e *yb* são lagartos com garganta amarela e *bb* representa o fenótipo azul. Já as fêmeas apresentam uma garganta laranja quando possuem pelo menos um alelo *o* e garganta amarela nos demais casos.

Por simplicidade, definiremos cada lagarto pela cor de sua garganta. Os lagartos laranjas geralmente são mais agressivos e dominam territórios maiores, os azuis costumam defender territórios menores e com menos fêmeas e os amarelos se camuflam como fêmeas para se infiltrar nos territórios dos demais lagartos (SINERVO et al., 2000).

Devido às características específicas de cada fenótipo, nota-se que a dominância dos indivíduos laranjas é derrotada pela furtividade dos amarelos, enquanto estes perdem para a estratégia defensiva dos azuis, que são derrotados pela estratégia dos lagartos laranjas (SINERVO; LIVELY, 1996). Essas estratégias comportamentais formam um ciclo de vantagem competitiva semelhante a um jogo de pedra-papel-tesoura, onde cada fenótipo prevalece sobre o outro em um padrão cíclico.

Embora a teoria dos jogos tenha sido inicialmente utilizada para descrever o comportamento humano por meio de decisões racionais, ela vem cada vez mais sendo utilizada para analisar conflito e

cooperação em plantas e animais (HAMMERSTEIN; SELTEN, 1994), como é o caso dos lagartos-de-mancha-lateral.

Portanto, o objetivo desse trabalho é estudar as interações entre lagartos da espécie *Uta stansburiana* por meio da construção de dois modelos discretos, que levam em consideração tanto a dinâmica pedra-papel-tesoura dos diferentes fenótipos quanto as relações de dominância dos alelos que compõem os genótipos de cada indivíduo.

Modelo não espacial

De forma semelhante ao que foi realizado por Barreto (2017), definimos $f_{O,t}$, $f_{Y,t}$ e $f_{B,t}$ as proporções fenotípicas de lagartos laranjas, amarelos e azuis, respectivamente, no instante de tempo t . Assim, determinamos também as proporções genotípicas $g_{oo,t}$, $g_{oy,t}$, $g_{ob,t}$, $g_{yy,t}$, $g_{yb,t}$ e $g_{bb,t}$ referentes a cada um dos possíveis genótipos. Os indivíduos com pelo menos um alelo o apresentam fenótipo laranja, os genótipos yy e yb são de lagartos amarelos e o fenótipo azul é representado unicamente pelo genótipo bb . As relações dessas proporções são dadas na Equação (1),

$$f_{O,t} = g_{oo,t} + g_{oy,t} + g_{ob,t}, \quad f_{Y,t} = g_{yy,t} + g_{yb,t}, \quad f_{B,t} = g_{bb,t}. \quad (1)$$

Podemos determinar as proporções alélicas $a_{o,t}$, $a_{y,t}$ e $a_{b,t}$, referentes às proporções dos alelos o , y e b na população, através da Equação (2),

$$a_{o,t} = g_{oo,t} + \frac{1}{2}g_{oy,t} + \frac{1}{2}g_{ob,t}, \quad a_{y,t} = g_{yy,t} + \frac{1}{2}g_{oy,t} + \frac{1}{2}g_{yb,t}, \quad a_{b,t} = g_{bb,t} + \frac{1}{2}g_{ob,t} + \frac{1}{2}g_{yb,t}. \quad (2)$$

Por se tratarem de proporções, cabe salientar que em todo instante de tempo t valem as relações $\sum_I f_{I,t} = \sum_{i,j} g_{ij,t} = \sum_i a_{i,t} = 1$, no qual $I \in \{O, Y, B\}$ e $i, j \in \{o, y, b\}$. As interações entre lagartos de diferentes fenótipos são regidas por uma matriz de recompensas M , que é exemplificada na Tabela 1, onde o elemento M_{IJ} representa a recompensa que um indivíduo que utiliza a estratégia I ganha ao interagir com um indivíduo de estratégia J . Nesse caso, as estratégias correspondem aos fenótipos dos lagartos, portanto $I, J \in \{O, Y, B\}$.

Tabela 1: Matriz de recompensas.

	O	Y	B
O	1	0.5	2
Y	2	1	0.5
B	0.5	2	1

Como os lagartos laranjas possuem vantagem quando os azuis são comuns e desvantagem quando os amarelos são comuns, notamos que $M_{OB} > M_{OO} > M_{OY}$. De forma análoga, temos $M_{YO} > M_{YY} > M_{YB}$ e $M_{BY} > M_{BB} > M_{BO}$, de forma semelhante ao que ocorre no jogo pedra-papel-tesoura. Para cada instante de tempo t , o *fitness* $W_{ij,t}$ de um genótipo ij , que apresenta um fenótipo I , e o *fitness* médio W_t da população são determinados pela Equação (3),

$$W_{ij,t} = W_{I,t} = \sum_J M_{IJ} f_{J,t}, \quad W_t = \sum_I f_{I,t} W_{I,t}. \quad (3)$$

Para determinar as proporções da geração seguinte, definimos então $\tilde{g}_{ij,t}$ como variáveis auxiliares de acordo com a Equação (4) com o intuito de aumentar as proporções dos genótipos com maior *fitness*,

$$\tilde{g}_{ij,t} = g_{ij,t} \frac{W_{ij,t}}{W_t} \quad (4)$$

Em seguida, utilizamos esses valores para calcular as proporções alélicas da próxima geração. Assim,

$$a_{o,t+1} = \tilde{g}_{oo,t} + \frac{1}{2}\tilde{g}_{oy,t} + \frac{1}{2}\tilde{g}_{ob,t}, \quad a_{y,t+1} = \tilde{g}_{yy,t} + \frac{1}{2}\tilde{g}_{oy,t} + \frac{1}{2}\tilde{g}_{yb,t}, \quad a_{b,t+1} = \tilde{g}_{bb,t} + \frac{1}{2}\tilde{g}_{ob,t} + \frac{1}{2}\tilde{g}_{yb,t}.$$

Finalmente, as proporções genótípicas da geração seguinte são determinadas pela Equação (5),

$$g_{ij,t+1} = \begin{cases} (a_{i,t+1})^2, & \text{se } i = j \\ 2 \cdot a_{i,t+1} \cdot a_{j,t+1}, & \text{se } i \neq j \end{cases} \quad (5)$$

Analisando a Equação (4), pode-se concluir que os pontos de equilíbrio desse modelo genotípico não espacial são tais que, para cada $i, j \in \{o, y, b\}$, temos $\bar{g}_{ij} = 0$ ou $W_{ij,t} = W_t$. Dessa forma, para o equilíbrio de coexistência é necessário que $W_{O,t} = W_{Y,t} = W_{B,t} = W_t$. Para a matriz de recompensas apresentada na Tabela 1, isso resulta em um equilíbrio trivial em que $\bar{f}_O = \bar{f}_Y = \bar{f}_B = \frac{1}{3}$.

Para gerar resultados mais interessantes, podemos alterar a matriz de recompensas para focar em cada uma das estratégias. Dessa forma, na Tabela 2 são apresentadas essas matrizes de recompensas, onde $0 < d < 1$ e $v > 1$ representam as recompensas que um indivíduo com a estratégia focal recebe ao interagir com outro indivíduo com quem ele possui uma desvantagem ou uma vantagem, respectivamente.

Tabela 2: Matrizes de recompensas com estratégias focais O, Y e B.

(a) Estratégia focal O				(b) Estratégia focal Y				(c) Estratégia focal B			
	O	Y	B		O	Y	B		O	Y	B
O	1	d	v	O	1	0.5	2	O	1	0.5	2
Y	2	1	0.5	Y	v	1	d	Y	2	1	0.5
B	0.5	2	1	B	0.5	2	1	B	d	v	1

Dessa forma, as proporções fenotípicas dos pontos de equilíbrio de coexistência para as estratégias focais O, Y e B são apresentados, respectivamente, nas Equações (6), (7) e (8),

$$\bar{f}_O = \frac{4v - 2d}{10v - 8d + 5}, \quad \bar{f}_Y = \frac{6v - 5}{10v - 8d + 5}, \quad \bar{f}_B = \frac{10 - 6d}{10v - 8d + 5}. \quad (6)$$

$$\bar{f}_O = \frac{10 - 6d}{10v - 8d + 5}, \quad \bar{f}_Y = \frac{4v - 2d}{10v - 8d + 5}, \quad \bar{f}_B = \frac{6v - 5}{10v - 8d + 5}. \quad (7)$$

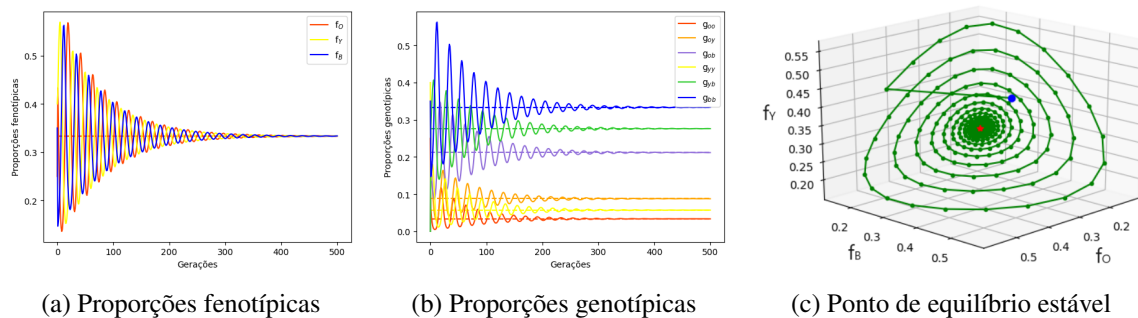
$$\bar{f}_O = \frac{6v - 5}{10v - 8d + 5}, \quad \bar{f}_Y = \frac{10 - 6d}{10v - 8d + 5}, \quad \bar{f}_B = \frac{4v - 2d}{10v - 8d + 5}. \quad (8)$$

Tomando $v = 2$ e $d = 0.5$ com qualquer uma das estratégias focais, obtemos o equilíbrio trivial $\bar{f}_O = \bar{f}_Y = \bar{f}_B = \frac{1}{3}$. Para todos os casos, as proporções alélicas e genótípicas do ponto de equilíbrio são apresentadas, respectivamente, nas Equações (9) e (10),

$$\bar{a}_o = 1 - \sqrt{\bar{f}_B + \bar{f}_Y}, \quad \bar{a}_y = \sqrt{\bar{f}_B + \bar{f}_Y} - \sqrt{\bar{f}_B}, \quad \bar{a}_b = \sqrt{\bar{f}_B}. \quad (9)$$

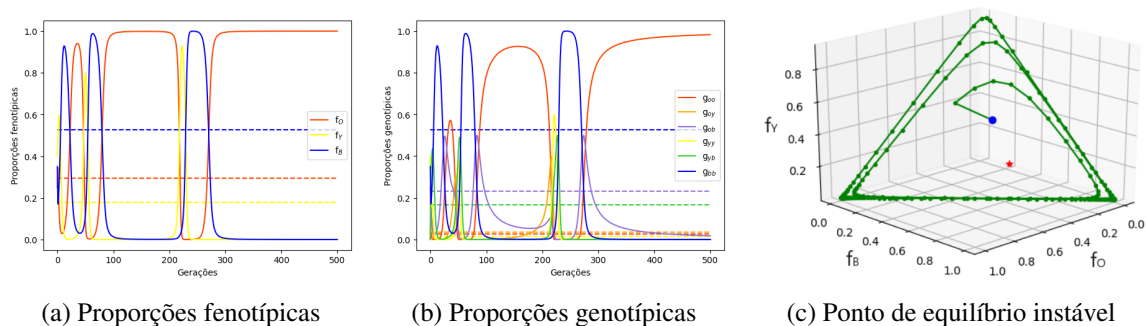
$$\bar{g}_{ij} = \begin{cases} (\bar{a}_{i,t})^2, & \text{se } i = j \\ 2 \cdot \bar{a}_{i,t} \cdot \bar{a}_{j,t}, & \text{se } i \neq j \end{cases} \quad (10)$$

Para ilustrar esse modelo, serão realizadas duas simulações utilizando a matriz de recompensas com estratégia focal O, representada na Tabela 2a, com 500 gerações cada e condições iniciais $g_{oo,0} = 0.25$, $g_{yy,0} = 0.4$, $g_{bb,0} = 0.35$ e $g_{oy,0} = g_{ob,0} = g_{yb,0} = 0$. Os resultados obtidos com $v = 2$ e $d = 0.5$ podem ser observados na Figura 1.


 Figura 1: Simulações com $\nu = 2$ e $d = 0.5$.

Nota-se que o sistema oscila, mas converge para o ponto de equilíbrio representado pelas retas tracejadas nas Figuras 1a e 1b e pela estrela vermelha na Figura 1c. Assim, o ponto de equilíbrio de coexistência com esses parâmetros é estável.

Enquanto isso, utilizando as mesmas condições iniciais e tomando $\nu = 1.3$ e $d = 0.3$, encontra-se um sistema que não converge para o ponto de equilíbrio de coexistência. O resultado de 500 gerações é apresentado na Figura 2, onde o sistema se aproxima mais dos equilíbrios em que há apenas um fenótipo de lagarto, convergindo por fim para o equilíbrio em que há apenas o genótipo oo .


 Figura 2: Simulações com $\nu = 1.3$ e $d = 0.3$.

Modelo espacial

Este modelo é constituído por um autômato celular bidimensional, que é uma abordagem que integra elementos de tempo, espaço e estados discretos (Mistro, 1998). Podemos visualizar esse modelo como uma matriz de dimensões $N \times N$, onde cada elemento, referido como sítio ou célula, pode assumir um conjunto limitado de valores, que indicam que ele pode estar vazio ou ocupado por um lagarto com um dos genótipos oo , oy , ob , yy , yb ou bb .

Consideramos uma densidade inicial p_0 , que determina a probabilidade de cada sítio estar inicialmente ocupado. Nessa primeira distribuição, consideramos apenas indivíduos homozigotos, então cada sítio ocupado terá probabilidades p_O , p_Y ou p_B de conter um indivíduo com genótipo oo , yy ou bb , respectivamente. Além disso, todo indivíduo pode morrer com probabilidade μ no início de cada geração, exceto a primeira, deixando o sítio vazio.

Para simular as interações entre vizinhos, é necessário definir as vizinhanças de interação. Assim, utilizamos as vizinhanças de Von Neumann, de Moore e de Moore estendida, ilustradas nas Figuras 3a, 3b e 3c, onde os sítios vermelhos representam os vizinhos do sítio azul. Além disso, são utilizadas condições de contorno periódicas.

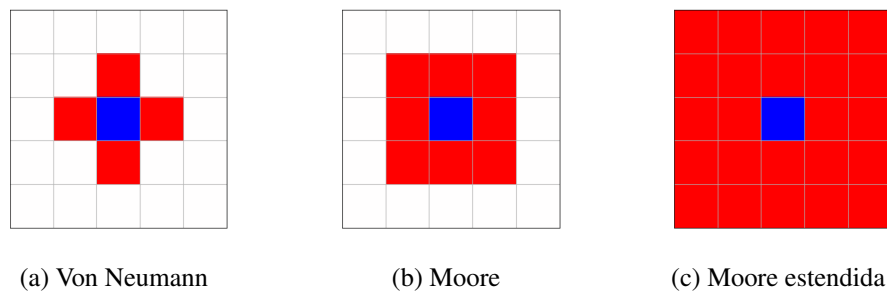


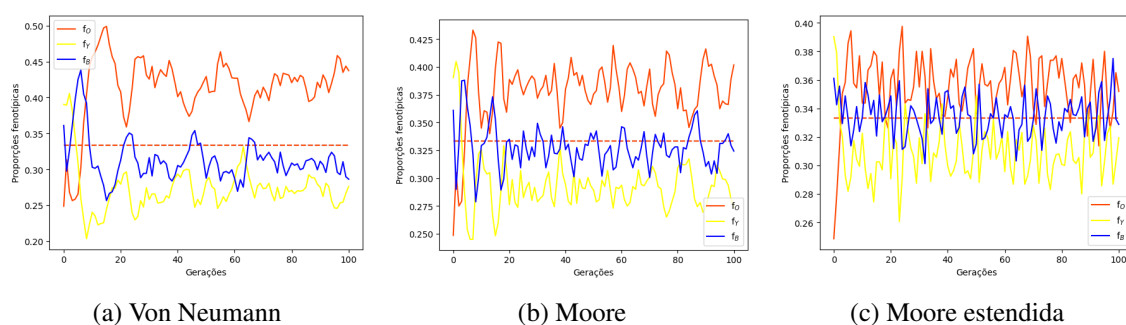
Figura 3: Vizinhanças.

Em cada geração, para que cada lagarto possa ter em média uma oportunidade de reprodução, são sorteados N^2 sítios. Em cada sorteio, se houver um lagarto no sítio escolhido, ele calcula o seu *fitness* e de cada um de seus vizinhos. O cálculo é realizado fazendo a média das recompensas obtidas nas interações com os sítios vizinhos não vazios com base na Tabela 2a. Por exemplo, ao considerar a vizinhança de Von Neumann, se um indivíduo amarelo tiver como vizinhos dois lagartos laranjas, um azul e uma célula vazia, seu *fitness* será $\frac{2+2+0.5}{3} = 1.5$.

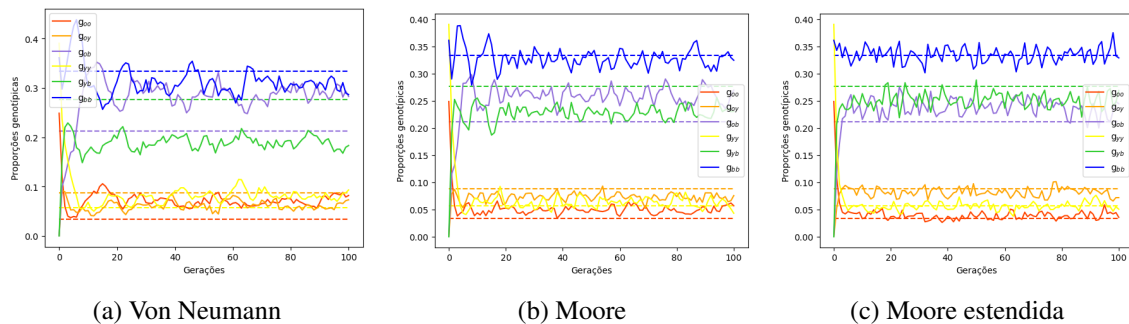
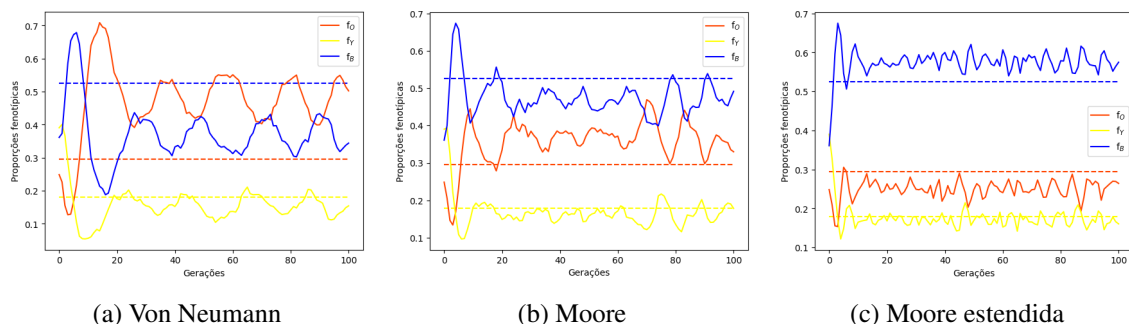
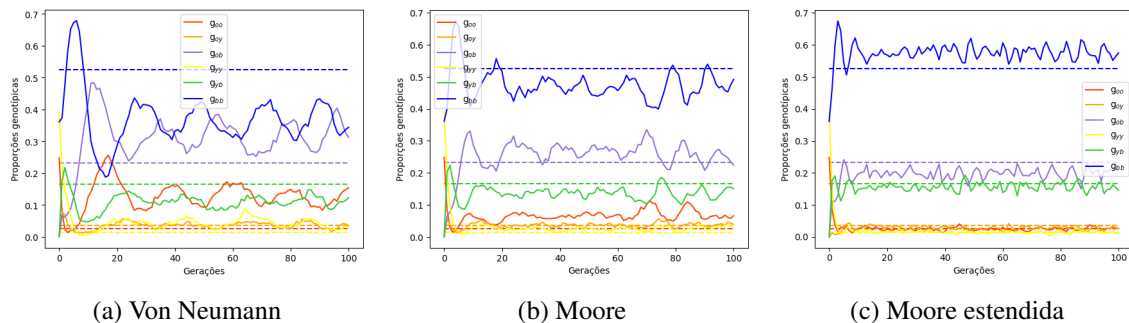
Sempre que houver pelo menos um outro lagarto na vizinhança do indivíduo sorteado, ele irá se reproduzir. Se o seu *fitness* for maior ou igual do que o de todos os seus vizinhos, ele sorteia aleatoriamente um lagarto da sua vizinhança para reproduzir. Caso contrário, a reprodução se dará com o lagarto de maior *fitness* da vizinhança (se houver mais de um, escolhe-se um aleatoriamente).

O genótipo do filho será composto por um alelo de cada um dos pais, portanto se os pais tiverem genótipos *bb* e *oy*, por exemplo, o primeiro obrigatoriamente fornecerá um alelo *b* enquanto o segundo pode fornecer um alelo *o* ou *y*. Notamos, nesse caso, que um indivíduo azul e um laranja, de genótipos *bb* e *oy*, podem gerar um filho com fenótipo amarelo, com genótipo *yb* por exemplo. Se houver células vazias na vizinhança de um dos dois indivíduos que estão se reproduzindo, o filho ocupará um desses sítios. Caso contrário, o filho ocupará o sítio do indivíduo de menor *fitness* entre os dois.

Serão realizadas simulações com os mesmos parâmetros utilizados no modelo não espacial, com estratégia focal *O* e probabilidades iniciais $p_O = 0.25$, $p_Y = 0.4$ e $p_B = 0.35$. Além disso, definimos $N = 40$, $p_0 = 0.9$ e $\mu = 0.1$. As proporções fenotípicas e genotípicas ao longo de 100 gerações com $v = 2$ e $d = 0.5$ para as diferentes vizinhanças são apresentadas, respectivamente, nas Figuras 4 e 5.


 Figura 4: Proporções fenotípicas para $v = 2$ e $d = 0.5$.

É possível observar que para a vizinhança de Von Neumann a proporção de indivíduos laranjas é muito maior do que a do ponto de equilíbrio do modelo não espacial, o que pode ser justificado ao considerar que o alelo *o* é dominante em relação aos demais. Por outro lado, ao aumentar o tamanho da vizinhança, nota-se que os resultados se aproximam cada vez mais do equilíbrio do modelo não espacial e apresentam oscilações menores, como se observa principalmente na Figura 5. Nas Figuras 6 e 7 são apresentados os resultados das simulações com $v = 1.3$ e $d = 0.3$ para as diferentes vizinhanças.


 Figura 5: Proporções genotípicas para $v = 2$ e $d = 0.5$.

 Figura 6: Proporções fenotípicas para $v = 1.3$ e $d = 0.3$.

 Figura 7: Proporções genotípicas para $v = 1.3$ e $d = 0.3$.

De forma semelhante ao que foi observado anteriormente, pode-se notar que os lagartos laranjas novamente possuem uma vantagem nas vizinhanças menores, mas o aumento da vizinhança contribuiu para que os resultados se aproximem do ponto de equilíbrio de coexistência do modelo não espacial, mesmo com esse ponto sendo instável para o primeiro modelo.

Para finalizar, seria interessante analisar quão próximos os resultados obtidos por ambos os modelos propostos seriam ao tentar desconsiderar as dependências espaciais desse segundo modelo. Assim, na Figura 8 são apresentados os resultados das 20 primeiras gerações das simulações utilizando uma vizinhança infinita, ou seja, onde os *fitnesses* dos lagartos são comparados com os de todos os outros indivíduos no momento da reprodução.

Enfim, conclui-se que os efeitos observados nas Figuras 4, 5, 6 e 7 são causados de fato pelo tamanho da vizinhança. Utilizando uma vizinhança infinita, as proporções fenotípicas se aproximaram rapidamente do ponto de equilíbrio do modelo não espacial, mas notou-se uma redução no número de indivíduos com genótipos *oo* e *yy*, que foi compensada com o crescimento dos genótipos *ob* e *yb*.

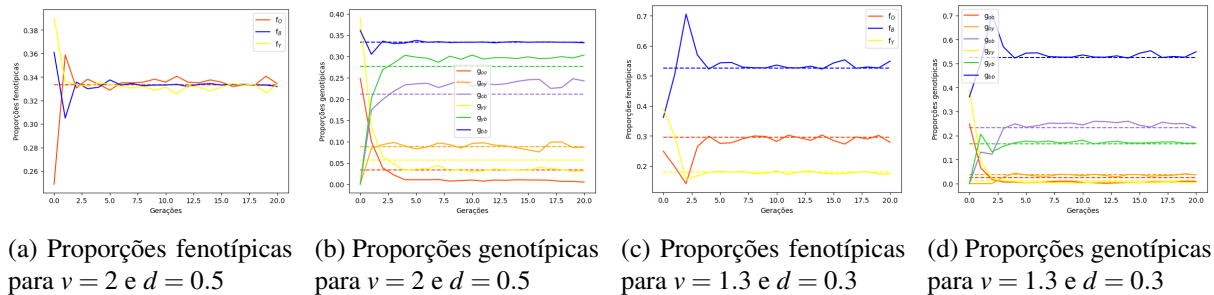


Figura 8: Simulações com vizinhança infinita.

Conclusões

Nesse trabalho foram apresentados dois modelos discretos genotípicos para descrever a dinâmica dos lagartos da espécie *Uta stansburiana* por meio de uma representação semelhante a um jogo pedra-papel-tesoura. Com o modelo não espacial, encontramos uma situação em que o equilíbrio de coexistência das espécies é estável e outra em que a coexistência é instável e os resultados se aproximam do equilíbrio em que todos os indivíduos possuem genótipo oo .

Por meio do modelo espacial, observamos que o tamanho da vizinhança influencia muito os resultados, favorecendo o fenótipo laranja, cujo alelo é dominante em relação aos demais, quando a vizinhança é pequena. Ao aumentar a vizinhança, observa-se que os resultados das simulações, principalmente no que se refere aos fenótipos, se aproximam cada vez mais do equilíbrio de coexistência encontrado para o modelo não espacial, independentemente de ele ser um equilíbrio estável ou instável no primeiro modelo.

Além disso, pode-se notar que o aumento da vizinhança reduz consideravelmente as oscilações das proporções fenotípicas e genotípicas ao longo das gerações. Por fim, as proporções fenotípicas das simulações com vizinhança infinita convergiram rapidamente para o equilíbrio do modelo não espacial, enquanto as proporções genotípicas se aproximaram desse equilíbrio, mas os genótipos heterozigotos com um alelo b se sobressaíram um pouco enquanto os homozigotos oo e yy chegaram a valores um pouco menores do que no outro modelo.

Referências

- BARRETO, W. P. **Teoria de jogos e genética de populações: um estudo de dinâmica de populações guiado por competição e cooperação entre indivíduos polimórficos**. 2017. 75 f. Tese (Doutorado em Física) – Unicamp, Campinas, 2017.
- FORD, E. B. Polymorphism. **Biological Reviews**, v. 20, n. 2, p. 73-88, 1945.
- HAMMERSTEIN, P.; SELTEN, R. Game theory and evolutionary biology. **Handbook of game theory with economic applications**, v. 2, p. 929-993, 1994.
- MISTRO, D. C. **Modelos para dispersão de abelhas africanizadas: um zoom matemático**. 1998. 182 f. Tese (Doutorado em Matemática Aplicada) – Unicamp, Campinas, 1998.
- SINERVO, B.; LIVELY, C. M. The rock–paper–scissors game and the evolution of alternative male strategies. **Nature**, v. 380, n. 6571, p. 240-243, 1996.
- SINERVO, B. et al. Testosterone, endurance, and Darwinian fitness: natural and sexual selection on the physiological bases of alternative male behaviors in side-blotched lizards. **Hormones and Behavior**, v. 38, n. 4, p. 222-233, 2000.

UM MODELO PARA A COMPETIÇÃO ENTRE O MEXILHÃO DOURADO E SUA VERSÃO GENETICAMENTE MODIFICADA

Charles Henrique Xavier Barreto Barbosa
Universidade Federal Rural do Rio de Janeiro
charles.hxbb@gmail.com

Claudia Mazza Dias
Universidade Federal Rural do Rio de Janeiro - UFRRJ
mazza@ufrrj.br

Dayse Haime Pastore
Centro Federal de Educação Tecnológica - CEFET-RJ
dayse.pastore@cefet-rj.br

Resumo: A presença do invasor mexilhão dourado em águas brasileiras representa uma sorte de impactos ecológicos ou econômicos. Diante desse cenário, uma alternativa de controle populacional é promissora: a imersão de uma espécie geneticamente modificada com o objetivo de, em contato com o molusco selvagem, produzir futuras gerações inférteis. Uma proposta de modelo matemático que contemple tal competição torna-se de grande valia, uma vez que pode fornecer informações para nortear experimentos de laboratório. No entanto, o modelo de competição necessita de uma dinâmica de crescimento bem estabelecida. Esse trabalho apresenta dois modelos para a dinâmica do mexilhão selvagem, estudados de forma analítica. O estudo expõe as soluções de equilíbrio, sua estabilidade local nos pontos estacionários e uma análise de sensibilidade do números de reprodução basal (R_0 's) para ambos os modelos. Através desses resultados analíticos, define-se o mais adequado para receber as parcelas que modelem a competição. São realizadas simulações numéricas a fim de aferir possíveis desfechos da competição, fazendo uso do método de Runge-Kutta de quarta ordem. O estudo é capaz de apresentar dois pontos conclusivos importantes: que através da competição planejada a erradicação pode ser atingida e que há necessidade de vantagens competitivas para que a estratégia funcione, indicando que uma maior eficiência reprodutiva e um mexilhão modificado menos suscetível à predação são as vantagens mais promissoras. Finalmente, à vista de um tópico tão importante, esse trabalho é necessário para que essa estratégia de controle populacional seja estudada e testada para auxiliar no desenvolvimento da espécie modificada e respeitar a ecologia.

Palavras-chave: Modelagem Matemática, Modelo de Competição, Dinâmica Populacional, Análise de Estabilidade, Biomatemática.

Introdução

O equilíbrio de um ecossistema é definido pelas contínuas adaptações das espécies que nele vivem. A invasão de uma nova espécie tem o potencial de perturbar esse meio, através de uma competição por recursos com espécies nativas. Quando os invasores não encontram pedradores naturais e mostram-se resilientes ao meio, desequilibram a cadeia trófica e acarretam danos à biodiversidade e às atividades econômicas (BOLTOVSKOY *et al.*, 2006; SOUZA *et al.*, 2009; RICCIARDI, 1998). Este trabalho estuda a espécie invasora *Limnoperna fortunei* (Dunker, 1857), conhecida como mexilhão dourado. Trata-se de um molusco bivalve e de água doce que foi registrado pela primeira vez na América Latina, na planta industrial de Atucha em 1996 (BOLTOVSKOY *et al.*, 2006). A reprodução do molusco é externa, dá-se após o encontro dos gametas na água, e em sequência, inicia-se a fase larval, que pode durar até 30 dias (CATALDO *et al.*, 2005). Com o fim da fase larval, o mexilhão se estabelece como recruta e busca um substrato liso e sólido para aderência. Com frequência, ambientes antrópicos oferecem este tipo ideal de substrato, implicando em problemas de coexistência (CATALDO *et al.*, 2005; RICCIARDI, 1998;

BOLTOVSKOY *et al.*, 2006). Como as larvas são direcionadas pelo fluxo de corrente, são admitidas com facilidade nas plantas industriais hidrelétricas, onde encontram ambiente de constante fluxo de água, o que facilita a filtração, e com substratos antrópicos para aderência (SILVA *et al.*, 2016; BOLTOVSKOY *et al.*, 2015). Consequentemente, a presença de altas densidades do mexilhão nas plantas hidroelétricas, principalmente em tubulações, acarreta em perdas de carga e frequentes paradas para limpeza (SIMEÃO *et al.*, 2011). Ademais, além das usinas hidrelétricas, a alta densidade populacional desequilibra os níveis de nutrientes do meio, fitoplâncton e microalgas, atingindo a cadeia trófica e prejudicando a atividade de pesca (BOLTOVSKOY *et al.*, 2015; SILVA *et al.*, 2016).

À vista da dificuldade para combater a infestação por meios convencionais, novas estratégias se destacam, como a contenção (ou erradicação) da infestação de uma espécie invasora por meio de uma competição planejada (NDII *et al.*, 2012; SCHRAIBER *et al.*, 2012; ZHANG *et al.*, 2015; HASTINGS, 1994). Este trabalho estuda a competição planejada entre o mexilhão dourado selvagem e outro, que será modificado geneticamente. Ou seja, a estratégia admite uma competição na qual o encontro dos gametas mexilhões selvagens com os modificados promovam larvas modificadas. Tais larvas podem produzir mexilhões modificados, que transmitem a infertilidade, culminando na redução populacional. A estratégia se encontra em desenvolvimento por meio do projeto de P&D ANEEL: Controle da Infestação do Mexilhão Dourado por Indução Genética de Infertilidade (PD-10381-0419/2019), e esse trabalho tem como objetivo estudar a viabilidade desta técnica, embora admita considerações próprias para que a competição possa ser analisada, uma vez que maiores informações sobre a espécie modificada estão indisponíveis.

Modelos de Crescimento I e II

O modelo I consiste em um sistema com três equações diferenciais ordinárias, elaborado a partir dos fatores que influenciam na variação populacional dos envolvidos: larvas, mexilhões adultos e algas, respectivamente L , M e A . As algas são entendidas como principal fonte de alimento do mexilhão, como aponta PADILLA (1997). Entende-se que a variação populacional das larvas depende da reprodutividade, r_1 , do mexilhão e seu nível populacional. E seu crescimento populacional é arrefecido por uma capacidade suporte, K_L . Além disso, também se admite perdas relativas às taxas de mortalidade e maturação, μ e λ . Para os mexilhões, admite-se variação positiva de acordo com a taxa de maturação e o nível populacional das larvas, bem como também variam de acordo com o encontro mexilhão-alga, modelado pela resposta funcional de Michaelis-Menten (EDELSTEIN-KESHET, 2005; MURRAY, 2002), sob uma taxa máxima de predação β e uma constante de saturação média c . Assim como as larvas, a variação populacional do mexilhão sofre perdas pela taxa de mortalidade, b .

Finalmente, entende-se que as algas sofrem variação positiva de acordo com sua taxa de crescimento, r_2 , e negativas de acordo com a taxa de predação, p . As equações (1)-(3) apresentam o modelo I, cujo desenvolvimento completo pode ser acompanhado em BARBOSA (2022), tal que t é tempo.

$$\frac{dL}{dt} = r_1 M \left(1 - \frac{L}{K_L} \right) - (\mu + \lambda) L, \quad (1)$$

$$\frac{dM}{dt} = \lambda L + \beta \left(\frac{MA}{c + A} \right) - bM, \quad (2)$$

$$\frac{dA}{dt} = r_2 A \left(1 - \frac{A}{K_A} \right) - p \left(\frac{AM}{c + A} \right). \quad (3)$$

Para o Modelo I, foram encontradas analiticamente três soluções estacionárias:

$$P_1 = (0, 0, 0), \quad (4)$$

$$P_2 = \left(\frac{K_L(r_1\lambda - bs)}{r_1\lambda}, \frac{K_L(r_1\lambda - bs)}{r_1b}, 0 \right), \quad (5)$$

$$P_3 = (0, 0, K_A). \quad (6)$$

A primeira solução estacionária, P_1 , não desperta muito interesse, uma vez representa a extinção de todas as populações envolvidas. P_2 , por outro lado, representa a extinção apenas das algas. Observe que esta é uma direta consequência da modelagem do problema, uma vez que o modelo entende que a interação mexilhão-alga é benéfica ao mexilhão, mas não necessária para sua subsistência. P_3 representa uma solução interessante, que mostra a extinção do invasor e o prevalecimento das algas. Uma vez que esse cenário pode exemplificar o meio antes da infestação, a estabilidade dessa solução estacionária pode contribuir para o entendimento da infestação.

Uma análise de estabilidade é realizada localmente, na região próxima às soluções estacionárias, de modo a, através da aplicação da Matriz Jacobiana, obter as partes reais dos autovalores associados às soluções (MURRAY, 2002; EDELSTEIN-KESHET, 2005). A desigualdade (7) apresenta a condição encontrada para que P_3 seja localmente estável e consiga resistir à infestação, devidamente atrelada às partes reais de seus autovalores associados. Ou seja, trata-se do número de reprodução basal do Modelo I, representado por R_0^1 , que implica no Teorema 1.

Teorema 1: Se $R_0^1 < 1$ então o ponto de equilíbrio $P_3 = (0, 0, K_A)$ possui três autovalores reais negativos, sendo portanto estável. Quando $R_0^1 > 1$ temos dois autovalores negativos e um autovalor positivo, sendo consequentemente instável.

$$R_0^1 = \frac{\beta s K_A + r_1 \lambda (c + K_A)}{(c + K_A) b s} < 1. \quad (7)$$

Ademais, as taxas de crescimento das larvas (reprodutividade dos mexilhões), de mortalidade e maturação foram estimadas através das referências (CATALDO *et al.*, 2005; PAOLUCCI *et al.*, 2010; VANDERPLOEG *et al.*, 1996). A Tabela 1 apresenta os valores para as simulações, ressalta-se que os demais valores foram adotados.

Tabela 1: Valores dos parâmetros da simulação do Modelo I.

Parâmetro	r_1	μ	λ	β	b	r_2	p	K_L	K_A	c
Valor	0,07	0,004	0,03	0,01	0,01	0,12	0,02	20	0,01	0,001
Unidade	dia^{-1}	dia^{-1}	dia^{-1}	dia^{-1}	dia^{-1}	dia^{-1}	dia^{-1}	g/l	g/l	g/l

As equações (8)-(10) formam o sistema relacionado ao Modelo II (SILVA *et al.*, 2022). Este, difere-se do Modelo I em dois importantes pontos. Primeiro, ele considera que não apenas a interação mexilhão-alga é benéfica ao mexilhão, mas necessária para a subsistência do molusco. Além disso, essa interação passa a ser modelada pela resposta funcional de Holling tipo 3 (HOLLING, 1965).

$$\frac{dL}{dt} = r_1 M \left(1 - \frac{L}{K_L} \right) - \mu L - \lambda L, \quad (8)$$

$$\frac{dM}{dt} = \lambda \frac{A^2}{c + A^2} L \left(1 - \frac{M}{K_M} \right) - b M, \quad (9)$$

$$\frac{dA}{dt} = r_2 A \left(1 - \frac{A}{K_A} \right) - p \frac{A^2}{c^2 + A^2} M. \quad (10)$$

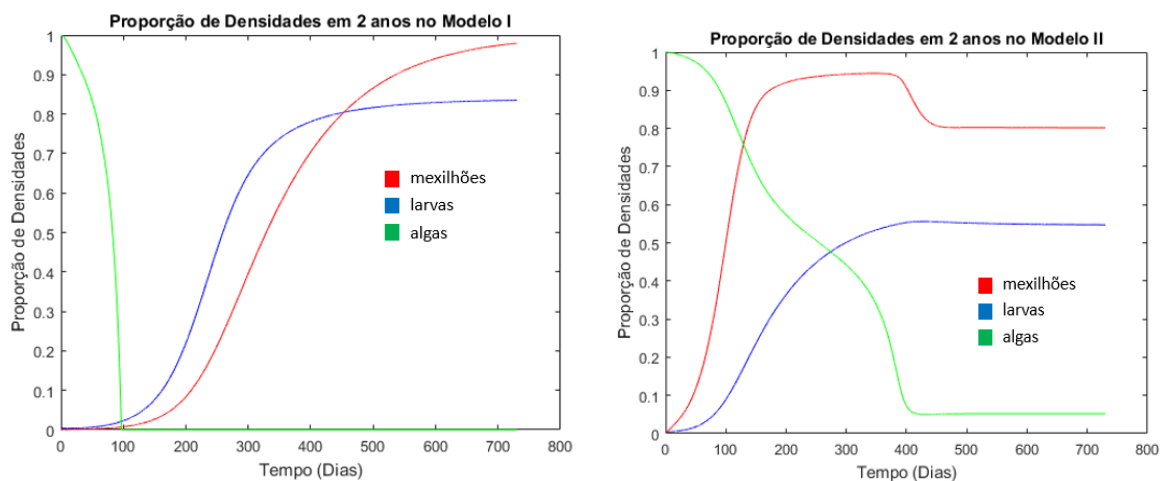
Para o Modelo II, as soluções estacionárias são: $S_1 = (0, 0, 0)$ e $S_2 = (0, 0, K_A)$. Note que, ao tornar necessária a existência de alga para o desenvolvimento do mexilhão, a possibilidade de uma solução estacionária que culmine na extinção das algas é eliminada. Realizando o mesmo procedimento de avaliação da estabilidade local das soluções de equilíbrio para o Modelo II, S_1 e S_2 , encontra-se a condição expressa pela desigualdade (11) para que S_2 seja localmente estável. Ou seja, trata-se do número de reprodução basal para o Modelo II, R_0^2 , que acarreta no Teorema 2.

Teorema 2: Se $R_0^2 < 1$ então o ponto de equilíbrio $S_2 = (0, 0, K_A)$ possui três autovalores reais negativos, sendo portanto estável. Quando $R_0^2 > 1$ temos dois autovalores negativos e um autovalor positivo, sendo consequentemente instável.

$$R_0^2 = \frac{r_1 \lambda K_A^2}{(K_A^2 + c_1^2)b(\lambda + \mu)} < 1. \quad (11)$$

Diferentemente do Modelo I, o Modelo II foi desenvolvido pelo grupo de pesquisa (SILVA *et al.*, 2022), e a adoção dos parâmetros estimada por observações de campo. A capacidade suporte para o Modelo II é de 1732 gm^{-2} , que evidencia uma unidade diferente e, portanto, exigiu um ajuste para as simulações numéricas. Além disso, a taxa p para o segundo modelo foi adotada em $0,2 \text{ dia}^{-1}$, por semelhante ajuste dimensional. Os demais parâmetros permanecem os mesmos da Tabela 1.

A Figura 1 apresenta a dinâmica populacional, evidenciando a proporção da capacidade suporte no tempo, visando uma análise qualitativa do comportamento. Como condições iniciais, considera-se a população de algas em sua capacidade suporte K_A , sem a presença de mexilhões adultos e a entrada de $0,5\%K_L$. Como esperado, o Modelo I não representa a realidade, uma vez que, para os atuais parâmetros, as algas são extintas e a solução caminha para P_2 . Por outro lado, o Modelo II apresenta a coexistência entre as três populações, o que é observado na prática. Além disso, a queda da densidade de algas culmina em uma queda para as demais populações. Essa consequência tem importância quando o problema é analisado em uma região, de modo que a presença do invasor pode ser confirmada de acordo com a maior ausência de algas (SILVA *et al.*, 2022).



(a) Comportamento da proporção de densidades populacionais em 2 anos de simulação para o Modelo I. (b) Comportamento da proporção de densidades populacionais em 2 anos de simulação para o Modelo II.

Figura 1: Resultados para proporção de densidade de todas as espécies em 2 anos de simulação.

Modelo de Competição

Para a elaboração do Modelo de Competição, escolheu-se o Modelo II, para servir de base por adequar melhor a realidade. Sobre a estratégia de competição que deu origem a modelagem do problema, considera-se que a infertilidade será promovida da seguinte forma: as larvas selvagens são modeladas em função do encontro entre os gametas do mexilhão macho selvagem e da fêmea selvagem. As larvas modificadas foram modeladas em função do encontro dos gametas do mexilhão macho selvagem com os da fêmea modificada. Nesse contexto, os gametas do mexilhão macho modificado não produzem prole.

As equações (12)-(16) apresentam o modelo proposto,

$$\frac{dL_s}{dt} = r_1 \frac{M_s^2}{M_s + M_m} \left(1 - \frac{L_s + L_m}{K_L} \right) - \mu L_s - \lambda L_s, \quad (12)$$

$$\frac{dM_s}{dt} = (L_s + (1 - \chi)L_m) \lambda \frac{A^2}{c^2 + A^2} \left(1 - \frac{M_s + M_m}{K_M} \right) - bM_s, \quad (13)$$

$$\frac{dL_m}{dt} = \bar{r}_1 \frac{M_s M_m}{M_s + M_m} \left(1 - \frac{L_s + L_m}{K_L} \right) - \bar{\mu} L_m - \bar{\lambda} L_m, \quad (14)$$

$$\frac{dM_m}{dt} = \chi L_m \lambda \frac{A^2}{c^2 + A^2} \left(1 - \frac{P}{K_M} \right) - \bar{b} M_m, \quad (15)$$

$$\frac{dA}{dt} = r_2 A \left(1 - \frac{A}{K_A} \right) - p \frac{A^2}{c^2 + A^2} (M_s + M_m). \quad (16)$$

Assim, os mexilhões modificados são representados por M_m , os selvagens por M_s . A representação é análoga para as larvas. Ainda, o parâmetro χ deve ser entendido como a proporção (efetividade) de transmissão dessa modificação genética para futuras gerações. Embora a engenharia genética para controle populacional de algumas espécies já vem sendo discutida há certo tempo no meio acadêmico (BURT, 2003; HAMMOND *et al.*, 2016; HASTINGS, 1994), a obtenção de χ está relacionada a diversos fatores de experimentação.

Diversos experimentos numéricos foram realizados. Experimentos que avaliavam um cenário sem vantagens competitivas, isto é, o qual todas as características dos selvagens (valores de parâmetros) são mantidas para os modificados, e casos, os quais, foram atribuídas à espécie modificada vantagens competitivas. Avaliou-se ao longo da dissertação uma vantagem em cada um dos parâmetros envolvidos, de modo a confrontar o resultado com um análise de sensibilidade dos parâmetros que fazem parte do número de reprodução basal do Modelo II, R_0^2 . Ademais, também foram estudadas as condições iniciais das populações envolvidas. Para os modificados, os valores ficaram fixos em 2 gl^{-1} para as larvas e $0,175 \text{ kgm}^{-2}$ para os mexilhões. Quanto aos selvagens, os valores foram alterados ao longo das simulações, visando estudar como sua redução (limpeza do meio) poderia acelerar a dinâmica de infestação do modificado, a Tabela 2 indica essas reduções.

Tabela 2: Condições iniciais para diferentes níveis de limpeza.

População	Limpeza de 70%	Limpeza de 90%
Larvas Selvagens	3 gl^{-1}	1 gl^{-1}
Mexilhões Selvagens	$0,405 \text{ kgm}^{-2}$	$0,135 \text{ kgm}^{-2}$
Algas	$0,5 \text{ mgl}^{-1}$	$0,5 \text{ mgl}^{-1}$

A Figura 2a apresenta as proporções populacionais em 10 anos de simulação. Considera-se a limpeza de 70% do meio, atribui-se à espécie modificada uma taxa de reprodutividade 30% mais eficiente e uma mortalidade 15% menor. Enquanto que a Figura 2b considera uma limpeza no meio de 90%, com os modificados sendo 40% mais eficientes na taxa de reprodutividade e uma mortalidade 30% menor. Ambos os casos evidenciam, nessas condições ideais, a erradicação do molusco, que ocorre em um longo prazo.

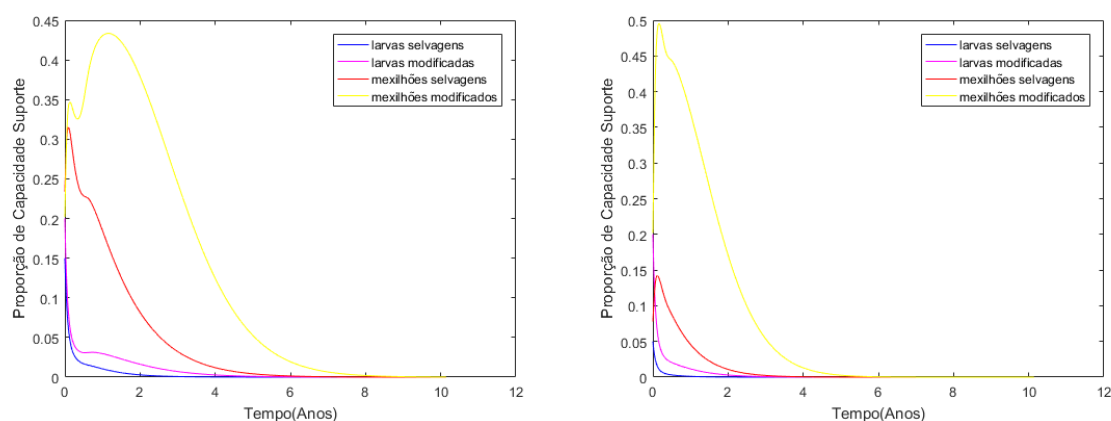


Figura 2: Proporção das densidades populacionais em 10 anos de simulação. À esquerda com limpeza de 70% e à direita com limpeza de 90%.

Considerações Finais

Esse trabalho analisou como inferências de modelagem podem atuar sobre a dinâmica de infestação nos modelos de crescimento e estudar a viabilidade da estratégia de contenção proposta pela inserção de uma espécie geneticamente modificada em um meio. Nota-se que o primeiro resultado se relaciona bem com o segundo, uma vez que é necessário adotar um modelo de crescimento para propor o modelo de competição.

Dadas as simulações numéricas apresentadas para os diferentes cenários para avaliar a estratégia de competição planejada, tornou-se evidente a necessidade de vantagens competitivas para que esta seja eficiente. Além disso, o trabalho também evidenciou que, ainda que a erradicação seja obtida sob condições ideais, a estratégia demanda um longo prazo de tempo. No entanto, baixos níveis populacionais do invasor podem ser obtidos em menor tempo e com vantagens competitivas mais razoáveis (maior eficiência reprodutiva e um mexilhão modificado menos suscetível à predação). Por fim, ressalta-se que não foram avaliados casos de reinserção de modificados no meio, o que poderia influenciar diretamente na viabilidade da estratégia.

Referências

- BARBOSA, C. H. X. B. **Um Estudo Hipotético Baseado em Modelos de Crescimento e Competição para a Dinâmica Populacional do Mexilhão Dourado**. Dissertação de mestrado. UFRRJ, 2022.
- BOLTOVSKOY, D. *et al.*. Dispersion and ecological impact of the invasive freshwater bivalve *Limnoperna fortunei* in the Río de la Plata watershed and beyond. **Biological Invasions**, 8, pp. 947–963, 2006.
- BOLTOVSKOY, D. *et al.*. Nutrient recycling, phytoplankton grazing, and associated impacts of *Limnoperna fortunei*. **Limnoperna fortunei: the ecology, distribution and control of a swiftly spreading invasive fouling mussel**, pp. 153–176, 2015.
- BURT, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. **Proceedings of the Royal Society of London. Series B: Biological Sciences**, 270.1518, pp. 921–928, 2003.
- CATALDO, D. *et al.*. Temperature-dependent rates of larval development in *Limnoperna fortunei* (Bivalvia: Mytilidae). **Journal of Molluscan Studies**, 71(1), pp. 41–46, 2005.

- EDELSTEIN-KESHET, L. **Mathematical models in biology**. SIAM, 2005.
- HAMMOND, A. *et al.*. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. **Nature biotechnology**, 34(1), pp. 78–83, 2016.
- HASTINGS, I. M. Selfish DNA as a method of pest control. **Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences**, 344.1309, pp. 313–324, 1994.
- HOLLING, C. S. The functional response of predators to prey density and its role in mimicry and population regulation. **The Memoirs of the Entomological Society of Canada**, 97S45, pp. 5–60, 1965.
- MURRAY, J. D. **Mathematical biology: I. An introduction**. Springer, 2002.
- NDII, M. Z.; HICKSON, R. I.; MERCER, G. N. Modelling the introduction of Wolbachia into *Aedes aegypti* mosquitoes to reduce dengue transmission. **The ANZIAM Journal**, 53(3), pp. 213–227, 2012.
- PADILLA, D. K. The effects of *Dreissena polymorpha* (Pallas) invasion on aquatic communities in eastern Europe. **Journal of Shellfish Research**, 16, pp. 187–203, 1997.
- PAOLUCCI, E. M. *et al.*. Veligers of an introduced bivalve, *Limnoperna fortunei*, are a new food resource that enhances growth of larval fish in the Paraná River (South America). **Freshwater Biology**, 55(9), pp. 1831–1844, 2010.
- RICCIARDI, A. Global range expansion of the Asian mussel *Limnoperna fortunei* (Mytilidae): another fouling threat to freshwater systems. **Biofouling**, 13(2), pp. 97–106, 1998.
- SCHRAIBER, J. G. *et al.*. Constraints on the use of lifespan-shortening Wolbachia to control dengue fever. **Journal of theoretical biology**, 297, pp. 26–32, 2012.
- SILVA, F. A. *et al.*. Mexilhão-Dourado no Brasil: detecção de um perigoso invasor. **Ciência Hoje**, 338(57), pp. 38–42, 2016.
- SILVA, J. C. R. *et al.*. Population growth of the golden mussel (*L. fortunei*) in hydroelectric power plants: a study via mathematical and computational modeling. **Revista Brasileira de Recursos Hídricos**, 27(9), pp. 1-15, 2022.
- SIMEÃO, C. M. G.; RESENDE, M. F.; MARTINEZ, C. B. Variação das características hidráulicas em condutos forçados operando sob condições de infestação por *Limnoperna fortunei*. **Revista Brasileira de Recursos Hídricos**, 16(2), pp. 13–24, 2011.
- SOUZA, R. C. C. L.; CALAZANS, S. H. Calazans; SILVA, E. P. Impacto das espécies invasoras no ambiente aquático. **Ciência e cultura**, 61(1), pp. 35–41, 2009.
- VANDERPLOEG, H. A.; LIEBIG, J. R.; GLUCK, A. A. Evaluation of different phytoplankton for supporting development of zebra mussel larvae (*Dreissena polymorpha*): the importance of size and polyunsaturated fatty acid content. **Journal of Great Lakes Research**, 22(1), pp. 36–45, 1996.
- ZHANG, X.; TANG, S.; CHEKE, R. A. Models to assess how best to replace dengue virus vectors with Wolbachia-infected mosquito populations. **Mathematical Biosciences**, 269, pp. 164–177, 2015.

REFORÇO DAS OPERAÇÕES BÁSICAS ATRAVÉS DE UMA ATIVIDADE DE MODELAGEM MATEMÁTICA SOBRE O PLANEJAMENTO DE VIAGENS

Victor Hugo de Paula Oliveira

Universidade Federal Rural do Rio de Janeiro

victor.h_paula@hotmail.com

Claudia Mazza Dias

Universidade Federal Rural do Rio de Janeiro

mazza@ufrj.br

Marcela Lima Santos Pereira

Universidade Federal Rural do Rio de Janeiro

marcelalsantos@ufrj.br

Resumo: A Modelagem Matemática pode ser utilizada como alternativa ao ensino tradicional de Matemática. Assim, este trabalho apresenta uma proposta de atividade por antecipação, tendo como tema o planejamento de uma viagem. A atividade foi aplicada junto aos alunos de uma turma do 5º ano de uma Escola Municipal, que puderam entender que para a determinação dos custos de uma viagem é indispensável o uso das operações básicas, que são assim revisitadas para reforço de seu entendimento e significação de seu uso. A atividade também proporcionou aos estudantes a oportunidade de ampliar seu conhecimento sobre o Turismo e sobre o planejamento financeiro.

Palavras-chave: Modelagem Matemática. Educação Financeira. Planejamento de Viagens. Operações Básicas.

Introdução

Neste trabalho, utilizou-se a Modelagem Matemática como ferramenta para o ensino, através da investigação de uma situação cotidiana, onde busca-se contextualizar o conteúdo matemático com a realidade dos alunos (D'AMBROSIO, 1989; BIEMBENGUT, 2014; BASSANEZI, 2002). O planejamento para se fazer uma viagem é tema de uma atividade para reforço dos conceitos matemáticos relativos às quatro operações fundamentais (básicas), e aplicado junto às turmas do 5º ano do Ensino Fundamental. Ensinar Matemática é um desafio frente às dificuldades apresentadas por muitos estudantes. Nem sempre se reconhece a utilidade dos conteúdos na solução de problemas do cotidiano, como é o caso do planejamento financeiro, através do levantamento dos custos necessários para realização de uma viagem. Por esse motivo, pensou-se em apresentar em sala de aula a ideia de planejar uma viagem, que tem potencial não só de aguçar a curiosidade dos alunos como também seria uma oportunidade de realizar uma aula diferente, na qual se pode explorar diversos conteúdos interdisciplinares, evidenciando que fazer uma viagem é mais do que pensar apenas no lazer (CECIERJ, 2010). É preciso estar atento às recomendações que nos garantem uma certa tranquilidade e conforto para poder aproveitar todos os momentos, garantir antecipadamente os recursos monetários necessários para utilização durante o período da viagem. Logo, é importante desenvolver estratégias de economia doméstica e planejamento.

Além das quatro operações básicas em consonância com a Base Nacional Comum Curricular (BRASIL, 2017), a atividade permitiu também mostrar a importância e a utilidade das tabelas. A interdisciplinaridade foi evidenciada pela abordagem de conteúdos relativos à geografia e à economia doméstica, bem como conhecimentos gerais.

Metodologia

A atividade, planejada por antecipação, levou em conta a produção de material de apoio para os estudantes, como formulários com as informações necessárias, e até mesmo uma ficha de avaliação a ser preenchida por eles. Os alunos precisaram levantar, com ajuda desse material, quais os recursos financeiros necessários para a realização de uma viagem. Foram considerados 4 destinos turísticos, todos na região Sudeste. Os alunos tiveram que pesquisar sobre os destinos turísticos, consultar tabelas com valores aproximados de hospedagem e passagens, consultar a Internet para levantar quais os principais pontos turísticos de cada localidade, custos de ingressos, alimentação, etc. Com as informações em mãos, foram realizados os cálculos das despesas necessárias considerando as categorias: passagens, hospedagem (2 diárias), alimentação, ingressos e lembrancinhas de viagem. Assim, eles compararam os recursos necessários para visitar cada destino, observando qual o destino mais caro e qual o mais barato. Para nortear a pesquisa, uma história fictícia foi apresentada aos alunos. A atividade foi ministrada por um então licenciando do curso de Matemática da UFRRJ, estagiário da escola, que pediu “ajuda” da turma na definição do destino e estimativa dos gastos para uma viagem para comemoração de sua formatura. Para tal, os alunos tiveram que levar em consideração que como recursos mensais teria disponível apenas 1 salário mínimo que deveria também suprir suas necessidades básicas. Os alunos calcularam quanto poderia ser reservado por mês para a viagem e em quanto tempo seria possível viajar com tais recursos.

Conclusões

A partir da participação dos alunos durante a atividade e das respostas dos alunos através da ficha de avaliação da atividade, percebe-se que o objetivo da atividade foi alcançado, uma vez que os alunos compreenderam a utilidade da matemática, não só no planejamento de viagem, como também em outras situações cotidianas. Os alunos observaram naturalmente a necessidade do uso das operações básicas, ainda que tenham apresentado dificuldades no cálculo das despesas. Também entenderam a importância de economizar para adquirir bens e serviços. Adicionalmente, a atividade abordou a importância do salário mínimo, assim como os conhecimentos gerais e geografia, o que foi crucial para a abordagem interdisciplinar.

Referências

- BASSANEZI, R. C. *Ensino e aprendizagem com modelagem matemática*. São Paulo: Contexto, 2002.
- BIEMBENGUT, M. S. *Modelagem Matemática no Ensino Fundamental*. 1ª Ed, Blumenau: Edifurb, 2014.
- BRASIL. Base Nacional Comum Curricular, Ministério da Educação-Secretaria de Educação Básica. Brasília, DF. 2017. Disponível em: <http://basenacionalcomum.mec.gov.br/>. Acesso em: 22 jun. 2023.
- CECIERJ. Centro de Ciências e Educação Superior a Distância do Estado do Rio de Janeiro. Fundamentos do Turismo. Rio de Janeiro. v. 1. 2010. 293p. Disponível em: <https://canal.cecierj.edu.br/recurso/6457>. Acesso em: 11 jul. 2023.
- D'AMBROSIO, B. S. Como ensinar matemática hoje?. *Temas e Debates*. SBEM. n. 2, p. 15-19, 1989.

UMA PROPOSTA DE ATIVIDADE DE PROGRAMAÇÃO LINEAR PARA O 9º ANO DO ENSINO FUNDAMENTAL

Raphael da Silva Bacelar

*Universidade Federal Rural do Rio de Janeiro
raphaelbacelar@yahoo.com.br*

Marcela Lima Santos Pereira

*Universidade Federal Rural do Rio de Janeiro
marcelalsantos@ufrj.br*

Ronaldo Malheiros Gregorio

*Universidade Federal Rural do Rio de Janeiro
rgregor@ufrj.br*

Resumo: Ao estudar um conteúdo novo de matemática, é muito comum o aluno perguntar: “Por que eu preciso aprender isso?” ou, ainda, “Em que situação, eu irei usar?”. O presente trabalho visa apresentar uma proposta de atividade elaborada objetivando tornar o Ensino de Matemática mais significativo, fazendo uso da Resolução de Problemas como recurso didático trabalhando o conceito de inequações, sistemas de inequações e representação gráfica no plano cartesiano, para solucionar um Problema de Programação Linear. A Programação Linear é uma das mais importantes técnicas de Pesquisa Operacional que se originou no final da década de 40 (DANTZIG, 2002), consiste em minimizar (ou maximizar) uma função objetivo, comumente o custo, sujeita a algumas restrições que reduzem o espaço de possibilidade para as variáveis em questão, de modo que as relações estabelecidas sejam lineares. A Resolução de Problemas é um recurso que visa romper o paradigma de que para saber matemática basta memorizar algumas fórmulas e aplicá-las corretamente, permitindo que os alunos desenvolvam uma compreensão mais significativa dos conceitos matemáticos e a forma adequada de utilizá-los. Segundo Polya (1995), a Resolução de Problemas requer a execução de uma sequência de fases, a saber: i) compreender o problema; ii) estabelecer um plano para busca de solução; iii) executar o plano; iv) retrospecto, que consiste em examinar a solução obtida. Este recurso didático quando utilizado para trabalhar problemas cotidianos permite que os alunos identifiquem a relevância da Matemática para além da sala de aula, conforme defendido por D’Ambrósio (2012). Além disso, alunos podem desenvolver habilidades críticas, como raciocínio lógico, tomada de decisão e pensamento criativo. Quando trabalhado coletivamente, a Resolução de Problemas envolve discussões e colaboração entre os alunos, permitindo que estes aprendam a comunicar suas estratégias e soluções de maneira clara, tornando-os agentes ativos na construção do próprio saber e colaborando na formação de seus colegas. Dessa forma, o presente trabalho visa apresentar uma proposta de atividade idealizada para estudantes do 9º ano do Ensino Fundamental, baseada na Resolução de Problemas buscando definir, hipoteticamente, o quantitativo de dois bolos a serem produzidos e vendidos na cantina de uma escola com o material já disponível e visando maximizar o ganho com a venda destes. Para desenvolvimento da atividade será apresentado o plano cartesiano e toda sua estrutura, bem como a definição do espaço solução para inequações lineares, trabalhando a definição do espaço interseção para o caso de estudo de sistemas de inequações lineares, conceito fundamental para a Programação Linear. A revisão do conceito de equações do primeiro grau pode ser requerida. Embora o estudo de sistemas lineares seja previsto apenas para o Ensino Médio, a atividade permite a antecipação deste conceito sem a apresentação formalizada do mesmo. Acreditamos que o tratamento do problema sob a representação gráfica facilitará o entendimento do problema permitindo que os alunos visualizem o papel das restrições na definição do espaço solução.

Palavras-chave: Resolução de Problemas. Programação Linear. Problema da Dieta.

Agradecimentos

A pesquisa relatada neste trabalho foi apoiada pela FAPERJ por meio Edital nº 45/2021 - Apoio à melhoria das escolas da rede pública sediadas no Estado do Rio de Janeiro - 2021.

Referências

D'AMBROSIO, U. **Educação Matemática: da teoria à prática**. 23. ed., São Paulo: Papirus Editora, 2012.

DANTZIG, G. B. Linear Programming. **Operations Research**. v. 50, n. 1, p. 42-47, 2002. Disponível em: <<https://pubsonline.informs.org/doi/10.1287/opre.50.1.42.17798>>. Acesso em: 06 set. 2023.

POLYA, G. **A arte de resolver problemas: um novo aspecto do método matemático**. Tradução e adaptação ARAÚJO, H. L. Rio de Janeiro: Interciência, 1995.

A PROGRAMAÇÃO LINEAR COMO FERRAMENTA PARA ELABORAÇÃO DE UM CARDÁPIO DA MERENDA ESCOLAR PARA O ALMOÇO

Caio Vasconcellos de Paulo

*Universidade Federal Rural do Rio de Janeiro
caiovpaulo@gmail.com*

Marcela Lima Santos Pereira

*Universidade Federal Rural do Rio de Janeiro
marcelalsantos@ufrj.br*

Ronaldo Malheiros Gregorio

*Universidade Federal Rural do Rio de Janeiro
rgregor@ufrj.br*

Resumo: A preocupação com uma dieta balanceada e de baixo custo é uma questão importante em diversos contextos, incluindo a alimentação escolar. A oferta de uma alimentação saudável e equilibrada é fundamental para a saúde e o desenvolvimento dos estudantes, mas também pode apresentar desafios em relação à disponibilidade de recursos e à eficiência dos processos envolvidos (BRASIL, 2019). Quando pensada sob o ponto de vista matemático, a alimentação saudável em escolas públicas deve atender às exigências nutricionais previstas pelos órgãos de saúde e de educação, visando sem agregar grandes custos. Portanto, este problema (problema da dieta) pode ser facilmente tratado como um problema de Programação Linear (OLIVEIRA *et al.*, 2011). A Programação Linear é uma técnica que se destina a resolver problemas de maximização ou minimização que possam ser representados através de expressões lineares (MARINS, 2011). A Programação Linear permite encontrar a combinação ideal de alimentos para alcançar metas nutricionais específicas, tornando-se uma ferramenta valiosa para melhorar a qualidade de vida e a eficiência na gestão de recursos alimentares. Dessa forma, buscando realizar um estudo sob a realidade vivenciada no Colégio Estadual Engenheiro Arêa Leão, localizado no município de Nova Iguaçu-RJ, está sendo desenvolvido uma pesquisa visando determinar a melhor dieta para os estudantes a partir dos princípios da Programação Linear e utilizando a Modelagem Matemática como metodologia científica. O presente trabalho destina-se a apresentação dos primeiros resultados para o estudo de elaboração de um cardápio para o almoço. Para a construção dos modelos, foi utilizado como referência para os alimentos e seus valores nutricionais a Tabela Brasileira de Composição de Alimentos - TACO (TACO, 2011), que consistem em um banco de dados de informações relativas a composição nutricional de diversos alimentos consumidos no país. A tabela TACO contém informações sobre os macronutrientes (proteínas, carboidratos e gorduras), micronutrientes (vitaminas e minerais) e fibras presentes nos alimentos, além de informar as calorias e massa de porções comuns. A alimentação escolar é um importante componente para a saúde e o bem-estar dos estudantes, e o café da manhã é uma das principais refeições do dia (MARINS, 2011). Nesse sentido, é importante que o cardápio da merenda escolar das escolas seja planejado de forma a garantir a oferta de alimentos nutritivos e equilibrados, respeitando as necessidades nutricionais dos estudantes. Inicialmente, o objetivo era criar um cardápio semanal para o café da manhã, considerando restrições nutricionais para carboidratos, proteínas, fibras alimentares, cálcio, ferro e vitamina C, fundamentais na dieta diária. O objetivo do modelo foi minimizar o custo dos alimentos, levando em consideração as restrições nutricionais estabelecidas. Devido ao grande número de variáveis envolvidas no problema de otimização, se fazia inviável a resolução manual do problema e, para tanto, foi utilizado um software de planilha eletrônica - Excel®, e a ferramenta Solver. Para o cardápio do almoço foram adicionadas uma série de novas restrições estabelecidas em Brasil (2022), como o consumo máximo de 600mg de sódio por refeição ou 800 mg em duas refeições e o consumo mínimo obrigatório de frutas duas vezes por semana e de hortaliças três vezes totalizando 280g semanais. Para o primeiro cardápio do almoço, foram consideradas 121 variáveis de decisão associadas aos alimentos da Tabela TACO, divididos em 13 grupos alimentares. Essas variáveis foram submetidas a 12 restrições no processo

de otimização através da ferramenta Solver. A simulação permitiu concluir que, se todos os alimentos tivessem o mesmo custo, apenas aqueles que apresentam uma contribuição significativamente maior seriam utilizados. Isso ressalta a importância de considerar não apenas o custo, mas também a contribuição de cada alimento ao problema em questão. Além disso, essa análise proporcionou uma compreensão mais profunda do problema, destacando a necessidade de realizar novos estudos para estabelecer restrições adicionais. Isso sugere que uma abordagem mais refinada e completa é necessária para abordar o problema de forma eficaz e eficiente, levando em consideração múltiplos fatores e variáveis.

Palavras-chave: Modelagem Matemática. Programação Linear. Problema da Dieta. Merenda Escolar.

Agradecimentos

A pesquisa relatada neste trabalho foi apoiada pela FAPERJ por meio Edital nº 45/2021 - Apoio à melhoria das escolas da rede pública sediadas no Estado do Rio de Janeiro - 2021.

Referências

BRASIL. 2019. **Programa Nacional de Alimentação Escolar (PNAE)**. Brasília: Ministério da Educação-MEC/Secretaria de Educação Básica-SEB. Disponível em: <<https://www.gov.br/fnde/pt-br/aceso-a-informacao/acoes-e-programas/programas/pnae>>. Acesso em: 06 set. 2023.

BRASIL. **Planejamento de cardápios para a Alimentação Escolar**. Brasília: Ministério da Educação-MEC/Fundo Nacional de Desenvolvimento da Educação-FNDE, 2022. Disponível em: <https://www.gov.br/fnde/pt-br/aceso-a-informacao/acoes-e-programas/programas/pnae/manuais-e-cartilhas/MANUAL_V8.pdf>. Acesso em: 06 set. 2023.

MARINS, Fernando Augusto Silva. **Introdução à Pesquisa Operacional**. São Paulo, Cultura Acadêmica, 2011.

OLIVEIRA, Danilo E; BORGES, Ana C. A; SILVA, Vinycius V. **Uma aplicação do problema da dieta para se encontrar o menor custo de refeições diárias para idosos na cidade de Monte Carmelo – MG**. Brazilian Journal of Development, Curitiba, 2020.

TACO. **Tabela Brasileira de Composição de Alimentos**. Versão 4. Campinas: Unicamp, 2011. <https://www.cfn.org.br/wp-content/uploads/2017/03/taco_4_edicao_ampliada_e_revisada.pdf>. Acesso em: 06 set. 2023.

A PROGRAMAÇÃO LINEAR COMO FERRAMENTA PARA ELABORAÇÃO DE UM CARDÁPIO DA MERENDA ESCOLAR PARA O CAFÉ DA MANHÃ

Stephanie Pimentel da Silva

Universidade Federal Rural do Rio de Janeiro

stephaniepsilva16@gmail.com

Marcela Lima Santos Pereira

Universidade Federal Rural do Rio de Janeiro

marcelalsantos@ufrj.br

Ronaldo Malheiros Gregorio

Universidade Federal Rural do Rio de Janeiro

rgregor@ufrj.br

Resumo: É de conhecimento geral que uma boa alimentação é indispensável para a sobrevivência do ser humano. Manter uma dieta balanceada, com os nutrientes essenciais, previne regularmente doenças relacionadas à má-alimentação, como anemia e obesidade (BRASIL, 2013). Logo, a essencialidade da escolha consciente de alimentos que irão figurar em um cardápio é indiscutível. O Problema da Dieta teve sua origem na década de 30 com George Stigler (DANTZIG, 2002). Stigler utilizou Programação Linear para determinar um conjunto de alimentos que fornecesse a quantidade essencial de nutrientes ao menor custo possível. A necessidade de encontrar uma dieta ideal para as tropas durante a Segunda Guerra Mundial impulsionou o desenvolvimento de técnicas de otimização para a formulação de dietas equilibradas e de baixo custo. Os pesquisadores buscavam encontrar a combinação perfeita de alimentos que garantisse a ingestão adequada de nutrientes ao mesmo tempo em que minimizavam os gastos. Atualmente, o Problema da Dieta continua sendo um tema de estudo importante, podendo ser uma ferramenta crucial na elaboração de cardápios para a merenda escolar. Através de modelos matemáticos e técnicas de otimização, é possível criar um cardápio balanceado e saudável para atender às necessidades nutricionais dos alunos. Nessa perspectiva, o presente trabalho visa apresentar o estudo para desenvolvimento de um cardápio de café da manhã para uma unidade escolar da região de Nova Iguaçu. Para tal, estão sendo realizadas simulações que envolvem a criação de diferentes cardápios, levando em consideração os diferentes tipos de alimentos oferecidos e a quantidade nutricional de cada um. A otimização se dará por meio da minimização do custo do cardápio oferecido e as restrições a serem atendidas, de maneira generalizada, serão as quantidades mínimas de nutrientes a serem ofertados. Por meio deste modelo, em síntese, será possível suprir a quantidade de nutrientes previamente escolhidos na alimentação dos indivíduos afetados, garantindo que recebam uma dieta balanceada e saudável com o menor custo possível. Segundo Brasil (2022), um cardápio adequado e saudável deve considerar a quantidade adequada de calorias e nutrientes, priorizar alimentos de qualidade, equilibrar os nutrientes nas refeições e adaptar-se às necessidades individuais, levando em conta fatores como idade, saúde e preferências pessoais. Neste mesmo texto também se pode encontrar diversas restrições para montagem de um cardápio acerca de: i) Alimentos ultraprocessados; ii) Frutas, verduras e legumes; iii) Alimentos fontes de ferro heme e não heme; iv) Gordura trans e saturada; v) Gorduras; vi) Sódio e açúcar simples; vii) Alimentos in natura ou minimamente processados. Dessa forma, foi feita uma primeira simulação de cardápio utilizando a ferramenta Solver do software Excel[®] para obtenção de solução via método simplex. Inicialmente, considerou-se apenas o café da manhã da unidade escolar, sendo utilizados 72 alimentos (TACO, 2011) e 5 restrições. O custo de todos os alimentos para esta simulação foi adotado como unitário com o objetivo de verificar os alimentos com melhores nutrientes. Esta hipótese permite verificar os alimentos capazes de melhor atender as restrições nutricionais estabelecidas. Como conclusão desta simulação, observou-se que a aveia, independente do preço, é um alimento capaz de fornecer todos os nutrientes necessários, contribuindo significativamente em todas as restrições. Obtiveram-se também resultados com soluções não viáveis e resultados com uma variedade maior de alimentos (aveia, abacate, goiaba, queijo prato

e ovo), em outras soluções retornaram o achocolatado como um alimento passível de figurar no cardápio. Em linhas gerais, pode-se perceber que os alimentos que figuraram o cardápio foram: aveia, queijo, pão e ovo. Esta simulação inicial permitiu melhor compreender o problema a ser modelado, considerando as particularidades a alimentação escolar.

Palavras-chave: Modelagem Matemática. Programação Linear. Problema da Dieta.

Agradecimentos

A pesquisa relatada neste trabalho foi apoiada pela FAPERJ por meio Edital nº 45/2021 - Apoio à melhoria das escolas da rede pública sediadas no Estado do Rio de Janeiro - 2021.

Referências

BRASIL. **Política Nacional de Alimentação e Nutrição**. Brasília : Ministério da Saúde/Secretaria de Atenção à Saúde, 2013. Disponível em:

<https://bvsmms.saude.gov.br/bvs/publicacoes/politica_nacional_alimentacao_nutricao.pdf>. Acesso em: 06 set. 2023.

BRASIL. **Planejamento de cardápios para a Alimentação Escolar**. Brasília: Ministério da Educação-MEC/Fundo Nacional de Desenvolvimento da Educação-FNDE, 2022. Disponível em:

<https://www.gov.br/fnde/pt-br/aceso-a-informacao/acoes-e-programas/programas/pnae/manuais-e-cartilhas/MANUAL_V8.pdf>. Acesso em: 06 set. 2023.

DANTZIG, G. B. Linear Programming. **Operations Research**. v. 50, n. 1, p. 42-47, 2002. Disponível em: <<https://pubsonline.informs.org/doi/10.1287/opre.50.1.42.17798>>. Acesso em: 06 set. 2023.

TACO. **Tabela Brasileira de Composição de Alimentos**. Versão 4. Campinas: Unicamp, 2011. Disponível em:

<https://www.cfn.org.br/wp-content/uploads/2017/03/taco_4_edicao_ampliada_e_revisada.pdf>. Acesso em: 06 set. 2023.

Estudo de um pêndulo paramétrico com mola torcional de liga com memória de forma para colheita de energia

Matheus Abrahão

*Universidade Federal Fluminense
matheusabrahao@id.uff.br*

Wesley Luiz da Silva Assis

*Universidade Federal Fluminense
wesleyassis@id.uff.br*

Alberto Paiva

*Universidade Federal Fluminense
paiva.ufrj@gmail.com*

Resumo: Este estudo investiga um pêndulo excitado parametricamente com uma mola de memória de forma, visando a aplicação na colheita de energia das ondas do mar. Compara a eficiência com um sistema similar usando uma mola convencional e analisa o comportamento do sistema em casos conservativos, dissipativos e de excitação paramétrica através de simulações numéricas.

Palavras-chave: Pêndulo Paramétrico. Dinâmica não linear. Colheita de Energia

Introdução

Nos últimos anos, houve um aumento significativo na população global, levando a uma demanda crescente por energia, estimando-se um acréscimo de 50% até 2023. Nesse contexto, a energia das ondas dos oceanos tem emergido como uma fonte limpa e renovável (Xu, 2005), portanto este estudo visa analisar a dinâmica não linear de um sistema composto por um pêndulo simples e uma mola de memória de forma, com excitação paramétrica, explorando seu potencial para a colheita de energia e contribuindo para soluções sustentáveis.

Metodologia

A análise abrangeu a modelagem matemática com o método Newton-Euler e a descrição mecânica da mola de memória de forma usando o modelo de Falk (1980). Após a obtenção das equações de movimento, os parâmetros foram adimensionalizados para simplificar a análise, com a integração numérica realizada pelo método de Runge-Kutta de quarta ordem em C++. As análises dinâmicas utilizaram ferramentas como seção de Poincaré, diagrama de bifurcação e expoente de Lyapunov. A pesquisa abordou um modelo composto por pêndulo, base oscilante, haste e mola torcional de memória de forma, considerando a excitação paramétrica das ondas para análise do sistema.

Resultados

Na Figura 1, para as diferentes temperaturas (baixa, intermediária e alta), é perceptível que a quantidade de pontos de equilíbrio estáveis diminui à medida que a temperatura aumenta, resultando também em mudanças no comportamento dinâmico do sistema. Na Figura 2, notamos que o pêndulo tende a adotar amplitudes de movimento maiores durante comportamento caótico, tornando-o favorável para a colheita de energia.

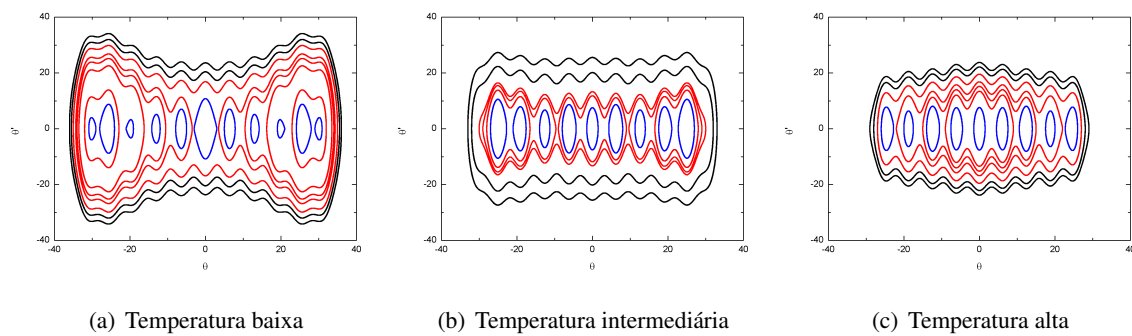


Figura 1: Curvas do espaço de fase de θ' (velocidade angular) por θ (deslocamento angular).

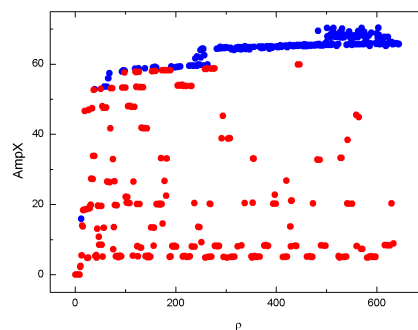


Figura 2: Gráfico da amplitude máxima obtida por simulação com relação a ρ (amplitude de forçamento). Onde os pontos em azul são para comportamentos caóticos (coeficiente de Lyapunov > 0) e os pontos em vermelho são para comportamentos periódicos (coeficiente de Lyapunov ≤ 0).

Conclusões

Os resultados revelaram que, à medida que a temperatura aumenta, o sistema exibe uma diminuição nos pontos de equilíbrio estáveis, indicando mudanças significativas em seu comportamento dinâmico. Além disso, durante comportamento caótico, o pêndulo tende a adotar amplitudes de movimento mais amplas, o que o torna favorável para a colheita de energia. Essas descobertas oferecem perspectivas promissoras para o desenvolvimento de soluções sustentáveis na captação de energia das ondas do mar, um recurso limpo e renovável em crescimento, contribuindo para atender à crescente demanda global por energia.

Referências

- FALK, F. Model free energy, mechanics, and thermodynamics of shape memory alloys. **Acta Metallurgica**, v. 28, n. 12, p. 1773–1780, 1980.
- PAIVA, A.; SAVI, M. A. An overview of constitutive models for shape memory alloys. **Mathematical problems in engineering**, v. 2006, p. 1–30, 2006.
- XU, X. **Nonlinear dynamics of parametric pendulum for wave energy extraction**. 2015. Thesis (Doctor of Philosophy) – University of Aberdeen, Aberdeen, 2005.

Homogeneização Assintótica Aplicada a Uma Barra Elástica Cilíndrica

Luiz Gustavo da Silva Pereira
Universidade Federal Fluminense
luizgsp@id.uff.br

Reinaldo Rodriguez Ramos
Universidade Federal Fluminense
reinaldorr@id.uff.br

Panthers Rodríguez Bermudez
Universidade Federal Fluminense
pantersrb@id.uff.br

Jorge Alberto Rodriguez Durán
Universidade Federal Fluminense
jorgeduran@id.uff.br

Resumo: Neste trabalho será abordado o **Método da Homogeneização assintótica** produzindo uma **Solução Assintótica Formal** de segunda ordem, a partir destes são determinados os coeficientes efetivos e proposta a formulação matemática do problema homogeneizado. Por fim, diferentes simulações numéricas são realizadas para a resolução das equações diferenciais, com o intuito de demonstrar a aproximação entre a SAF e a solução analítica.

Palavras-chave: Matemática Aplicada. Homogeneização Assintótica.

Introdução

O seguinte trabalho visa apresentar o **Método da Homogeneização Assintótica**, que vem sendo usado extensivamente nas áreas de matemática, engenharia e materiais, pois é capaz de facilitar a análise de sólidos heterogêneos com estrutura microperiódica. Para isso será analisado o deslocamento de um sólido heterogêneo submetido à tensões puramente radiais.

Submissão

O sólido cujo deslocamento será calculado apresenta uma periodicidade espacial do módulo de Young e a razão de Poisson e sua célula periódica unitária está representada na Figura 1. A equação que descreve o deslocamento do sólido e cuja solução será abordada neste trabalho é

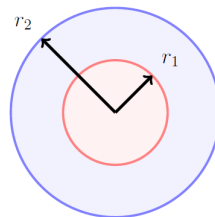


Figura 1: Célula periódica unitária do cilindro não homogeneizado

$$\frac{d}{dr}(C_r^\varepsilon \frac{du_r^\varepsilon}{dr}) + \frac{C_r^\varepsilon}{r}(\frac{du_r^\varepsilon}{dr}) - f_r^\varepsilon = 0, \quad u^\varepsilon(0) = 0, \quad u^\varepsilon(1) = 1, \quad (1)$$

onde $C_r^\varepsilon = 2\mu_r^\varepsilon + \lambda_r^\varepsilon = \frac{E_r^\varepsilon(1-\nu_r^\varepsilon)}{(1+\nu_r^\varepsilon)(1-2\nu_r^\varepsilon)}$ e ε um parâmetro geométrico pequeno definido como $\varepsilon = \frac{l}{L}$, e y , geralmente chamada de variável lenta, uma nova variável radial que varia de forma mais lenta em relação a r . Sendo assim, supõe-se a seguinte solução assintótica formal de segunda ordem para a homogeneização do problema

$$u^{(2)}(r, \varepsilon) = u_0(r, y) + \varepsilon u_1(r, y) + \varepsilon^2 u_2(r, y), \quad (2)$$

substituindo (2) em (1) e solucionando, chega-se em (3), a equação para o problema local, e (4) a homogeneização do problema original (1). Ressalte-se que a função $N_1(y)$ é de grande importância, pois será usada no cálculo de u_1 e u_2 da SAF (2)

$$\frac{d}{dy} \left(C(y) \frac{dN_1(y)}{dy} + C(y) \right) = 0, \quad N_1(0) = 0, \quad (3)$$

$$\frac{d^2 u_0}{dr^2} \left(\lambda_0 \hat{C} + \frac{\lambda_1}{r} \right) + \frac{du_0}{dr} \left\langle C \left(\frac{r}{\varepsilon} \right) \right\rangle = \hat{f} \left(\frac{r}{\varepsilon} \right), \quad \lambda_0 = \frac{1}{2} \left(\frac{r_2}{\varepsilon} \right)^3, \quad \lambda_1 = \int_0^{\frac{r_2}{\varepsilon}} C(y) N_1(y) y dy. \quad (4)$$

Ressalte-se que a equação homogeneizada (4) é a de maior importância, já que irá ditar o comportamento efetivo do sólido na macroescala. A Figura 2 demonstra as soluções para (3) e (4) para diversos valores de ε utilizando $C(r) = 1 + 0.25 \cos[2\pi(\frac{r}{\varepsilon})]$ e $f(r) = \cos[2\pi(\frac{r}{\varepsilon})]$.

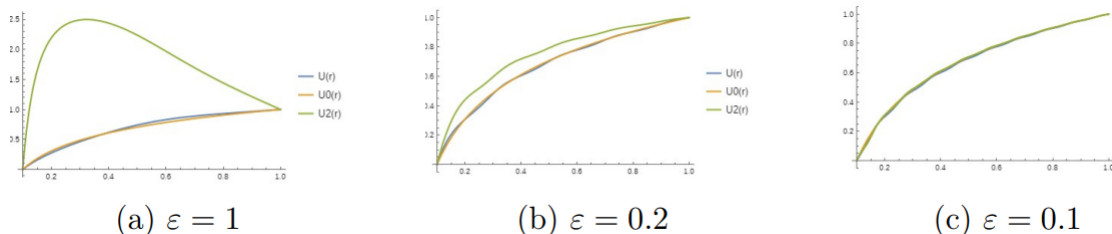


Figura 2: Ilustração da aproximação entre as curvas $u(r)$ solução analítica e a Solução Assintótica Formal representada por $u_0(r)$ e $u_2(r)$.

Conclusões

Como foi possível observar o **Método da Homogeneização Assintótica** produz soluções efetivas que tendem assintoticamente ao problema original, tornando a análise de tensões significativamente mais eficiente.

Referências

LEITZKE, Bruna da Silva. **Soluções assintóticas formais de segunda ordem na homogeneização assintótica de meios microperiódicos**. 2017. Dissertação de Mestrado. Universidade Federal de Pelotas.

SADD, Martin H. **Elasticity: theory, applications, and numerics**. Academic Press, 2009.

ALGUMAS CLASSES DE GRAFOS COM ÍNDICE INTEIRO

Bruna Lima de Oliveira

Universidade Federal Fluminense

Bl_oliveira@id.uff.br

Andre Ebling Brondani Universidade Federal Fluminense

andrebrondani@id.uff.br

Francisca Andrea Macedo Franca

Universidade Federal Fluminense

francisca_franca@id.uff.br

Resumo: Seja $G = (V, E)$ um grafo simples com n vértices. A matriz de adjacência de G , $A(G) = [a_{ij}]$, é a matriz quadrada de ordem n cujas entradas são dadas por $a_{ij} = 1$ quando $\{v_i, v_j\} \in E$ e 0 nos demais casos. O polinômio característico de G é definido por $\varphi(G, x) = \det(xI - A)$ e suas raízes são denominadas autovalores de G . Como A é uma matriz simétrica, todos os autovalores de G são números reais e, portanto, podem ser ordenados em ordem não crescente $\lambda_1(G) \geq \lambda_2(G) \geq \dots \geq \lambda_n(G)$. O maior autovalor de G é conhecido como índice de G e o multiconjunto constituído por todos os autovalores de G é denominado espectro de G e é denotado por $\sigma(G)$. O grafo G é dito integral se $\sigma(G)$ é composto somente por números inteiros. “Which graphs have integral spectra?” Este problema foi proposto em (HARARY; SCHWENK, 1974) com a imediata observação de que o problema em geral parece ser intratável. De fato, o número de grafos integrais não é apenas infinito, mas podemos encontrá-los em todas as classes de grafos e entre grafos de todas as ordens. Para mais detalhes, sugerimos (BALINSKA *et al.*, 2002). Grafos cujo índice é inteiro mostraram-se importantes no estudo de redes quânticas, como pode ser visto em (SAXENA *et al.*, 2007), o que propicia um vasto campo de estudo interdisciplinar. Motivados por este problema, pesquisamos subclasses de grafos integrais dentre as classes de grafos que definimos a seguir, que ilustramos posteriormente na Figura 1.

Definição. Dados os inteiros n e ℓ , onde $n \geq 3$ e $1 \leq \ell \leq n$, o grafo H_n^ℓ é obtido de duas cópias do grafo completo K_n ao adicionarmos ℓ arestas entre ℓ vértices de uma cópia de K_n e seus vértices correspondentes na outra cópia.

Definição. Dados os inteiros $p \geq 3$ e $q \geq 1$, o grafo abacaxi, K_p^q , é obtido do grafo completo K_p ao adicionarmos q arestas pendentes a um de seus vértices.

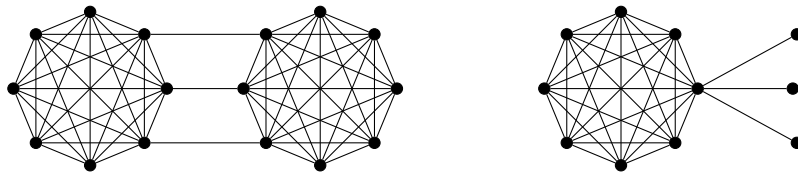


Figura 1: Grafo H_8^3 e o grafo abacaxi K_8^3 , respectivamente.

Uma condição necessária para que um grafo seja integral é que seu índice seja um número inteiro. Seguindo essa linha, determinamos algumas subclasses de grafos do tipo H_n^ℓ com índice inteiro e obtemos uma caracterização para que tais grafos sejam integrais, de acordo com o seguinte resultado.

Teorema. Para que o grafo H_n^ℓ seja integral é necessário e suficiente que $\ell = n$.

Na Figura 2 exibimos os grafos H_3^2 e H_3^3 cujos espectros são

$$\sigma(H_3^2) = \{-2, -\sqrt{2}, 1 - \sqrt{3}, 0, \sqrt{2}, 1 + \sqrt{3}\} \text{ e } \sigma(H_3^3) = \{-2, -2, 0, 0, 1, 3\}.$$

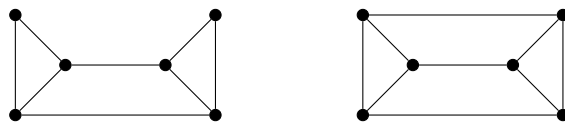


Figura 2: Grafos H_3^2 e H_3^3 .

Dentre os resultados obtidos ao investigarmos a classe dos grafos abacaxi, destacamos o seguinte resultado.

Teorema. Se $q \geq 1$ e $p \geq \frac{-1+\sqrt{8q+1}}{2}$, então o grafo abacaxi K_p^q não é integral. Além disso, existem duas subclasses infinitas de grafos abacaxi com índice inteiro.

Para exemplificar o teorema anterior, considere os grafos abacaxi K_3^6 e K_4^{10} cujos polinômios característicos são

$$\varphi(K_3^6, x) = x^5(x+1)(x-3)(x^2+2x-2) \quad \text{e} \quad \varphi(K_4^{10}, x) = x^9(x+1)^2(x-4)(x^2+2x-5).$$

Observe que os dois polinômios quadráticos são irredutíveis sobre o corpo dos números racionais e que $\lambda_1(K_3^6) = 3$ e $\lambda_1(K_4^{10}) = 4$.

Palavras-chave: Matriz de adjacência. Espectro. Índice inteiro. Grafos integrais.

Agradecimentos

A primeira autora agradece o apoio financeiro recebido da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro - FAPERJ, processo número E-26/204.492/2022.

Referências

- BALINSKA, K.; CVETKOVIĆ, D.; RADOSAVLJEVIĆ, Z.; SIMIĆ, S.; STEVANOVIĆ, D. A survey on integral graphs. **Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat.**, v. 13, p. 42-65, 2002.
- HARARY, F.; SCHWENK, A. J. Which graphs have integral spectra? In: BARI, R.; HARARY, F. (Edit.). **Graphs and Combinatorics**. Berlin: Springer-Verlag, 1974. p. 45-51.
- SAXENA, N.; SEVERINI, S.; SHPARLINSKI, I. Parameters of integral circulant graphs and periodic quantum dynamics. **International Journal of Quantum Information**, v. 5, p. 417-430, 2007.

A DECOMPOSIÇÃO EM VALORES SINGULARES E UMA APLICAÇÃO NA COMPRESSÃO DE IMAGENS

Viviane da Silva Pimentel

*Universidade Federal Fluminense, Instituto de Ciências Exatas, Volta Redonda, Rio de Janeiro
viviane.pimentel@id.uff.br*

Alan Prata de Paula

*Universidade Federal Fluminense, Instituto de Ciências Exatas, Volta Redonda, Rio de Janeiro
alanprata@id.uff.br*

Marina Sequeiros Dias de Freitas

*Universidade Federal Fluminense, Instituto de Ciências Exatas, Volta Redonda, Rio de Janeiro
msdias@id.uff.br*

Resumo: A Decomposição em Valores Singulares emerge como uma das mais proeminentes ferramentas de fatoração matricial. Neste estudo, exploraremos esse método de fatoração, juntamente com uma aplicação concisa na compressão de imagens, visando a redução de dimensionalidade e mantendo uma qualidade visual aceitável.

Palavras-chave: Decomposição em Valores Singulares. Fatoração Matricial. Compressão de Imagens.

Introdução

A Decomposição em Valores Singulares (SVD, sigla do nome em inglês *Singular Value Decomposition*) é uma técnica que desempenha um papel fundamental no processamento de imagens. Este estudo visa aprofundar a compreensão da SVD, explorando sua aplicação na compressão de imagens, mostrando sua capacidade em aprimorar a análise visual das imagens, além de destacar as métricas de comparação do método.

A Decomposição em Valores Singulares

A Decomposição em Valores Singulares garante, basicamente, que qualquer matriz pode ser decomposta em três partes fundamentais: duas matrizes de rotação e uma matriz diagonal. A seguir, observe o teorema que garante a sua existência.

Teorema 1 *Seja $A \in M_{m \times n}(\mathbb{R})$ uma matriz com posto r . Então existe uma matriz*

$$\Sigma_{m \times n} = \begin{bmatrix} D_{r \times r} & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix},$$

onde os elementos da diagonal principal da matriz D são os r valores singulares não-nulos de A , isto é, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$. Mais ainda, existem matrizes ortogonais $U_{m \times m}$ e $V_{n \times n}$ tais que $A = U \Sigma V^T$. O produto $U \Sigma V^T$ é denominado uma Decomposição em Valores Singulares para a matriz A .

Além disso, existem resultados que possibilitam compactar os dados da matriz ao ignorar os menores valores singulares. O primeiro resultado garante que “Dada uma matriz $A \in M_{m \times n}(\mathbb{R})$ que possui exatamente r valores singulares não-nulos, isto é, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ e $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$. Então, o posto de A é r ”. Somado a isso, Trefethen (1997, p. 35) garante que “ A é a soma de r matrizes de posto 1, isto é, $A = \sum_{j=1}^r \sigma_j u_j v_j^T$ ” e que “Para algum v com $0 \leq v \leq r$, definindo $A_v = \sum_{j=1}^v \sigma_j u_j v_j^T$, se $v = p = \min\{m, n\}$ e $\sigma_{v+1} = 0$, então $\|A - A_v\| = \inf \|A - B\|$, onde $B \in \mathbb{C}^{m \times n}$ e $\text{posto}(B) \leq v$ ”.

Utilizando esses resultados, é possível garantir a melhor aproximação da matriz utilizando os primeiros v valores singulares, além de reduzir a dimensão da matriz. Esse processo facilita a transmissão de dados via redes e a extração de informações relevantes. A seguir, de maneira prática, exploraremos uma aplicação na compressão de imagens.

Compressão de Imagens Digitais

Na compressão de imagens, o foco é reduzir o armazenamento computacional das imagens, tornando-as de mais fácil transferência pelas redes. Nesse trabalho, através da SVD, abordaremos essa aplicação com imagens em tons de cinza (8 bits). Matricialmente, podemos representar essa imagem digital como uma matriz A bidimensional ($m \times n$) em que cada elemento da matriz representa o valor de intensidade de um pixel específico. Através da decomposição SVD podemos descrever a matriz A como $A = \sum_{i=1}^r \sigma_i u_i v_i^T$. Considerando apenas os k primeiros valores singulares, teremos $A = \sum_{i=1}^k \sigma_i u_i v_i^T$. Além disso, segundo Cao (2006), podemos avaliar a qualidade da compressão através de duas métricas: taxa de compressão e erro quadrático médio (*Mean Square Error* - MSE). A taxa de compressão C_R é dada por $C_R = \frac{mn}{k(m+n+1)}$. Já o erro quadrático médio é dado por $MSE = \frac{1}{mn} \sum_{x=1}^n \sum_{y=1}^m (f(x,y) - f_c(x,y))^2$, onde $f(x,y)$ é o valor do pixel na imagem original e $f_c(x,y)$, o valor do pixel após a compressão. Note que quanto menor for número k , maior será a taxa de compressão C_R , ou seja, a imagem estará mais distante de sua versão original, basta observar que, consequentemente, o MSE será maior. Ambas as métricas medem a qualidade entre a imagem original e a compactada. Na Tabela 1 da próxima seção, isso será mais perceptível.

Conclusões

Nesse estudo, notamos que a redução de dimensionalidade obtida ao descartar os valores singulares menos significativos permite uma representação mais compacta da imagem, o que resulta em economia de espaço de armazenamento e maior eficiência na transmissão. As métricas de comparação utilizadas são úteis e mostram, de maneira eficiente, a proporção dessa redução. Observe a Tabela 1 abaixo.

Tabela 1: Tabela de comparação das métricas.

Valores singulares	C_R	MSE	Armazenamento (bytes)
20	0.04	386.3451	701739
60	0.13	125.8883	781739
300	0.65	7.2076	1261739
500	1.08	0.5047	1661739
720	0.64	3.8094e-8	2101739

Referências

- BORTOLOSSI, H. J. **Tópicos em Álgebra Linear: Decomposições Matriciais e Aplicações**. Universidade Federal Fluminense, 2017.
- CAO, L. Singular value decomposition applied to digital image processing. **Division of Computing Studies, Arizona State University Polytechnic Campus**, 2006.
- TREFETHEN, L. N.; BAU, I. D. Numerical Linear Algebra. **Philadelphia: Society for Industrial and Applied Mathematics**, 1997.

Caotificação de um circuito eletrônico análogo ao oscilador de Duffing

Josiane Campos, André Gusso, Wellington G. Dantas, Sebastian Ujevic
Departamento de Ciências Exatas, EEIMVR, UFF
andregusso@id.uff.br

Resumo: A excitação de duas frequências emergiu recentemente como um método eficiente para gerar forte caotificação de sistemas dinâmicos do tipo Duffing e Duffing com potenciais de poço único e duplo, tendo sido um método proposto pelos autores. Neste trabalho demonstramos experimentalmente a caotificação da dinâmica de um circuito eletrônico análogo ao oscilador de Duffing com poço duplo.

Palavras-chave: Caos. Caos robusto. Dinâmica não-linear. Circuito eletrônico.

Modelagem do circuito eletrônico

Neste trabalho modelamos fisicamente o circuito análogo ao oscilador de Duffing com poço duplo proposto por Tamaseviciute (2008). A equação de Duffing tem a forma geral

$$m\ddot{x} + c\dot{x} - kx + k_3x^3 = F(t). \quad (1)$$

O circuito eletrônico possui uma equação análoga a esta, apresentando o característico potencial de poço duplo resultante do termo de força não-linear $-kx + k_3x^3$. As variáveis físicas escolhidas para analisar a dinâmica do circuito são a corrente no indutor (I_L) e a tensão no capacitor (V_C) e resultam no seguinte sistema de equações

$$\begin{aligned} L \frac{d}{dt} I_L &= gV_D(V_C) - V_C - RI_L - g'V(t), \\ C \frac{d}{dt} V_C &= I_L - I_D(V_C). \end{aligned} \quad (2)$$

onde $g = 1 + R_2/R_1$ e $g' = R_1/R_2$. Com uma adequada mudança de variáveis o sistema $x = V_C, y = \rho I_L, \tau = t/\sqrt{LC}, \Omega = \omega\sqrt{LC}, b = R'/\rho$, e $\rho = \sqrt{L/C}$ obtemos o fluxo em termos das três equações

$$\begin{aligned} \frac{dx}{d\tau} &= y - \rho I_D(x) \\ \frac{dy}{d\tau} &= F(x) - by + g'A[\sin(z) + a\sin(z/r)] \\ \frac{dz}{d\tau} &= \Omega, \end{aligned} \quad (3)$$

onde

$$F(x) = (g - 1)x - \text{sign}(x)gI_D(|x|)R_3. \quad (4)$$

corresponde o potencial não-linear análogo ao potencial de poço duplo do oscilador de Duffing. Nestas equações R, R_1, R_2 e R_3 denotam resistências do circuito e L e C o indutor e o capacitor do circuito RLC série nele implementado.

Neste trabalho resolvemos numericamente o sistema de equações (3) considerando excitação tanto por uma quanto por duas frequências, buscando encontrar as regiões no espaço de parâmetros relevante onde obtivéssemos uma dinâmica caótica. Também construímos o circuito e desenvolvemos um sistema de medidas automatizadas que permitiram obter resultados experimentais para compararmos com as previsões teóricas.

Resultados

Na figura 1(a) temos uma comparação entre os resultados experimentais e os teóricos para o caso de excitação por uma frequência. Esta é a configuração normalmente encontrada na literatura. Na figura 1(b) temos uma comparação entre os resultados experimentais e os teóricos para o caso de excitação por duas frequências com razão $r = 1.07$ entre si. As subfiguras mostram a dinâmica obtida variando-se a frequência de excitação f e sua amplitude, A .

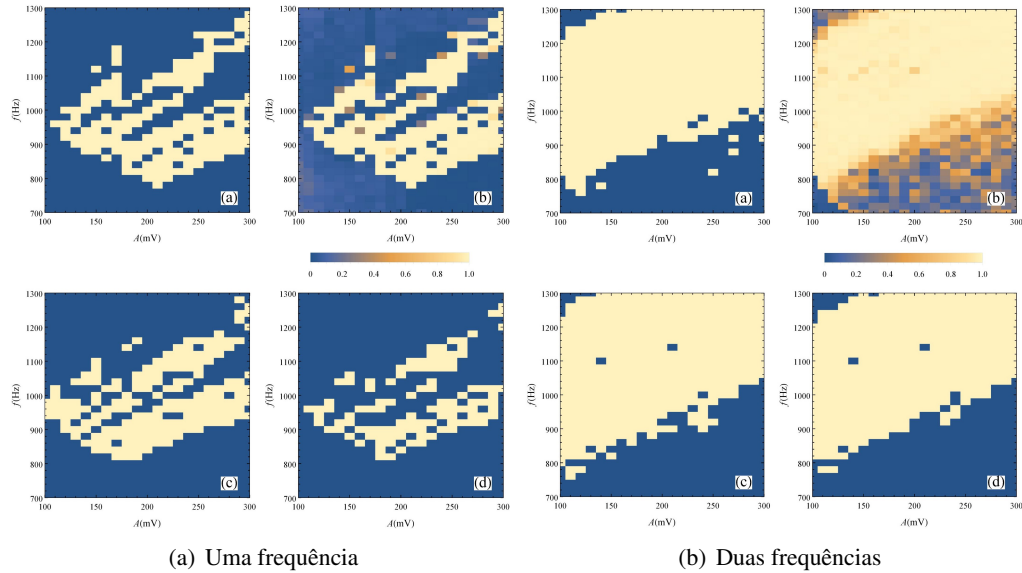


Figura 1: Resultado experimentais que identificam em cores claras pontos com dinâmica caótica e em cores escuras aqueles com dinâmica regular para os casos de excitação por uma (a) e duas frequências (b) com razão $r = 1.07$. Para cada subfigura em (a) e (b) resultados experimentais, (c) previsão teórica do modelo e (d) pontos de superposição teórico-experimental.

Conclusões

Na figura 1(a), observamos regiões com caos misturadas com aquelas com regularidade. Esta situação torna difícil usar o circuito como uma fonte de sinais caóticos. Vemos, também, que há boa concordância entre os resultados experimentais e as expectativas teóricas. Por sua vez, na figura 1(b) pode-se ver que se obteve uma grande região praticamente contínua com caos. Assim, fica demonstrada experimentalmente a eficácia do método de excitação por duas frequências como um novo método de caotificação, como previsto inicialmente por Gusso et al. (2019). A boa concordância entre teoria e experimento também neste caso mostra que o modelo teórico foi suficientemente preciso e realista.

Referências

GUSSO, A.; DANTAS, W.G.; UJEVIC, S. Prediction of robust chaos in micro and nanoresonators under two-frequency excitation. **Chaos**, v. 29, 033112, 2019.

TAMASEVICIUTE, E. et al. Analogue electrical circuit for simulation of the Duffing-Holmes equation. **Nonlinear Analysis Modelling and Control**, v. 13, p. 241-252, 2008.

DESENVOLVENDO UM AMBIENTE COMPUTACIONAL PARA IDENTIFICAR OS PARÂMETROS QUE CARACTERIZAM AS CURVAS DE POLARIZAÇÃO

Letícia Vieira Dias
ICE/UFRRJ, Seropédica, RJ
leticiavdias04@gmail.com

Fabício Magalhães Coutinho
PESQ/UFRRJ, Rio de Janeiro, RJ
fabricao.mrj@hotmail.com

Renan de Souza Teixeira
PPGMMC/UFRRJ, Seropédica, RJ
renannovoser@gmail.com

Wilian Jeronimo dos Santos
PPGMMC/UFRRJ, Seropédica, RJ
wilian@ufrj.br

Resumo: O objetivo deste trabalho é desenvolver um software que encontre, de forma ótima, os parâmetros da curva de polarização que se ajustem aos dados observados. Para isso, é resolvido um problema de mínimos quadrados não linear com auxílio de um algoritmo genético (AG) e do método da região de confiança (MRC).

Palavras-chave: Algoritmo Genético. Otimização não linear. Problemas de corrosão.

Introdução

No que tange o problema da corrosão, uma equação muito utilizada nesse caso é a equação de Wagner-Traud (WOLYNEC, 2003, p.94), a qual busca ajustar as curvas de polarização e é dada por

$$\Delta I = I^* \left[\exp\left(\frac{2,303\Delta E}{b_a}\right) - \exp\left(\frac{2,303\Delta E}{b_c}\right) \right], \quad (1)$$

onde ΔI é a diferença entre a corrente de polarização anódica e catódica, I^* é a corrente de corrosão, ΔE é a diferença de potencial e b_a e b_c são os declives de Tafel anódico e catódico, respectivamente. Para encontrar esses parâmetros é necessário o auxílio de métodos numéricos para solucionar um problema de mínimos quadrados não linear. Neste trabalho, um AG é inicialmente utilizado para resolver tal problema de mínimos quadrados, de forma a explorar o espaço de busca. Em seguida, o resultado que o AG obtém serve como chute inicial para o método determinístico MRC, a fim de buscar um ajuste mais preciso, objetivando a convergência para o ótimo global. O software está em fase de desenvolvimento, sendo implementado em Python. O algoritmo genético foi implementado a partir do apresentando por Linden (2008) e para MRC foi utilizado a biblioteca Scipy. Para as telas foi utilizada a biblioteca PyQt, sendo ilustradas nas Figuras 1 e 2.

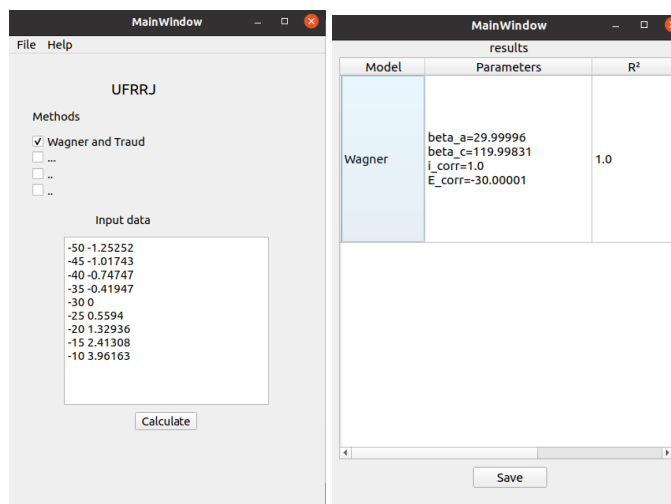


Figura 1: A tela inicial com dos dados de potencial e densidade de corrente (à esquerda) e a tela de resultados, com os parâmetros de ajuste encontrados (à direita).

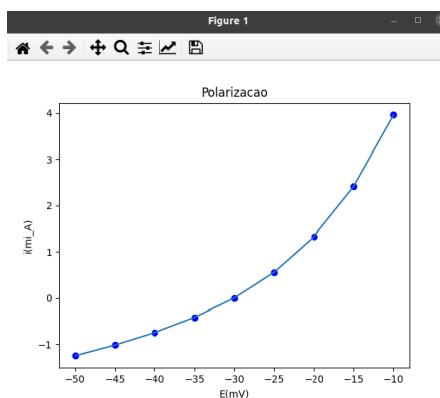


Figura 2: A tela de gráfico com a curva de polarização ajustada.

Conclusões

Apesar do software ainda se encontrar em fase de desenvolvimento e aprimoramentos, os primeiros testes se mostraram promissores na busca da elaboração de uma ferramenta computacional que possa ser útil para estudantes e pesquisadores que trabalham com problemas relacionados à corrosão.

Agradecimentos

A Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) e a Universidade Federal Rural do Rio de Janeiro (UFRRJ).

Referências

- LINDEN, R. **Algoritmos genéticos - Uma importante ferramenta da inteligência computacional**. 2. ed. Rio de Janeiro: Brasport, 2008.
- WOLYNEC, S. **Técnicas eletroquímicas em corrosão**. EDUSP, 2003.

ESTUDO DA DIFUSÃO DO HIDROGÊNIO EM AGLOMERADO DE Fe₈ ATRAVÉS DO MÉTODO DA TEORIA DO FUNCIONAL DE DENSIDADE

Nathan Matheus da Silva

Universidade Federal Fluminense/EEIMVR

nathanmatheus@id.uff.br

Lilian Weitzel Coelho Paes

Universidade Federal Fluminense/EEIMVR/VCE

lilianweitzel@id.uff.br

Resumo: O processo de difusão de hidrogênio é uma questão crucial relacionada à fragilização por hidrogênio. É bem conhecido que o hidrogênio degrada as propriedades do ferro e aços. A presença de interstícios na estrutura cúbica de corpo centrado (CCC) do ferro influencia significativamente a difusividade de H, portanto, é um aspecto importante do problema na fragilização por hidrogênio (DÍAZ; ALEGRE; CUESTA, 2016). Os efeitos causados pela difusão do hidrogênio nas propriedades dos materiais estão bem documentados, mas não são bem compreendidos. Uma investigação experimental além de mais dispendiosa teria algumas limitações do ponto de vista prático, como por exemplo o acompanhamento do caminho de difusão dos átomos (ITAKURA; MIYAUCHI; MURASE; YAKABE *et al.*, 2021). Os modelos de difusão predizem que estruturas cúbicas de corpo centrado (CCC) são propensas ao enfraquecimento. Os átomos de hidrogênio podem saltar de um sítio intersticial para outro vago vizinho e podem se difundir dessa maneira por grandes distâncias através do metal. Desta forma, é de suma importância adquirir um conhecimento do mecanismo de difusão de átomos em pequenos aglomerados metálicos para compreender o fenômeno mais detalhadamente, por esta razão investigações são realizadas em escalas atômicas sobre estado eletrônico de menor energia (BODA; ALI, 2022). Os cálculos da teoria do funcional de densidade (TFD) têm se apresentado como um método eficiente para o estudo das propriedades geométricas e eletrônicas, sendo uma das vantagens a facilidade que esse método permite de se explorar as propriedades dos materiais em nível atômico/molecular. Neste trabalho, cálculos TFD foram realizados em um pequeno aglomerado de ferro, com o intuito de investigar o comportamento da difusão de hidrogênio. Os cálculos foram realizados utilizando o funcional BPW91 e o conjunto de bases Lanl2dz para o átomo de ferro e 6-311+G(d,p) para o átomo de hidrogênio. O átomo de hidrogênio deslocou-se no aglomerado de ferro do interstício octaédrico inicial para o interstício tetraédrico localizado no centro da aresta da célula. Foi observado que a célula permanece com a mesma geometria (quatro átomos de Fe) em torno de H, porém ocorreu um alongamento na distância Fe-Fe, cerca de 0,06 Å quando H está presente no sítio tetraédrico. Isso demonstra que H se encaixa bem no interstício tetraédrico. Essa distorção está em concordância com dados da literatura (JIANG; CARTER, 2004). A variação de entalpia para a difusão calculada entre os interstícios octaédrico e tetraédrico foi de 2,18 eV (50,2 Kcal/mol). A utilização do método TFD comprovou ser possível a simulação da difusão em pequenos aglomerados de ferro e mostrou ser viável em velocidade computacional quando comparado à métodos mais sofisticados.

Palavras-chave: TFD. Difusão. Hidrogênio. Aglomerado de ferro.

CNPq/FAPERJ/UFF

BODA, A.; ALI, S. M. First principles DFT analysis on the diffusion kinetics of hydrogen isotopes through bcc iron (Fe): Role of temperature and surface coverage. **International Journal of Hydrogen Energy**, 47, n. 73, p. 31481-31498, 2022.

DÍAZ, A.; ALEGRE, J. M.; CUESTA, I. I. A review on diffusion modelling in hydrogen related failures of metals. **Engineering Failure Analysis**, 66, p. 577-595, 2016.

ITAKURA, A. N.; MIYAUCHI, N.; MURASE, Y.; YAKABE, T. *et al.* Model of local hydrogen permeability in stainless steel with two coexisting structures. **Scientific Reports**, 11, n. 1, p. 8553, 2021.

JIANG, D. E.; CARTER, E. A. Diffusion of interstitial hydrogen into and through bcc Fe from first principles. **Physical Review B**, 70, n. 6, p. 064102, 2004.

Análise do impacto do quadro de horário de disciplinas sobre a frequência estudantil

Carolina Pinheiro dos Santos
Universidade Federal Rural do Rio de Janeiro
carol30p@ufrj.br

Ronaldo Malheiros Gregório
Universidade Federal Rural do Rio de Janeiro
rgregor@ufrj.br

Felipe Leite Coelho da Silva
Universidade Federal Rural do Rio de Janeiro
felipeleite@ufrj.br

Yasmin Vitória Lima de Jesus
Universidade Federal Rural do Rio de Janeiro
yasminvitorialima@ufrj.br

Introdução

A Lei nº 11.274 de 6 de fevereiro de 2006 institui a matrícula obrigatória, gratuita em escolas públicas, a partir dos 6 anos de idade, assegurando tempo prolongado de permanência escolar aos estudantes através dos nove anos do ensino fundamental. Baseado nisso, o Plano Nacional de Educação de 2014, que estabelece metas e estratégias para políticas educacionais até o ano de 2024, objetiva a universalização do ensino fundamental e sua conclusão na idade recomendada (14 anos) para, pelo menos, 95% dos alunos. É também meta do PNE “fomentar a qualidade da educação básica em todas as etapas e modalidades, com melhoria do fluxo escolar e da aprendizagem” (Lei 13.005/2014). Uma das estratégias mencionadas é que, em seu último ano de vigência, os estudantes de ensino fundamental e médio tenham alcançado nível suficiente de aprendizado, e pelo menos 80% tenham alcançado um nível desejável. Outra medida governamental, visando a universalização da renda básica de cidadania, o Programa Bolsa Família, requer o cumprimento da frequência escolar mínima de 75% para beneficiários de 6 a 18 anos de idade incompletos que não tenham concluído a educação básica. Para o cumprimento das metas educacionais não basta, portanto, que se garanta números inflados de matrículas sem que esses alunos se mantenham na escola com bom aproveitamento.

Há numerosa literatura voltada aos parâmetros externos ao ambiente escolar que influenciam a frequência estudantil por perspectivas econômicos, sociais, étnicos, entre outros. No entanto, existem ações internas, passíveis de implementação pela gestão escolar, que podem tornar a escola mais convidativa, incentivando a participação estudantil de forma regular. Aqui apresentaremos um estudo da relação observada entre o quadro de horários semanal e a presença dos alunos em sala, visando apontar melhoramentos que poderiam influenciar positivamente a taxa de frequência estudantil.

Metodologia e resultados

O estudo parte do questionamento da existência de disciplinas com pesos positivos ou negativos à frequência estudantil, afetando a taxa de frequência ao longo da semana. Para estabelecer essas conexões, foi feito o levantamento dos dados de 42 turmas ao longo do ano de 2022 de uma escola estadual, contando com turmas de diferentes turnos e níveis, do Ensino Fundamental II, Ensino Médio e NEJA. Os dados coletados foram: frequência por turma, quadro de horários e alocação de professores nas disciplinas. Em cima dos dados de frequência atrelados aos dias da semana, foi realizada uma análise exploratória e, em seguida, utilizada a análise de

variância de 1-fator (ANOVA) que, utilizando o teste F, avalia se dentre os dias da semana há igualdade das médias de frequência. Para complementar essa análise, foi acrescentado o teste de Tukey para avaliar a existência ou inexistência de correlações entre a frequência estudantil e os quadros de horários formulados para as turmas. Mais detalhes sobre ANOVA podem ser encontrados em Bussab e Morettin (2011).

Os resultados das análises de variância apontaram para a sexta-feira comumente relacionada à queda significativa da frequência estudantil. No Ensino Médio, observou-se que 77% das turmas apresentava a menor média de frequência às sextas e 71% apresentava a maior média às terças. Investigações acerca dos problemas relacionados à sexta-feira apontaram a alocação preferencial de disciplinas eletivas, disciplinas sem professores e tempos vagos. Já as terças-feiras concentravam mais tempos de aula de matemática e língua portuguesa. Além disso, no Ensino Fundamental II pouco se difere: 83% das maiores médias de frequência observadas encontravam-se entre as terças e quartas, enquanto 89% das menores médias de frequência continuavam relacionadas às sextas-feiras. Matemática continuava predominantemente relacionada aos dias de maior frequência, enquanto as sextas concentravam arte, ciências e língua portuguesa. Já para o NEJA, cujo dia relacionado às maiores médias de frequência foram em 63% das turmas as terças, focadas em matemática e língua portuguesa, e dia relacionado às menores médias foram em 88% das turmas as sextas, concentrando disciplinas de biologia e arte.

De uma maneira geral, 84% das turmas apresentaram menor frequência às sextas-feiras. Essas escolhas ao se organizar o quadro de horários estudantil podem ter impactado diretamente a escolha dos estudantes de se ausentar às sextas, implicando em possíveis faltas desnecessárias que prejudicam as matérias que foram dadas nesses dias.

Conclusão

Considerando o interesse de se manter o engajamento estudantil constante, não somente para que as normas nacionais que requerem altas taxas de frequência estudantil sejam cumpridas, mas também buscando o sucesso escolar de todos os estudantes, deve-se agir para corrigir a distribuição desigual das médias de frequência ao longo da semana. Sabendo que a alocação das disciplinas e tempos vagos afeta a decisão estudantil de frequentar ou não a escola em um determinado dia, essa pesquisa abre precedentes para a criação de um conjunto de ações de controle que podem ser incorporadas a modelagem do quadro de horários das turmas.

Referências

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Censo da Educação Básica 2022: notas estatísticas**. Brasília, DF: Inep, 2023.

BRASIL. **Lei Federal n.º 11.274, de 06/02/2006**. Brasília: MEC, 2006.

BRASIL. **Medida Provisória n.º 1.164, de 02/03/2023**. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF. 2023.

BRASIL. Plano Nacional de Educação (PNE). **Lei Federal n.º 13.005, de 25/06/2014**. Brasília: MEC, 2014.

BUSSAB, W.; MORETTIN, P. **Estatística Básica**. Editora Saraiva, 7ª edição. São Paulo. 2011.

Aplicação do modelo de Amortecimento Exponencial aos dados de consumo de energia elétrica do setor industrial da região sudeste no Brasil

Marcos F. S. Carvalho
DEMAT/UFRRJ, Seropédica, RJ
marcoscarvalho56789@gmail.com

Felipe L. C. Silva
DEMAT/UFRRJ, Seropédica, RJ
felipeleite@ufrj.br

Resumo: A avaliação do consumo de energia elétrica desempenha um papel crucial no planejamento energético do Brasil. Nesse contexto, é relevante enfatizar que o setor industrial assume a posição de liderança como o maior consumidor de energia no país, sendo a região Sudeste aquela com a maior demanda. Neste estudo, foi investigado o modelo de amortecimento exponencial para prever o consumo de energia com base em séries temporais. Utilizam-se técnicas de erro para avaliar a precisão desse modelo e, em seguida, são comparados com o modelo ingênuo sazonal, onde o modelo de amortecimento exponencial apresenta maior precisão nas suas previsões, comparado ao modelo ingênuo sazonal.

Palavras-chave: Energia Elétrica. Modelos. Séries Temporais.

Introdução

O consumo de energia elétrica do setor industrial brasileiro tem sido investigado nos últimos anos devido principalmente pela eficiência energética, para o desenvolvimento do setor e para o planejamento do sistema elétrico [1]. O setor industrial da região sudeste foi responsável por aproximadamente 52,3% do consumo de energia elétrica em 2022 [2]. Assim, as projeções do consumo de energia elétrica são informações importantes para o planejamento de setores da indústria brasileira.

O objetivo deste trabalho é aplicar os modelos de amortecimento exponencial aos dados de consumo de energia elétrica do setor industrial na região sudeste do Brasil. Desta forma, o interesse é obter um modelo de amortecimento exponencial que melhor se ajuste aos dados e que tenha uma melhor capacidade preditiva.

Metodologia

Neste trabalho, foram utilizados os dados mensais do consumo de energia elétrica do setor industrial brasileiro na região sudeste, considerando o período de janeiro de 2004 até dezembro de 2022. Os dados foram obtidos no sítio eletrônico da Empresa de Pesquisa Energética (EPE).

O modelo de amortecimento exponencial foi obtido através da função ETS do pacote R forecast [3], o qual seleciona o melhor dos modelos de amortecimento exponencial que se ajusta aos dados. Como os dados utilizados no trabalho tem tendência e sazonalidade, o modelo que melhor se ajusta aos dados é o modelo de Holt-Winters [4], e para avaliar a capacidade preditiva do modelo de amortecimento exponencial foi comparado ao modelo ingênuo sazonal utilizando a métrica do erro médio absoluto percentual (MAPE). Os modelos foram comparados usando diferentes conjuntos de treino (ajuste) e teste (previsão). Assim, o primeiro período considerou o conjunto de treinamento entre 2004 e 2018, e o período de teste sendo os meses do ano de 2019; segundo período considera o conjunto de treinamento entre 2004 e 2019, e o conjunto de teste sendo os meses do ano de 2020; o terceiro período de treinamento entre 2004 e 2020, e o conjunto de teste sendo os meses do ano de 2021; último período (ou período 4) sendo o conjunto de treino entre 2004 e 2021, e o conjunto de teste sendo os meses do ano de 2022.

As análises estatísticas foram realizadas através do programa R [5].

Resultados

A figura 1(a) apresenta o ajuste dos modelos de amortecimento e o ingênuo sazonal aos dados de consumo de energia elétrica das indústrias da região sudeste entre 2004 e 2021. Pode-se observar nesta figura que o modelo de amortecimento exponencial conseguiu capturar a trajetória do consumo de energia elétrica durante a crise econômica de 2009 e a pandemia do COVID-19 em 2020. A figura 1(b) mostra a previsão mensal para o ano de 2022.

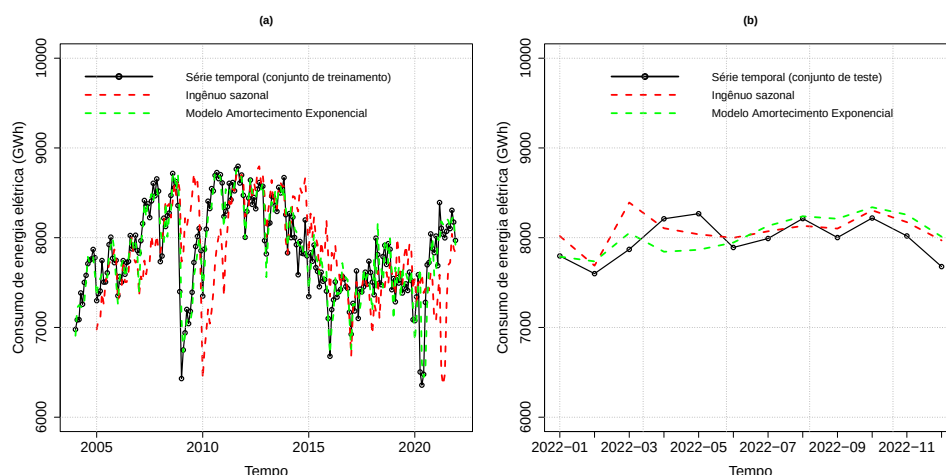


Figura 1: (a) Ajuste dos modelos aos dados observados. (b) Previsão dos modelos considerados no estudo.

Fonte: Próprio autor.

A tabela 1 mostra o MAPE para cada período do conjunto de treino e teste, e a média entre os períodos considerados. Os resultados apresentados nesta tabela indicam que o modelo de amortecimento exponencial tem uma melhor capacidade preditiva, com um MAPE médio inferior a 4%.

Tabela 1: Comparação entre os modelos de amortecimento exponencial e o ingênuo sazonal.

Período	Métricas	Período 1	Período 2	Período 3	Período 4	Média
Treino	MAPE(AE*)	1.629	1.703	1.924	1.919	1.794
	MAPE(Ingênuo)	4.625	4.551	4.708	4.989	4.718
Teste	MAPE(AE*)	2.576	7.083	2.692	2.308	3.665
	MAPE(Ingênuo)	3.514	7.096	9.487	2.173	5.567

* AE - Amortecimento Exponencial.

Fonte: Próprio autor.

Referências

- [1] F. L. C. Silva, R. C. Souza, F. L. Cyrino Oliveira, P. M. Lourenco e R. F. Calili. “A bottom-up methodology for long term electricity consumption forecasting of an industrial sector - Application to pulp and paper sector in Brazil”. Em: **Energy** 144 (2018), pp. 1107–1118. DOI: 10.1016/j.energy.2017.12.078.
- [2] EPE. **Empresa de Pesquisa Energética**. Online. Acessado em 30/03/2023, <https://www.epe.gov.br/pt>.
- [3] Rob J Hyndman e Yeasmin Khandakar. “Automatic time series forecasting: the forecast package for R”. Em: **Journal of Statistical Software** 26.3 (2008), pp. 1–22. DOI: 10.18637/jss.v027.i03.
- [4] Rob J Hyndman e George Athanasopoulos. “Forecasting: principles and practice”. Em: OTexts, 2018. Cap. 7.3–7.4.
- [5] R Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing. Vienna, Austria, 2021. URL: <https://www.R-project.org/>.

Consistência do Algoritmo de Aprendizado Supervisionado: Estudando as Cotas de Generalização do Perceptron

Yeonatan Mauhnoom
ICEX-UFF Volta Redonda
yeonatanmauhnoom@id.uff.br

Marina S. Dias de Freitas
ICEX-UFF Volta Redonda
msdias@id.uff.br

Alan P. de Paula
ICEX-UFF Volta Redonda
alanprata@id.uff.br

Resumo: Neste trabalho estamos interessados no aprendizado supervisionado e na fundamentação desse método através do princípio da minimização do risco empírico (MRE). Este princípio será utilizado como ferramenta para garantir a consistência de um algoritmo de aprendizado sob condições específicas. Por fim, vamos aplicar esses conceitos fundamentalmente teóricos no caso prático de um algoritmo de classificação do tipo perceptron e, assim, estudar como sua complexidade pode determinar sua capacidade de aprendizado e eficiência.

Palavras-chave: Aprendizado Supervisionado. Perceptron. Dimensão VC. Generalização.

Introdução

No estudo de aprendizado supervisionado denotamos por $\mathbb{X}$, $\mathbb{Y}$, e $P_{X,Y}$ respectivamente, o espaço de entrada, o espaço de saída e uma distribuição de probabilidade conjunta em $\mathbb{X} \times \mathbb{Y}$.

Seja $\mathcal{D}_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$ uma amostra aleatória independente e identicamente distribuída (i.i.d.) de $P_{X,Y}$ e (X, Y) um par de variáveis aleatórias distribuídas de acordo com $P_{X,Y}$.

Considerando que $P_{X,Y}$ deve ser fixa e desconhecida, o objetivo do algoritmo de aprendizado é selecionar, dentro de uma classe de funções, chamada de classe de hipóteses $\mathcal{H} \subset \mathcal{F}(\mathbb{X}, \mathbb{Y})^1$, o melhor classificador $h : \mathbb{X} \rightarrow \mathbb{Y}$ tal que $h(X)$ seja uma boa aproximação de Y (SHALEV-SHWARTZ e BEN-DAVID, 2014).

Para medir o desempenho de um preditor $h : \mathbb{X} \rightarrow \mathbb{Y}$ precisamos da definição de função de perda, que é dada por uma função $\ell : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}^+$, tal que $\ell(y_i, h(x_i))$ mede o quanto $h(x_i)$ dista de y_i para cada par $(x_i, y_i) \in X \times Y$.

Princípio de minimização do risco empírico

O risco verdadeiro representa em média o quanto uma função $h \in \mathcal{H}$ erra com relação a Y , pela função de perda ℓ , e é definido por

$$R(h) := \mathbb{E}_{X,Y}[\ell(Y, h(X))]. \quad (1)$$

Definimos o melhor classificador da classe e denotamos por h^* a função $h \in \mathcal{H}$ que minimiza o risco verdadeiro, ou seja,

$$h^* = \arg \min_{h \in \mathcal{H}} R(h) = \arg \min_{h \in \mathcal{H}} \mathbb{E}[\ell(Y, h(X))]. \quad (2)$$

O risco empírico é dado por $\hat{R}_n(h) = \frac{1}{n} \sum_{i=1}^n \ell(y_i, h(x_i))$. Assim, o princípio de minimização do risco empírico (MRE) consiste em escolher a função que minimiza o risco empírico, denotada por $\hat{h}_n$, isto é,

¹Definimos $\mathcal{F}(\mathbb{X}, \mathbb{Y})$ como o espaço de todas as funções $f : \mathbb{X} \rightarrow \mathbb{Y}$

$$\hat{h}_n = \arg \min_{h \in \mathcal{H}} \hat{R}_n(h) = \arg \min_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n \ell(y_i, h(x_i)). \quad (3)$$

Dado que o risco verdadeiro é desconhecido, precisamos achar uma forma de garantir que a função $\hat{h}_n$ é de fato uma boa aproximação do melhor classificador da classe h^* . Portanto, definimos o conceito de generalização por uma função $G(h)$ que quantifica a distância entre o risco empírico de uma função $h \in \mathcal{H}$ para o seu risco verdadeiro.

Mostraremos que, sob certas condições sobre $\mathcal{H}$, a probabilidade de se ter uma má generalização (com $G(h) = |\hat{R}_n(h) - R(h)| > \varepsilon$) está cotada por um valor δ tão pequeno quanto desejarmos, dada uma amostra de dados $\mathcal{D}_n$ i.i.d suficientemente grande. Diremos que o princípio de MRE é consistente com relação a classe $\mathcal{H}$ se

$$P(|R(\hat{h}_n) - R(h^*)| > \varepsilon) < \delta, \quad (4)$$

para todo $n > n_0$.

Diz-se que uma amostra $\mathcal{D}_n$ é fragmentada por $\mathcal{H}$, caso as funções de classificação em $\mathcal{H}$ produzam todas as possíveis classificações distintas em $\mathcal{D}_n$. Quando isso ocorre, dizemos que o coeficiente de fragmentação, denotado por $\mathcal{N}(\mathcal{H}, \mathcal{D}_n)$, tem um crescimento exponencial, ver (PONTI e DE MELO, 2018).

Definimos a dimensão Vapnik-Chervonenkis (ou dimensão VC) como o tamanho máximo, isto é número máximo de elementos n' tal que existe uma amostra $\mathcal{D}_{n'}$ fragmentável por $\mathcal{H}$. Veremos que na equação (4) acima $n_0 = n'$.

Da cota de Chernoff, conseguimos estabelecer a seguinte cota em função do coeficiente de fragmentação para a probabilidade do pior caso de generalização para classes de hipóteses com infinitas funções.

$$P\left(\sup_{h \in \mathcal{H}} |\hat{R}_n(h) - R(h)| > \varepsilon\right) \leq 2\mathcal{N}(\mathcal{H}, 2n) \exp\left(-n \frac{\varepsilon^2}{4}\right). \quad (5)$$

Afirmamos que o pior caso de generalização é de fato cota superior para a equação (4), ou seja, podemos substituir δ pelo lado direito da desigualdade na equação (5) acima.

Ao estudar o algoritmo de classificação binária perceptron, no qual a dimensão do espaço dos dados de entrada é $d = 2$. Pode-se mostrar que a dimensão VC de qualquer classe de hipóteses $\mathcal{H}$, cujas funções $h \in \mathcal{H}$ são hiper-planos lineares $d - 1$ dimensionais, é igual a $d + 1$ (PONTI e DE MELO, 2018, p.120).

Outro resultado importante obtido consiste em uma maneira de se calcular o coeficiente de fragmentação do perceptron, através da seguinte expressão

$$\mathcal{N}(\mathcal{H}, 2n) = 2^n - 2 \sum_{i=d+1}^n \binom{n-1}{i}. \quad (6)$$

Em particular, para o perceptron considerando $d = 2$ temos $\mathcal{N}(\mathcal{H}, 2n) = n^2 - n + 2$.

Conclusões

Concluimos assim que para garantir a consistência de um algoritmo é preciso avaliar sua complexidade, a partir do estudo de sua dimensão VC e do coeficiente de fragmentação. De modo que podemos fixar a precisão, dada por ε , e a taxa de assertividade, dada por δ , conforme desejarmos e descobrir a quantidade mínima de dados necessários para viabilizar o aprendizado.

Referências

SHALEV-SHWARTZ S.; BEN-DAVID S. **Understanding machine learning: From theory to algorithms**. Cambridge university press, 2014.

PONTI M. A.; DE MELOR. F. **Machine Learning: a Practical Approach on the statistical learning theory**. Springer International Publishing AG, 2018.

Estudo Sobre Modelos de Redes Neurais Artificiais Aplicados aos Dados de Consumo de Energia Elétrica no Setor Industrial do Estado de São Paulo

Daniel Lourenço dos Santos

Universidade Federal Rural do Rio de Janeiro

dlourenco@ufrj.br

Felipe Leite Coelho da Silva

Universidade Federal Rural do Rio de Janeiro

felipeleite@ufrj.br

Resumo: O consumo de energia elétrica no setor industrial brasileiro vem sendo investigado ao longo dos últimos anos. Esse interesse está relacionado com o desenvolvimento do setor, políticas de planejamento energético e eficiência energética. Desta forma, os modelos de previsões são importantes para tomadas de decisões. O objetivo deste trabalho é comparar diferentes modelos de previsão aplicado aos dados mensais do consumo de energia elétrica do setor industrial no estado de São Paulo. Foram utilizados os modelos de Redes Neurais Autorregressivas (NNAR), Perceptron Multicamadas (MLP) e o modelo SARIMA. Através das Métricas de erro utilizadas, o modelo SARIMA apresentou melhor resultado no período analisado.

Palavras-chave: Séries Temporais. Redes Neurais. Setor Industrial.

Introdução

Nos últimos anos têm-se investigado as projeções do consumo de energia elétrica em setores da indústria do Brasil (SILVA et al., 2018, p. 1107). Esse interesse está relacionado com o desenvolvimento do setor, políticas de planejamento e eficiência energética. Assim, os modelos de previsão podem contribuir para análise do consumo de energia elétrica.

O objetivo deste trabalho é avaliar a capacidade preditiva dos modelos NNAR (HYNDMAN e ATHANASOPOULOS, 2018) e MLP (HAYKIN, 2009) aplicados aos dados mensais do consumo de energia elétrica do setor industrial no estado de São Paulo.

Para avaliar o desempenho de previsão dos modelos NNAR e MLP comparou-se os resultados com modelo SARIMA (BOX et al., 2018).

Metodologia

Neste trabalho foram utilizados os dados mensais do consumo de energia elétrica do setor industrial no estado de São Paulo do ano de 2004 até o ano de 2022, disponíveis no sítio eletrônico da Empresa de Pesquisa Energética (EPE, 2023). A Tabela 1 mostra como foram divididos os dados para ajustes, testes e avaliação dos modelos.

Tabela 1: Divisão dos dados entre ajuste e teste.

	Ajuste	Teste
Período 1	2004 até 2019	2020
Período 2	2004 até 2020	2021
Período 3	2004 até 2021	2022

Para avaliarmos a capacidade preditiva dos modelos NNAR, MLP e SARIMA utilizou-se a raiz do erro quadrático médio (RMSE) e o erro médio absoluto percentual (MAPE).

Resultados e discussão

Neste trabalho, foram ajustados os modelos MLP com 12 nós de entrada e 2 camadas ocultas com 10 nós cada, o modelo $NNAR(1,1,2)_{12}$ e o modelo $SARIMA(1,0,1) \times (1,1,1)_{12}$ nos 3 períodos de ajustes e testes.

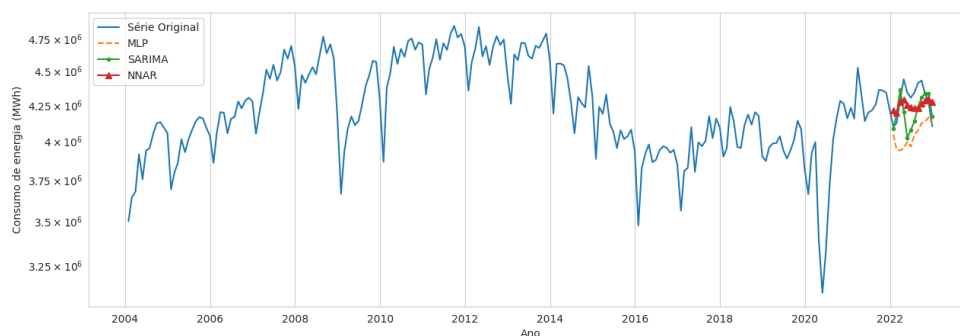


Figura 1: Série original seguida das previsões dos modelos.

Tabela 2: Comparação entre a capacidade preditiva dos modelos MLP, $NNAR(1,1,2)_{12}$ e $SARIMA(1,0,1) \times (1,1,1)_{12}$ através das métricas MAPE e RMSE.

Métrica	Modelo	Período 1	Período 2	Período 3	Média
MAPE	MLP	8,80%	7,43%	7,14%	7,8%
	NNAR	8,51%	10,1%	2,31%	6,98%
	SARIMA	7,71%	3,73%	2,97%	4,81%
RMSE	MLP	499.907,77	334.282,55	318.491,65	384.227,33
	NNAR	406.023,81	458.765,33	112.831,47	325.873,54
	SARIMA	355.713,29	206.438,21	159.642,23	240.597,91

Pelos resultados apresentados na Tabela 2, verificou-se que o modelo SARIMA apresentou uma melhor capacidade preditiva em relação aos modelos MLP e NNAR.

Referências

- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. 4.ed. New Jersey: John Wiley & Sons, Inc., 2008.
- EPE. Consumo Mensal de Energia Elétrica por Classe (regiões e subsistemas). 2023. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/consumo-de-energia-eletrica>. Acesso em: 5 sep. 2023.
- HAYKIN, S. S. **Neural Networks and Learning Machines**. 3.ed. New Jersey: Prentice Hall, 2009.
- HYNDMAN R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 2.ed. Melbourne: OTexts, 2018.
- SILVA, F. et al. A bottom-up methodology for long term electricity consumption forecasting of an industrial sector - Application to pulp and paper sector in Brazil. **Energy**. v. 144. p. 1107-1118, 2018.

Geração de Variáveis Aleatórias

Daniela Sano Adathi

Universidade Federal Fluminense (UFF)

danielasano@id.uff.br

Marina Sequeiros Dias de Freitas

Universidade Federal Fluminense (UFF)

msdias@id.uff.br

Alan Prata de Paula

Universidade Federal Fluminense (UFF)

alanprata@id.uff.br

Resumo: Neste trabalho primeiramente estudamos como um computador pode ser utilizado para gerar números aleatórios, mais precisamente, números pseudoaleatórios. Em seguida, estudamos como esses números podem ser utilizados para gerar os valores de variáveis aleatórias de distribuições arbitrárias. Esse conhecimento permite realizar simulações de modelos estocásticos ao longo tempo, tendo portanto inúmeras aplicações.

Palavras-chave: Geração de Variáveis Aleatórias. Processos Estocásticos. Simulação.

Introdução

Para o estudo de simulação uma das tarefas necessárias é a habilidade de gerar números aleatórios. O objetivo deste trabalho é o estudo de geração de variáveis aleatórias discretas e contínuas e, por esse motivo, primeiro precisamos entender o conceito de números aleatórios e pseudoaleatórios.

O conhecimento do processo de geração de variáveis aleatórias é um dos primeiros passos para simular modelos probabilísticos. Algumas aplicações podem ser feitas a partir da aproximação desses modelos por simulação de eventos discretos e incluem simulação de filas com um único servidor ou vários servidores em série ou em paralelo, controle de estoque de lojas, análise de risco de uma empresa de seguros, dentre outros, conforme Paula, 2014.

Números Aleatórios e Pseudoaleatórios

Números aleatórios representam o valor de uma variável aleatória com distribuição uniforme $[0,1]$, isto é, um número gerado conforme a função densidade $f(x) = 1$ se $x \in [0, 1]$ e $f(x) = 0$ caso contrário.

Já os números pseudoaleatórios formam uma sequência de valores que tentam simular uma sequência de realizações independentes de uma distribuição uniforme $[0,1]$. São gerados por algoritmos determinísticos, como o Método Congruente Multiplicativo, a partir de uma semente e seguem um padrão previsível. Apesar dos números gerados por algoritmos pseudoaleatórios parecerem aleatórios em curtos períodos, eles, em algum momento, irão repetir o mesmo ciclo de valores, conforme Ross (2012).

É importante notar que a partir de uma variável aleatória U com distribuição uniforme $[0,1]$ pode-se simular o valor de qualquer variável aleatória X , discreta ou contínua, como será feito na próxima seção.

Método da transformada inversa

Para simular o valor de uma variável aleatória discreta X , com função de probabilidade definida por

$$\mathbb{P}(X = x_j) = p_j, \quad j \geq 0 \text{ e } \sum_j p_j = 1,$$

gere um número aleatório U e considere:

$$X = \begin{cases} x_0, & \text{se } U < p_0 \\ x_1, & \text{se } p_0 \leq U < p_0 + p_1 \\ \vdots & \\ x_j, & \text{se } \sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^j p_i. \end{cases}$$

Como U tem distribuição uniforme $[0,1]$, temos $\mathbb{P}(a \leq U < b) = b - a$, com $0 < a < b < 1$. Logo, obtemos $\mathbb{P}(X = x_j) = \mathbb{P}\left(\sum_{i=0}^{j-1} p_i \leq U < \sum_{i=0}^j p_i\right) = p_j$. Ou seja, X terá distribuição desejada.

Por outro lado, para simular uma variável aleatória contínua com função de distribuição acumulada F , utilizamos o resultado seguinte.

Teorema 1 *Seja U com distribuição uniforme $[0,1]$. Para qualquer função de distribuição acumulada F de uma variável aleatória contínua, temos que X definida por $X = F^{-1}(U)$ tem distribuição F .*

Técnica da aceitação-rejeição

Se conhecemos o processo de geração de uma variável aleatória com função de probabilidade definida por $\{q_j, j \geq 0\}$, podemos utilizá-la como base para simular uma variável aleatória com função de probabilidade $\{p_j, j \geq 0\}$.

Considere c uma constante tal que $\frac{p_j}{q_j} \leq c$ para todo $j \geq 0$, com $p_j > 0$. Para simular uma variável aleatória X com função de probabilidade p_j , simule um valor de Y com função de probabilidade q_j , em seguida gere um número aleatório U . Por fim, se $U < \frac{p_Y}{cq_Y}$, faça $X = Y$. Caso contrário, retorne ao primeiro passo. Essa técnica é justificada pelo seguinte teorema.

Teorema 2 *O algoritmo da aceitação-rejeição gera uma variável aleatória X com função de probabilidade definida por $\mathbb{P}(X = x_j) = p_j$, $j \geq 0$, a partir de uma variável aleatória com função de probabilidade dada por $\{q_j, j \geq 0\}$. Além disso, o número de iterações do algoritmo necessárias para obter X é uma variável aleatória geométrica com média c .*

Analogamente, o mesmo processo pode ser utilizado para gerar variáveis aleatórias contínuas.

Conclusões

Nesse trabalho mostramos como podemos gerar números aleatórios (mais especificamente, números pseudoaleatórios) e como, utilizando-os, podemos simular variáveis aleatórias discretas e contínuas. Em particular, podemos gerar uma variável aleatória geométrica, de Poisson, exponencial ou uma sequência de variáveis aleatórias Bernoulli. Com isso, podemos simular diversos modelos probabilísticos.

Referências

- PAULA, R. R. **Método de Monte Carlo e Aplicações**. 2014. 81f. Trabalho de Conclusão de Curso – Instituto de Ciências Exatas, UFF, Volta Redonda, 2014.
- ROSS, S.M. **Simulation**. 5.ed. Berkeley: Elsevier Academic Press, 2012.

MODELOS DE PREVISÃO DO CONSUMO DE ENERGIA DO SETOR INDUSTRIAL DE ALIMENTOS E BEBIDAS DO BRASIL

Lucas Camaz Ferreira

*UFRRJ - Universidade Federal Rural do Rio de Janeiro
camazlucas@gmail.com*

Felipe Leite Coelho da Silva

*UFRRJ - Universidade Federal Rural do Rio de Janeiro
felipeleite@ufrj.br*

Resumo: O estudo sobre o consumo de energia elétrica é essencial para o planejamento energético do país, neste contexto, o setor industrial detém a liderança em demanda de energia no Brasil e dentro deste setor a indústria de alimentos e bebidas IAB é a segunda maior consumidora de energia. Neste trabalho são estudados três modelos de previsão do consumo de energia utilizando séries temporais: O modelo SARIMA, o de Suavização Exponencial e o modelo de Redes Neurais Autorregressivas utilizando métricas de erro para avaliar a acurácia de cada modelo.

Palavras-chave: Previsão. Séries Temporais. Consumo de Energia

Introdução

O consumo de energia elétrica do setor industrial brasileiro vem sendo investigado devido a sua importância no desenvolvimento nacional em termos sociais e econômicos. Em 2021, a indústria de alimentos e bebidas (IAB) foi responsável por 13,1% do consumo de energia elétrica do setor industrial (EPE, 2022). Assim, um estudo sobre os modelos de séries temporais é importante para investigar e prever a demanda de energia elétrica, auxiliando no planejamento energético nacional.

Neste trabalho, serão utilizados os modelos de amortecimento exponencial, o modelo de Box e Jenkins e o de Redes Neurais Autoregressivas para prever o consumo de energia elétrica da IAB.

Metodologia

Os dados de consumo de energia elétrica foram obtidos em sítio eletrônico da empresa de pesquisa energética (EPE), para o período de janeiro de 2015 até maio de 2022. Esses dados foram divididos em dois conjuntos, um de treino e outro de teste. Para o conjunto treino foram utilizados os dados de janeiro de 2015 até maio de 2021 e para o conjunto de teste de junho de 2021 até maio de 2022.

Para o modelo de previsão de suavização exponencial (SE) foi utilizado o método aditivo. Esse modelo é utilizado quando as variações sazonais são aproximadamente constantes ao longo da série. Nesse método, a componente sazonal é representada em termos absolutos na escala das séries observadas e na equação de nível a série é ajustada sazonalmente subtraindo a componente sazonal (MORETTIN; TOLOI, 2018).

Para o modelo de Box & Jenkins foi utilizado o modelo ARIMA Sazonal (SARIMA) para incorporar a componente de sazonalidade que está presente nos dados, este modelo possui a estrutura $SARIMA(p, d, q) \times (P, D, Q)_s$, com os parâmetros p, d e q indicando respectivamente as componentes autoregressivas, diferenciação e de médias móveis, enquanto que os parâmetros P, D e Q tem a mesma representação porém na ordem sazonal, indicada por s .

Ja o modelo de Redes Neurais Autoregressivas (NNAR) busca modelar a relação entre um conjunto de sinais de entrada e um sinal de saída. Sua estrutura $NNAR(p, P, k)_s$ tem os parâmetros p e P equivalentes aos passos de uma autoregressão, logo, comparando com o modelo SARIMA seria como

um $SARIMA(p, 0, 0) \times (P, 0, 0)_s$, com s também indicando a sazonalidade e o parametro k indicando o número de neurônios na camada oculta.

Para avaliar a capacidade preditiva dos modelos aplicados foi utilizada a métrica do Erro Médio Absoluto Percentual (MAPE).

Resultados e Conclusões

A Figura 1(a) mostra o ajuste do modelo de suavização exponencial(SE), do modelo SARIMA e do modelo de Redes Neurais Autoregressivas (NNAR) aos dados de consumo de energia elétrica da indústria de alimentos e bebidas no Brasil. A previsão do consumo de energia elétrica pode ser observada na Figura 1(b).

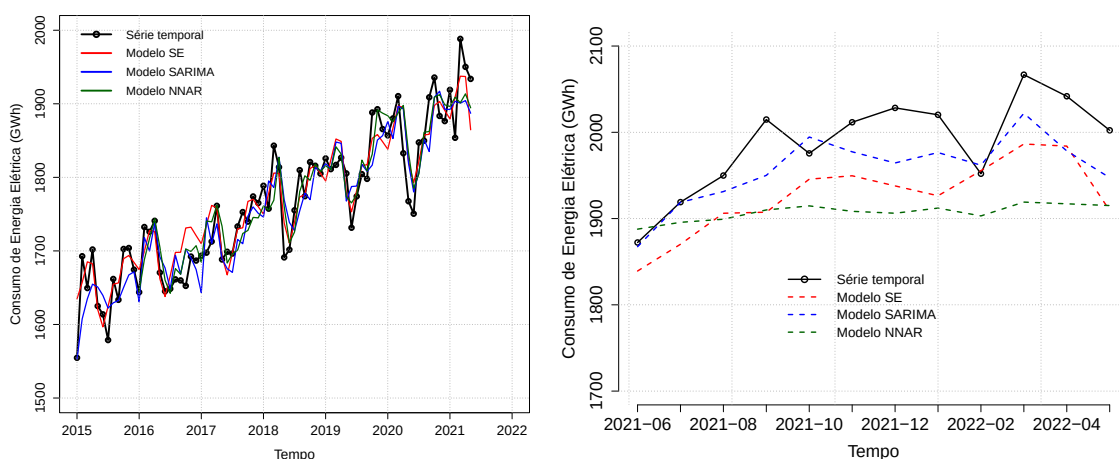


Figura 1: Comparativo entre os tres modelos de series temporais

Pela Tabela 1 podemos observar que o modelo NNAR forneceu um melhor ajuste ao conjunto de treino porém demonstrou a pior capacidade preditiva dentre os três, enquanto que o modelo SARIMA apesar de apresentar um ajuste pior demonstrou melhor capacidade preditiva.

Tabela 1: Métricas para o conjunto de treino e teste.

Modelo	SE	SARIMA	NNAR
Treino	1.43	1.63	1.41
Teste	3.10	1.75	4.13

Agradecimentos

Ao CNPq (407972/2021-8) e a FAPERJ (269297) pelo suporte financeiro.

Referências

- EPE. Empresa de Pesquisa Energética. 2022. Disponível em: <https://www.epe.gov.br/pt>. Acesso em: 27 mar. 2023.
- HYNDMAN R.J.; ATHANASOPOULOS, G. (2018) Forecasting: principles and practice, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.
- MORETTIN, P.A.; TOLOI, C.M.C. **Análise de Séries Temporais**. São Paulo, SP: Volume 1, 3ª edição. Blucher, 2018.

Uma análise exploratória para os dados de consumo de energia elétrica do setor industrial para os estados brasileiros

Guilherme Moreira da Fraga Filho
Universidade Federal Rural do Rio de Janeiro
guilhermef@ufrj.br

Felipe Leite Coelho da Silva
Universidade Federal Rural do Rio de Janeiro
felipeleite@ufrj.br

Palavras-chave: Energia elétrica. Setor industrial.

O setor industrial brasileiro tem uma importância para o crescimento econômico de uma região ou estado. Um dos fatores que influenciam a instalação de um segmento da indústria em uma região ou estado é o abastecimento de energia elétrica. Neste contexto, a análise estatística do consumo de energia elétrica por estado brasileiro é importante para o planejamento energético e para o desenvolvimento do setor.

Os dados de consumo de energia elétrica (MWh) do setor industrial utilizados neste trabalho foram obtidos através do site Empresa de Pesquisa Energética (EPE). Foram utilizados os dados mensais de consumo de energia elétrica por estado brasileiro, considerando o período de 2013 a 2022. Primeiramente, foram obtidas as médias anuais e o desvio padrão do consumo por estado e, seguida, uma classificação para os estados. Além disso, será utilizado os algoritmos de agrupamentos hierárquicos.

Estados	Média 2013 - 2022	Estados	Média 2013 - 2022	Estados	Média 2013 - 2022
Rondônia	39582.18	Ceará	206382.63	Rio de Janeiro	707165.43
Acre	3411.15	Rio Grande do Norte	104075.37	São Paulo	4131961.64
Amazonas	142640.60	Paraíba	120064.20	Paraná	1057438.82
Roraima	1931.38	Pernambuco	309260.62	Santa Catarina	860707.78
Pará	1005807.13	Alagoas	137625.96	Rio Grande do Sul	832127.91
Amapá	5842.21	Sergipe	109487.88	Mato Grosso do Sul	109487.63
Tocantins	27333.72	Bahia	786377.59	Mato Grosso	179937.96
Maranhão	187619.48	Minas Gerais	2529911.40	Goiás	454813.23
Piauí	19307.85	Espírito Santo	356919.52	Distrito Federal	52865.04

Tabela 1: Média do consumo de energia elétrica dos estados brasileiros para o período de 2013 a 2022.

Pela Tabela 1 pode-se observar que o estado de São Paulo é o maior consumidor de energia elétrica no setor industrial. A Tabela 2 mostra o desvio padrão do consumo de energia elétrica entre os estados brasileiros.

Estados	Desvio Padrão	Estados	Desvio Padrão	Estados	Desvio Padrão
Rondônia	4826.88	Ceará	13031.93	Rio de Janeiro	32812.30
Acre	592.51	Rio Grande do Norte	3759.54	São Paulo	250027.75
Amazonas	10021.50	Paraíba	2812.77	Paraná	71431.96
Roraima	246.83	Pernambuco	13770.63	Santa Catarina	63492.43
Pará	72948.82	Alagoas	41338.62	Rio Grande do Sul	34823.61
Amapá	2437.88	Sergipe	17485.81	Mato Grosso do Sul	14136.69
Tocantins	3671.71	Bahia	46520.06	Mato Grosso	12309.18
Maranhão	131740.01	Minas Gerais	141323.30	Goiás	34335.13
Piauí	2792.28	Espírito Santo	42862.70	Distrito Federal	8900.26

Tabela 2: Desvio padrão dos anos de 2013 a 2022 em relação a cada estado.

A Figura 1 mostra a média consumo de energia elétrica do setor industrial por estado brasileiro para

o ano de 2020, sendo estado de São Paulo o maior consumidor de energia elétrica e o estado de Roraima com o menor consumo.

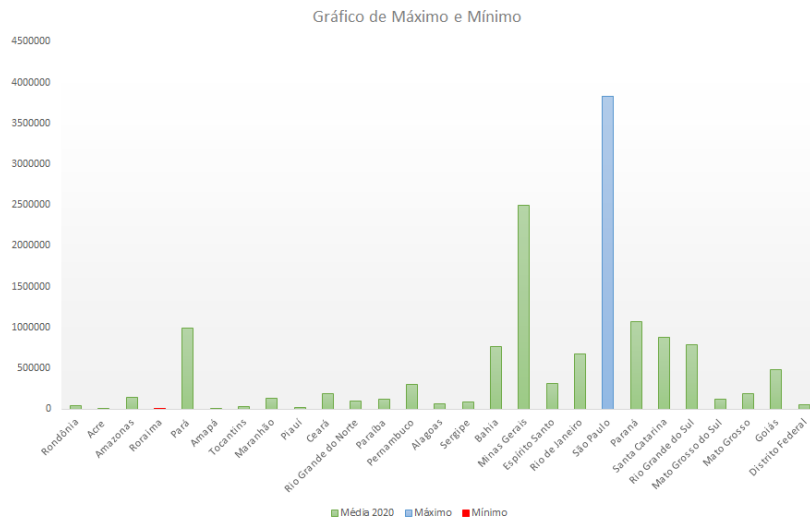


Figura 1: Consumo médio de energia elétrica do o ano de 2020.

Neste estudo, será utilizado os métodos de agrupamentos hierárquicos para classificar os estados brasileiros em relação ao consumo de energia elétrica. A análise do consumo de energia elétrica nas indústrias em cada estado brasileiro é importante para o planejamento energético do estado e por região.

Referências

- BUSSAB, W. O., MORETTIN, P. A.. Estatística básica. 5. ed.. Saraiva. 2002
- Manly, Bryan J. F.; ALBERTO, Jorge A. N. Multivariate Statistical Methods: A Primer. 4ed., CRC Press, Includes bibliographical references and index. 2017.

APRENDIZADO DE MÁQUINAS: PROBLEMAS DE SUBAJUSTE E SUPERAJUSTE

Matheus Moreira Nery
Universidade Federal Fluminense
matheusnery@id.uff.br

Marina Sequeiros Dias de Freitas
Universidade Federal Fluminense
msdias@id.uff.br

Alan Prata de Paula
Universidade Federal Fluminense
alanprata@id.uff.br

Resumo: O aprendizado de máquina, também conhecido como machine learning, é uma subárea da inteligência artificial que busca desenvolver algoritmos e técnicas que permitam que os computadores aprendam e melhorem seu desempenho em tarefas específicas sem serem explicitamente programados para fazê-lo. O aprendizado de máquinas tem como uma das suas principais subáreas o aprendizado supervisionado. Neste trabalho iremos apresentar dois problemas que encontramos nos algoritmos deste tipo de aprendizado, o subajuste e o superajuste, e posteriormente entenderemos alguns fatores que podem evitar esse problema, melhorando assim o desempenho do aprendizado.

Palavras-chave: Aprendizado de Máquinas. Aprendizado Supervisionado. Subajuste. Superajuste.

Introdução

No aprendizado supervisionado o algoritmo é treinado com um conjunto de dados rotulados, ou seja, dados onde as respostas corretas são conhecidas. Usamos $\mathbb{X}$ e $\mathbb{Y}$ para denotar o espaço de entrada e o espaço de saída e $P_{\mathbb{X},\mathbb{Y}}$ representa uma função de distribuição de probabilidade em $\mathbb{X} \times \mathbb{Y}$. Seja $(\mathbb{X}, \mathbb{Y})$ um par de variáveis aleatórias distribuídas de acordo com distribuição de probabilidade $P_{\mathbb{X},\mathbb{Y}}$ e $D_n = \{(x_1, y_1), \dots, (x_n, y_n)\}$ uma amostra aleatória independente e identicamente distribuída (i.i.d) de $P_{\mathbb{X},\mathbb{Y}}$. O objetivo do aprendizado supervisionado é encontrar um mapeamento $h : \mathbb{X} \rightarrow \mathbb{Y}$ baseado em D_n de modo que $h(\mathbb{X})$ seja uma boa aproximação de $\mathbb{Y}$.

A ideia do aprendizado é a partir do conjunto de treinamento D_n produzir uma boa predição h . Normalmente, o algoritmo de aprendizado se resume em um espaço de funções $\mathbb{H}$ que chamamos de espaço de hipóteses, contido no espaço de todas as funções que levam de $\mathbb{X}$ em $\mathbb{Y}$ denotado por $\mathcal{F}(\mathbb{X}, \mathbb{Y})$.

Uma função de perda é um mapeamento $\ell : \mathbb{Y} \times \mathbb{Y} \rightarrow \mathbb{R}^+$ onde ℓ é escolhido de acordo com o tipo de aprendizado. Na classificação por exemplo, a função de perda geralmente escolhida é $\ell(y, p) = \mathbb{I}(y \neq p)$, já em uma regressão é o erro quadrático $\ell(y, p) = (y - p)^2$. Para podermos mensurar a qualidade de um algoritmo de aprendizado, usaremos o **risco esperado** e o **risco empírico**, onde o risco esperado é definido utilizando o valor esperado da função de perda, dada por

$$R(h) = \mathbb{E}(\ell(\mathbb{X}, \mathbb{Y}, h(\mathbb{X}))) \quad (1)$$

e o risco empírico é calculado a partir de uma amostra aleatória D_n , onde aplicamos a função de perda em cada observação e calculamos a média, da seguinte forma:

$$R_{emp}(h) = \frac{1}{n} \sum_{i=1}^n \ell(x_i, y_i, h(x_i)). \quad (2)$$

Além disso, para relacionarmos esses dois erros, usamos o conceito de generalização

$$G = |R_{emp}(h) - R(h)|. \quad (3)$$

1 Erro Aproximado e Erro Estimado

Seja h_i o algoritmo de aprendizagem chamado de classificador. Ao escolher um classificador, a principal consideração deve ser a minimização do risco associado à classificação incorreta. O melhor classificador é o que apresenta o menor risco. Denotamos por $\hat{h}$ o melhor classificador em $\mathbb{H}$, e h^* o melhor classificador no espaço de todas as funções $\mathbb{X}$ em $\mathbb{Y}$.

Sabendo disso podemos definir algumas medidas de erros. O erro estimado representa a diferença entre nossa solução h_i e o melhor classificador $\hat{h} \in \mathbb{H}$. O erro aproximado representa a distância de $\hat{h}$ ao melhor classificador h^* dado todo o espaço de funções $\mathcal{F}(\mathbb{X}, \mathbb{Y})$, e o erro total é decomposto na soma desses dois erros, dado por

$$R(h_i) - R(h^*) = \underbrace{(R(\hat{h}) - R(h^*))}_{\text{erro-aproximado}} + \underbrace{(R(h_i) - R(\hat{h}))}_{\text{erro-estimado}}. \quad (4)$$

2 Superajuste e Subajuste

O superajuste ocorre quando um modelo de aprendizado de máquina se ajusta excessivamente aos dados de treinamento, capturando ruídos e detalhes irrelevantes, resultando em baixo desempenho em novos dados. Isso pode levar a uma falta de generalização, isto é, o modelo representa bem o conjunto D_n , entretanto, em dados não vistos tem um grande erro. Para uma classe de hipóteses $\mathbb{H}$ grande, o erro estimado é grande, mas o erro aproximado é pequeno.

O subajuste, por outro lado, acontece quando um modelo é muito simples para capturar a complexidade dos dados de treinamento, resultando em um desempenho insatisfatório tanto nos dados de treinamento quanto nos novos dados, levando também a um baixo poder de generalização. Para uma classe de hipóteses $\mathbb{H}$ pequena, o erro estimado é pequeno, mas o erro aproximado é grande.

O equilíbrio entre esses extremos é crucial para desenvolver modelos que generalizem bem para dados não vistos, sendo importante escolher estruturas de modelos apropriadas, ajustar parâmetros e utilizar conjuntos de dados adequados. Neste trabalho estudamos um exemplo de regressão polinomial, ajustando alguns polinômios com o intuito de entendermos na prática os problemas citados acima através da análise dos erros.

Conclusões

Neste trabalho, fazemos uma introdução ao aprendizado supervisionado, definindo seus principais conceitos, a decomposição do erro total e apresentamos os problemas de superajuste e subajuste.

Por fim, analisamos como minimizar a ocorrência desses problemas regulando a balança entre erro aproximado e erro estimado, vistos na decomposição do erro total, bem como isso implica diretamente no tamanho da nossa classe de hipóteses $\mathbb{H}$.

Referências

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding Machine Learning - From Theory to Algorithms**. Reino Unido: Cambridge University Press, 2014.

ZBICKI, R.; SANTOS, T. M. dos. *Aprendizado de máquina: uma abordagem estatística*. 2020. Disponível em: <http://www.rizbicki.ufscar.br/ame>. Acesso em 04 set. 2023.

Estudo de Convergência de Malha do Método Numérico de NT Aplicado ao Fenômeno de Sedimentação

Vívia de Souza Marins –
Universidade Federal Rural do Rio de Janeiro
vivasouza.dsm@gmail.com

Renan de Souza Teixeira –
Universidade Federal Rural do Rio de Janeiro
rsteixeira@ufrj.br

Resumo: Os estudos quanto a sedimentação são importantes e ainda bem explorados por contar com diferentes situações e aplicações. No seguinte trabalho mostraremos os diferentes resultados obtidos no refinamento de malha para o tempo de sedimentação, inicialmente estudando as características do fluido BR-Mul, manuseado em poços de petróleo, utilizando o método desenvolvido por Nessyahu-Tadmor e readaptado por Burguer para o problema aqui discutido. Os resultados obtidos foram também comparados a outro trabalho.

Palavras-chave: Método Numérico. Sedimentação. Nessyahu-Tadmor.

Introdução

O fenômeno de sedimentação em fluidos viscosos ocorre em diversas áreas, como citado em Silva et al. (2015), contudo, neste trabalho, focamos nos fluidos viscoelásticos utilizados na perfuração de poços de petróleo, descritos por Rocha (2018). A utilização desses fluidos nos poços tem como características notórias: (i) a retenção das paredes durante a perfuração, (ii) o transporte de cascalhos e sua suspensão durante paradas, (iii) resfriar e lubrificar a coluna de perfuração e a broca e (iv) manter o poço aberto até que o tubo de revestimento seja cimentado, anunciadas por BOURGOYNE et al. (1991). Porém, há partículas sólidas contidas nos fluidos e estes, por ação da gravidade, acabam se sedimentando no fundo, dificultando as etapas de perfuração e mantimento da abertura dos poços (NGUYEN et al., 2011).

O modelo de sedimentação é descrito por equações diferenciais não-linear e, consequentemente, a sua solução utilizando métodos implícitos eleva o custo computacional. Portanto, métodos explícitos capazes de capturar ondas de choque são importantes na solução dos problemas (BURGER et al., 2017; BURGER et al., 2000; BURGER et al., 2018; BURGER, KOZAKEVICIUS, 2007).

O objetivo deste trabalho é estudar o fenômeno utilizando diferentes métodos numéricos apropriados para a equação não-linear descrita abaixo, em especial o método NT de Nessyahu-Tadmor (1990).

$$\frac{\partial \varepsilon_s}{\partial t} + \frac{\partial (\varepsilon_s v_s)}{\partial z} = 0 \quad (1)$$

$$v_s = \left\{ \frac{K}{M(1 - \varepsilon_s)^{1-n}} \left[\frac{d_p}{\theta(\phi)} \right]^{n-1} \left(\frac{\rho_{susp}}{\rho_{susp} - \rho_s \varepsilon_s} \right) \left[\varepsilon_s (\rho_s - \rho_f) g - \frac{dP_s}{d\varepsilon_s} \frac{\partial \varepsilon_s}{\partial Z} \right] \right\}^{\frac{1}{n}} \quad (2)$$

Temos acima a equação da continuidade (1) e a equação de movimento (2) da fase sólida, válidos para o domínio $0 \leq z \leq L_0$ e $0 \leq t \leq t_f$, onde L_0 é a altura da suspensão e t_f o tempo final de sedimentação. Tendo como condição inicial e condições de contorno, respectivamente, as seguintes:

$$\varepsilon_s(z, t = 0) = \varepsilon_{s0} \quad (3a)$$

$$q_s = \varepsilon_s v_s(z = 0, t) = 0, \quad q_s = \varepsilon_s v_s(z = L_0, t) = 0 \quad (3b)$$

As equações são baseadas na teoria de sedimentação de Kynch, onde ε_s representa a concentração volumétrica, ρ_s e ρ_l são, respectivamente, a densidade do sólido e da fase líquida, v_s é a velocidade de

sedimentação e T_s tensão do sólido. Além disso temos t , z , g e m , que representam o tempo, a posição axial da coluna de sedimentação, a aceleração da gravidade e a força resistiva, respectivamente.

Além disso, temos as respectivas equações para a esfericidade da partícula ϕ , permeabilidade do ambiente e o gradiente de pressão:

$$\theta = -3,45\phi^2 + 5,25\phi - 1,41 \quad (4)$$

$$K = K_0 d_p^2 \left(\frac{\epsilon_{sm}}{\epsilon_s} - 1 \right)^\Lambda \quad (5)$$

$$\frac{dP_s}{d\epsilon_s} = \frac{P_{sref}\beta}{\epsilon_s^2} \exp \left[-\beta \left(\frac{1}{\epsilon_s} - \frac{1}{\epsilon_{sref}} \right) \right], \quad (6)$$

onde d_p é o diâmetro médio das partículas, K_0 e β são parâmetros de ajuste, P_s pressão nos sólidos e P_{sref} é a pressão nos sólidos na concentração de referência ϵ_{sref} .

Conclusões

Conforme as mudanças de malha os resultados se tornaram mais próximos aos esperados para esta situação, assim como ocorreu na observação de acordo com o tempo de simulação em cada malha.

Agradecimentos

Agradecemos a UFRRJ e a FAPERJ pela oportunidade e ajuda necessária para a realização do estudo.

Referências

- BOURGOYNE, A. et al. **Applied drilling engineering**. United States, 1991.
- BÜRGUER, R. et al. Numerical simulation of the settling of polydisperse suspensions of spheres. **Powder Technology**, v. 113, p. 30-54, 2000.
- BÜRGUER, R.; KOZAKEVICIUS, A. Adaptive multiresolution WENO schemes for multi-species kinematic flow models **Journal of Computational Physics**, v. 224, p. 1190-1222, 2007.
- BÜRGUER, R. et al. Entropy solutions of a scalar conservation law modeling sedimentation in vessels with varying cross-sectional area. **SIAM Journal on Applied Mathematics**, v. 77, p. 789-811, 2017.
- BÜRGUER, R. et al. A difference scheme for a degenerating convection-diffusion-reaction system modelling continuous sedimentation. **ESAIM: Mathematical Modelling and Numerical Analysis**, v. 52, p. 365-392, 2018.
- NESSYAHU, H. TADMOR, T. Non-oscillatory central differencing for hyperbolic conservation laws **Journal of Computational Physics**, v. 87, p. 408-463, 1990.
- NGUYEN, T. et al. Predicting dynamic barite sag in newtonian-oil based drilling fluids in pipe. **Journal of energy resources technology**, Orleans, out. 2009.
- ROCHA, R. R. **Estudo Teórico-Experimental da Sedimentação em Batelada: Monitoramento e Modelagem de Perfis de Concentração de Sólidos e Análise de Equações Constitutivas**. 2018. 116 f. Dissertação (Mestrado em Engenharia Química) – Universidade Federal Rural do Rio de Janeiro, Seropédica, 2018.
- SILVA, R. et al. Settling suspensions flow modelling: a review. **KONA Powder and Particle Journal**, v. 32, p. 41-56, fev. 2015.

MODELOS DISCRETOS PARA PESCA DE RENDIMENTO CONSTANTE E PESCA DE ESFORÇO CONSTANTE

Sarah Alves de Castro

Universidade Federal Rural do Rio de Janeiro

sarahalvesc@ufrj.br

Claudia Mazza Dias

Universidade Federal Rural do Rio de Janeiro

mazza@ufrj.br

Resumo: Considerando que a gestão racional dos recursos naturais e a proteção do meio ambiente são essenciais, esse trabalho propõe uma revisão de modelos discretos para a pesca de rendimento constante e a pesca de esforço constante, com base nos modelos clássicos da dinâmica populacional. Assim, pretende-se contribuir para o entendimento dos impactos da pesca no equilíbrio dos ecossistemas.

Palavras-chave: Modelagem Matemática, Modelos Discretos. Controle da Pesca. Dinâmica Populacional.

Introdução

A pesca exerce um papel fundamental na economia, tanto no Brasil, quanto do mundo. Além de garantir a segurança alimentar, é também garantia de emprego e renda para diversas famílias (OCEANA, 2022). Evidenciando os tipos de pesca, têm-se: a pesca tradicional, que é praticada pelo pescador autônomo desembarcado ou em pequenas embarcações, podendo ou não ter fins lucrativos; e a pesca industrial, que é realizada em grande escala. Apesar de serem economicamente indispensáveis, ambas apresentam riscos ao meio ambiente se forem feitas de forma desenfreada, ou sem a devida fiscalização. Por exemplo, na pesca tradicional, caso não haja apoio de órgãos fiscalizadores, essa atividade pode capturar espécies ameaçadas de extinção, ou até mesmo desrespeitar as restrições de áreas e períodos de pesca proibidos, podendo causar a devastação de habitats inteiros. Já, a pesca industrial, é a que mais promove riscos ao meio ambiente. Os navios pesqueiros e os arrastões das redes deixam um rastro de poluição, podendo contaminar ecossistemas marinhos. Nesse contexto, destaca-se a importância da modelagem matemática para prever e analisar o impacto da exploração, podendo, inclusive, evitar eventuais catástrofes (BORGES *et al.*, 2018). Esse trabalho propõe uma revisão de modelos discretos para a pesca de rendimento constante e a pesca de esforço constante.

Metodologia

Esse estudo considera a dinâmica de uma única espécie, e não leva em conta sua estrutura etária e sua distribuição espacial. Como variáveis discretas têm-se a população N variável no tempo, descrita pela sucessão N_k onde, $k = 1, 2, \dots$. Assim, N_k representa o nível da população em idade reprodutora (densidade de indivíduos ou biomassa da população) no ano, ou época reprodutora, ou geração k . Desta forma, existe uma dependência funcional entre o nível da população N_k no momento k , e o nível N_{k+1} , no momento $k + 1$, ou seja: $N_{k+1} = f(N_k)$. A partir do entendimento dessa dinâmica discreta, estuda-se o modelo proposto por Thomas Malthus, que observa a razão entre dois anos consecutivos de uma população, como sua razão de reprodutibilidade R_0 (COSTA, 2016). Dependendo dos fatores que podem influenciar o crescimento ou decrescimento da população, $R_0 > 1$ indica que a população da geração seguinte será superior a da geração presente, e se $R_0 < 1$, a população da geração presente será superior a da geração seguinte, e ainda se $R_0 = 1$, o nível da população não irá se alterar de uma geração para a seguinte. Estuda-se ainda o modelo logístico proposto por Verhust. A partir desses estudos, considera-se

uma população onde a introdução de pesca, captura ou colheita, na geração k , resulta numa diminuição da população em $k + 1$. Ou seja, a dinâmica do sistema passa a ser regida pela relação: $N_{k+1} = f(N_k) - H_k$, onde H_k é o número indivíduos na geração, ou época de pesca k , que são retirados do ecossistema pela atividade pesqueira. O trabalho estuda hipóteses sobre a forma como os indivíduos são pescados, ou seja, de que modo H_k depende de k . Descrever o comportamento de H_k como função de k , biologicamente significa conhecer a quantidade pescada em cada geração, ou em cada instante de tempo discreto. A proposta analisa duas formas de pesca. Na pesca de rendimento constante, retira-se do ecossistema uma quantidade constante $H > 0$, independente do tempo, ou da época de pesca k , e de N_k . Assim,

$$N_{k+1} = (1 + r)N_k \left(1 - \frac{N_k}{C}\right) - H, \quad (1)$$

satisfaz o modelo logístico, onde C é a capacidade de suporte do meio, e r é a taxa de variação da população. É evidente que tal estratégia pode ocasionar situações de não sustentabilidade se H for maior que o rendimento máximo sustentável, ou para valores inferiores próximos deste. A pesca de esforço constante, é uma prática que pode prevenir esse problema, que consiste em retirar do ecossistema, na unidade de tempo, uma quantidade de população proporcional ao nível populacional. Trata-se de considerar um rendimento de pesca dado por $H_k = E N_k$, onde $E = q\alpha$, sendo α o esforço de pesca, expresso em unidades de "barco padrão"BP, e $q \in (0, 1)$ a capturabilidade do recurso, medida em unidades de $BP^{-1} \times t^{-1}$, uma quantidade que traduz a proporção da população capturada por um barco padrão em cada unidade de tempo. Chama-se de esforço de pesca a quantidade E , assim têm-se a seguinte dinâmica:

$$N_{k+1} = (1 + r)N_k \left(1 - \frac{N_k}{C}\right) - E N_k. \quad (2)$$

Considerações Finais

Os modelos discretos de pesca de esforço constante e pesca de rendimento constante são analisados sob diferentes cenários, simulados através de uma implementação em Matlab. Assim, este trabalho, executado como tema de Iniciação Científica, contribui o estudo e planejamento da exploração dos recursos naturais (pesca), indispensável ao desenvolvimento econômico, considerando possíveis prejuízos em função da exploração excessiva e/ou desordenada desses recursos.

Referências

- COSTA, P. F. **Brevíssima Introdução a modelos matemáticos das pescas**, Universidade Aberta, Centro de Análise Matemática, Geometria e Sistemas Dinâmicos, IST. 2016. Disponível em: <http://hdl.handle.net/10400.2/5504>. Acesso em: 26 Ago 2023.
- BORGES, H.; DUARTE, R.; SEMPREBOM, T. R.; PEIRÓ, T. F. **Atividade Pesqueira e seus impactos no Meio Ambiente**. 2018. Disponível em: <https://www.bioicos.org.br/post/atividade-pesqueira-e-seus-impactos-no-meio-ambiente#:~:text=Na%20maioria%20das%20vezes%2C%20os,podem%20contaminar%20ecossistemas%20marinhos%20inteiros>. Acesso em: 26 Ago 2023.
- OCEANA. **Dia Mundial da Pesca: atividade é fonte de alimento e renda para milhões de pessoas**. 2022. Disponível em: <https://brasil.oceana.org/blog/dia-mundial-da-pesca-atividade-e-fonte-de-alimento-e-renda-para-milhoes-de-pessoas#:~:text=Cerca%20de%20600%20milh%C3%B5es%20de,em%20alguns%20pa%C3%ADses%20dessa%20regi%C3%A3o>. Acesso em: 26 Ago 2023.