



**INSTITUTO  
FEDERAL**

Goiás

---

Câmpus  
Goiânia

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE  
GOIÁS - CAMPUS GOIÂNIA**

**CURSO DE LICENCIATURA EM MATEMÁTICA**

**BRUNO RODRIGUES SOARES**

**EQUAÇÕES DIFERENCIAIS NO ESTUDO EPIDEMIOLÓGICO: MODELO SIR  
DE KERMACK E MCKENDRICK**

**GOIÂNIA - GO**

**2023**



**INSTITUTO FEDERAL**  
Goiás

**MINISTÉRIO DA EDUCAÇÃO**  
**SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA**  
**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**SISTEMA INTEGRADO DE BIBLIOTECAS**

### **TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO NO REPOSITÓRIO DIGITAL DO IFG - ReDi IFG**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia de Goiás, a disponibilizar gratuitamente o documento no Repositório Digital (ReDi IFG), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IFG.

#### **Identificação da Produção Técnico-Científica**

- |                                                                      |                                                         |
|----------------------------------------------------------------------|---------------------------------------------------------|
| <input type="checkbox"/> Tese                                        | <input type="checkbox"/> Artigo Científico              |
| <input type="checkbox"/> Dissertação                                 | <input type="checkbox"/> Capítulo de Livro              |
| <input type="checkbox"/> Monografia – Especialização                 | <input type="checkbox"/> Livro                          |
| <input checked="" type="checkbox"/> TCC - Graduação                  | <input type="checkbox"/> Trabalho Apresentado em Evento |
| <input type="checkbox"/> Produto Técnico e Educacional - Tipo: _____ |                                                         |

Nome Completo do Autor: Bruno Rodrigues Soares

Matrícula: 20191010940137

Título do Trabalho: Equações diferenciais no estudo epidemiológico: Modelo SIR de Kermack e McKendrick.

#### **Autorização - Marque uma das opções**

1. (X) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso aberto);
2. ( ) Autorizo disponibilizar meu trabalho no Repositório Digital do IFG somente após a data \_\_\_\_/\_\_\_\_/\_\_\_\_ (Embargo);
3. ( ) Não autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso restrito).

Ao indicar a opção **2 ou 3**, marque a justificativa:

- ( ) O documento está sujeito a registro de patente.  
( ) O documento pode vir a ser publicado como livro, capítulo de livro ou artigo.  
( ) Outra justificativa: \_\_\_\_\_

#### **DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA**

O/A referido/a autor/a declara que:

- i. o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- ii. obteve autorização de quaisquer materiais inclusos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- iii. cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

\_\_\_\_Goiânia-GO\_\_\_\_, \_\_\_\_11\_\_\_\_/\_\_\_\_12\_\_\_\_/2023\_\_\_\_.  
Local Data

*Bruno R. Soares*

Assinatura do Autor e/ou Detentor dos Direitos Autorais

**BRUNO RODRIGUES SOARES**

**EQUAÇÕES DIFERENCIAIS NO ESTUDO EPIDEMIOLÓGICO: MODELO SIR  
DE KERMACK E MCKENDRICK**

Trabalho de Conclusão de Curso apresentado ao curso de Licenciatura em matemática do Instituto Federal de Educação, Ciência e Tecnologia de Goiás - Câmpus Goiânia, como parte dos requisitos para obtenção grau de Licenciado em Matemática.

Orientador: Dr. Hugo Leonardo da Silva Belisário

**GOIÂNIA - GO**

**2023**

So11e Soares, Bruno Rodrigues.

Equações diferenciais no estudo epidemiológico: modelo SIR de Kermack e Mckendrick / Bruno Rodrigues Soares. – Goiânia: Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia, 2023.  
43 f.: il.

Orientador: Prof. Dr. Hugo Leonardo da Silva Belisário.

TCC (Trabalho de Conclusão de Curso) – Licenciatura em Matemática, Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia.

1. Equações diferenciais. 2. Modelo SIR. 3. Epidemiologia Matemática. I. Belisário, Hugo Leonardo da Silva (orientador). II. Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia. III. Título.

CDD 515.35

Ficha catalográfica elaborada pela Bibliotecária Karol Almeida da Silva CRB1/ 2.740  
Biblioteca Professor Jorge Félix de Souza,  
Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia.



**INSTITUTO FEDERAL**  
Goiás

MINISTÉRIO DA EDUCAÇÃO  
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA  
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE GOIÁS  
CÂMPUS GOIÂNIA

## EQUAÇÕES DIFERENCIAIS NO ESTUDO EPIDEMIOLÓGICO: MODELO SIR DE KERMACK E MCKENDRICK

por

**BRUNO RODRIGUES SOARES**

Trabalho de Conclusão de Curso apresentado ao Colegiado do Curso de Licenciatura em Matemática do Instituto Federal de Educação, Ciência e Tecnologia de Goiás – Câmpus Goiânia, como requisito parcial para a obtenção do Diploma de Licenciado em Matemática.

**Área de concentração:** Matemática

**Orientador:** Prof. Dr. Hugo Leonardo da Silva Belisário

Trabalho de Conclusão de Curso defendido e aprovado, em 24 de novembro de 2023, pela banca examinadora constituída pelos seguintes membros:

*(assinado eletronicamente)*

Prof. Dr. Hugo Leonardo da Silva Belisário (Presidente da Banca/IFG/Câmpus Goiânia)

*(assinado eletronicamente)*

Prof<sup>a</sup>. Dr<sup>a</sup>. Kamilla Machado Palhares (IFG/Câmpus Goiás)

*(assinado eletronicamente)*

Prof. Dr. Julio Cesar Saavedra Vasquez (IFG/Câmpus Goiânia)

Documento assinado eletronicamente por:

- Julio Cesar Saavedra Vasquez, PROFESSOR ENS BASICO TECN TECNOLÓGICO, em 24/11/2023 14:24:20.
- Kamilla Machado Palhares, PROFESSOR ENS BASICO TECN TECNOLÓGICO, em 24/11/2023 14:12:54.
- Hugo Leonardo da Silva Belisario, PROFESSOR ENS BASICO TECN TECNOLÓGICO, em 24/11/2023 14:09:09.

Este documento foi emitido pelo SUAP em 01/11/2023. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifg.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 471377

Código de Autenticação: 80519096d2



## **AGRADECIMENTOS**

Primeiramente, expresso minha profunda gratidão á minha família que sempre me deu amparo e garantiu minha permanência no curso. Agradeço também o Professor Hugo Belisário que me ajudou , por sua orientação, paciência e incentivo ao longo deste projeto. Suas sugestões e conhecimentos compartilhados foram essenciais.

À equipe docente do curso de Licenciatura em Matemática do IFG, em especial aos professores que ministraram as disciplinas relacionadas a cálculo, agradeço por transmitirem seu conhecimento de maneira envolvente e inspiradora. Cada aula, conversa e explicação contribuíram para a minha compreensão, interesse e permanência no curso.

Não posso deixar de expressar minha gratidão aos meus colegas e amigos de classe, cujas discussões e trocas de ideias enriqueceram meu aprendizado. As longas horas de estudo em grupo e os debates calorosos foram fundamentais para o meu crescimento acadêmico e pessoal.

Por fim, desejo estender meus agradecimentos a todos que, de alguma forma, contribuíram para este trabalho e para a minha formação. Cada interação, lição aprendida e experiência vivida deixaram uma marca indelével em minha trajetória acadêmica.

*“Vivemos em uma sociedade extremamente dependente da ciência e tecnologia, na qual pouquíssimos sabem alguma coisa sobre ciência e tecnologia. Isto é uma clara prescrição para o desastre”.*

---

Carl Sagan

## RESUMO

Este trabalho aborda o estudo da epidemiologia através do modelo SIR, que se fundamenta na modelagem de problemas com o uso de equações diferenciais. Esta pesquisa visa principalmente compreender, investigar e divulgar os conhecimentos básicos do Modelo SIR da epidemiologia matemática, para isto, é apresentada uma introdução que contextualiza a importância das equações diferenciais na epidemiologia, destacando o papel de William Ogilvy Kermack e Anderson Gray McKendrick no desenvolvimento do campo da epidemiologia matemática. É feita uma breve fundamentação teórica sobre resultados de equações diferenciais relacionados a este campo. São apresentados alguns conceitos básicos da epidemiologia, preparando o terreno para a análise do modelo SIR através da análise qualitativa de interpretação de resultados matemáticos que podem ser obtidos. Para a análise do modelo SIR, é abordado o uso do teorema de Hartman–Grobman para aproximações lineares de sistemas não lineares e ainda as ideias de estabilidade e instabilidade do sistema. O trabalho conclui com um exemplo prático de aplicação do modelo SIR usando a linguagem de programação R.

**Palavras-chave:** Equações Diferenciais, Modelo SIR, Epidemiologia Matemática.

## ABSTRACT

This monograph deals with the study of epidemiology using the SIR model, which is based on modeling problems using differential equations. This research mainly aims to understanding, investigating and disseminating the basic knowledge of the SIR model of mathematical epidemiology. equations in epidemiology, highlighting the role of William Ogilvy Kermack and Anderson Gray McKendrick in the development of the field of mathematical epidemiology. A brief theoretical foundation is given on the results of differential equations related to this camp. Some basic concepts of epidemiology are presented, which prepares the way for the analysis of the SIR model through a qualitative analysis of the interpretation of the mathematical results that can be obtained. For the analysis of the SIR model, the use of the theorem Hartman–Grobman for linear approximations of nonlinear systems and also the ideas of system stability and instability. The research concludes with a practical application example of the SIR model using the R programming language.

**Keywords:** 1. Differential Equations 2. SIR Model 3. Mathematical Epidemiology.

## LISTA DE FIGURAS

|          |                                                                                                                                                                                                                                                                                                                                  |    |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1 | – Diagrama do modelo de compartimentos SIR, onde $S$ representa os suscetíveis, $I$ os infectados e $R$ os removidos. . . . .                                                                                                                                                                                                    | 33 |
| Figura 2 | – Linha do código de definição de parâmetros em R. . . . .                                                                                                                                                                                                                                                                       | 39 |
| Figura 3 | – Tempo em que ocorreu a pandemia do Exemplo, do dia 1 ao dia 200, contado um-a-um. . . . .                                                                                                                                                                                                                                      | 39 |
| Figura 4 | – Descrição do modelo SIR e os parâmetros do exemplo. . . . .                                                                                                                                                                                                                                                                    | 40 |
| Figura 5 | – Código para resolver as equações do sistema. . . . .                                                                                                                                                                                                                                                                           | 40 |
| Figura 6 | – Código usado para plotar o gráfico. . . . .                                                                                                                                                                                                                                                                                    | 40 |
| Figura 7 | – A solução numérica obtida pela linguagem de programação R para um surto de Covid-19 na cidade de Aparecida de Goiânia-Goiás, com uma taxa de infecção $\beta = 0,2969558$ . Note que pouco antes do quinquagésimo dia o número de infectados começa a aumentar e atinge seu pico por volta do septuagésimo quinto dia. . . . . | 41 |

## SUMÁRIO

|            |                                                                                      |           |
|------------|--------------------------------------------------------------------------------------|-----------|
| <b>1</b>   | <b>INTRODUÇÃO . . . . .</b>                                                          | <b>12</b> |
| <b>1.1</b> | <b>Breve histórico das equações diferenciais . . . . .</b>                           | <b>12</b> |
| <b>1.2</b> | <b>Equações diferenciais no estudo epidemiológico . . . . .</b>                      | <b>13</b> |
| <b>1.3</b> | <b>William Ogilvy Kermack e Anderson Gray McKendrick . . . . .</b>                   | <b>14</b> |
| <b>1.4</b> | <b>Objetivos . . . . .</b>                                                           | <b>15</b> |
| <b>2</b>   | <b>NOÇÕES DE EQUAÇÕES DIFERENCIAIS . . . . .</b>                                     | <b>16</b> |
| <b>2.1</b> | <b>Equações diferenciais . . . . .</b>                                               | <b>16</b> |
| 2.1.1      | Solução explícitas e implícitas de uma equação diferencial . . . . .                 | 17        |
| 2.1.2      | Problema de valor inicial . . . . .                                                  | 17        |
| <b>2.2</b> | <b>Equações separáveis . . . . .</b>                                                 | <b>20</b> |
| <b>2.3</b> | <b>EDO's lineares de primeira ordem e o método de fatores integrantes . .</b>        | <b>22</b> |
| 2.3.1      | Método de resolução pelos fatores integrantes . . . . .                              | 22        |
| <b>2.4</b> | <b>Sistemas de equações diferenciais lineares de primeira ordem . . . . .</b>        | <b>24</b> |
| 2.4.1      | Sistema de equações lineares . . . . .                                               | 24        |
| 2.4.2      | Autovalores e autovetores . . . . .                                                  | 24        |
| 2.4.3      | Outros resultados importantes de álgebra matricial . . . . .                         | 25        |
| <b>2.5</b> | <b>Estabilidade e instabilidade . . . . .</b>                                        | <b>28</b> |
| <b>2.6</b> | <b>Aproximações lineares de sistemas não lineares . . . . .</b>                      | <b>28</b> |
| <b>3</b>   | <b>EPIDEMIOLOGIA MATEMÁTICA . . . . .</b>                                            | <b>30</b> |
| <b>3.1</b> | <b>Conceitos básicos . . . . .</b>                                                   | <b>30</b> |
| 3.1.1      | Lei de ação das massas . . . . .                                                     | 31        |
| <b>3.2</b> | <b>Convergência entre a biologia realista e a complexidade matemática . .</b>        | <b>31</b> |
| <b>3.3</b> | <b>Modelos compartimentais determinísticos . . . . .</b>                             | <b>32</b> |
| <b>4</b>   | <b>O MODELO SIR . . . . .</b>                                                        | <b>34</b> |
| <b>4.1</b> | <b>Número Básico de reprodução (<math>R_0</math>) e o limiar epidêmico . . . . .</b> | <b>37</b> |
| <b>4.2</b> | <b>Análise do modelo SIR . . . . .</b>                                               | <b>37</b> |
| <b>4.3</b> | <b>Um exemplo simples de aplicação do modelo SIR em linguagem R . . .</b>            | <b>38</b> |
| <b>5</b>   | <b>CONSIDERAÇÕES FINAIS . . . . .</b>                                                | <b>42</b> |
|            | <b>REFERÊNCIAS . . . . .</b>                                                         | <b>43</b> |

## 1 INTRODUÇÃO

A matemática é uma ciência fundamental que possui diversas aplicações em áreas como física, engenharia, finanças, computação, entre outras. No entanto, a formação em matemática envolve muito mais do que apenas a compreensão de suas aplicações práticas. Ela também envolve o desenvolvimento de habilidades analíticas e lógicas, a compreensão de conceitos abstratos e o estudo de diferentes áreas da matemática pura.

Por essa razão, muitos cursos de graduação em matemática têm pouco espaço para aprofundar as aplicações práticas da matemática, uma vez que a ênfase é dada à construção de bases sólidas em teoria matemática. No entanto, isso não significa que as aplicações sejam negligenciadas. Ao contrário, as aplicações são abordadas de forma mais ampla, e não apenas em uma única área de aplicação.

O trabalho de conclusão de curso, por sua vez, é uma oportunidade para os estudantes de matemática se aprofundarem em um tópico específico e explorarem suas aplicações de forma mais detalhada. Além disso, o trabalho de conclusão de curso também oferece a oportunidade de explorar novas áreas da matemática, que podem não ter sido abordadas durante o curso de graduação. Isso é importante não apenas para o desenvolvimento pessoal do estudante, mas também para o avanço do conhecimento matemático como um todo.

Neste sentido, o trabalho se justifica pela necessidade de discutir estudos que não são compreendidos pelos conteúdos das ementas do curso de Licenciatura em Matemática, a fim de complementar a formação acadêmica e integrar o conhecimento específico do campo de estudo da biologia e afins.

### 1.1 Breve histórico das equações diferenciais

Segundo Çengel e III (2014), a expressão "equações diferenciais" é usada desde 1676, quando Gottfried Leibniz (1646 – 1716) a criou, e desde então, essa categoria de equações vem sendo usada por uma gama de cientistas para resolver problemas práticos de diversas áreas do conhecimento. Boyce e DiPrima (2010) destacam também as contribuições de Isaac Newton (1642 – 1727) para o desenvolvimento das equações diferenciais, que de certa forma contribuiu pouco para esta área, entretanto, seus estudos sobre mecânica ajudaram a estabelecer os caminhos para as aplicações de equações diferenciais ainda no século XVII.

Boyce e DiPrima (2010) dão ênfase a força das notações matemáticas usadas por Leibniz, que usamos ainda atualmente, como a notação de integral  $dy/dx$ , além disso, criou o método de resolução de EDO's muito popular nos cursos de graduação atualmente, o método de separação de variáveis e o método de resolução de equações lineares de primeira ordem. Ainda temos o destaque das contribuições dos irmãos Jakob Bernoulli (1654 – 1705) e Johann Bernoulli (1667

– 1748) para o desenvolvimento de métodos de resolução de equações diferenciais.

No século XIX, a teoria das equações diferenciais continuaram a se desenvolver, com importantes contribuições de matemáticos como Lagrange (1707 – 1783), Joseph Fourier (1768 – 1830), Cauchy (1789 – 1857) e Carl Jacobi (1804 – 1851). Com o passar do tempo, o estudo das equações diferenciais ganhou uma disciplina própria nos cursos superiores de matemática, além disso, suas aplicações ganharam espaço em outras áreas de estudo para além da matemática em si através da modelagem de problemas reais e práticos. O uso dessas equações para modelagem de problemas reais é uma de suas aplicações mais importantes e significativas ao longo da história.

Segundo Gray (2021), o uso de equações diferenciais para resolver problemas práticos tem uma longa história que remonta aos séculos XVII e XVIII. Um exemplo notável de como as equações diferenciais foram usadas para resolver problemas reais é a modelagem do movimento dos corpos celestes. Isaac Newton usou equações diferenciais para descrever o movimento dos planetas em torno do Sol, e essa técnica se tornou uma das aplicações mais importantes das equações diferenciais.

Além da física, as equações diferenciais também foram usadas na modelagem de fenômenos biológicos. Por exemplo, em meados do século XIX, o matemático francês Joseph Fourier usou equações diferenciais para descrever o fluxo de calor em um corpo vivo. Esse trabalho é um dos primeiros exemplos da aplicação de equações diferenciais para entender a biologia. Outra aplicação importante das equações diferenciais é a modelagem de sistemas econômicos. O matemático francês Antoine-Augustin Cournot (1801 – 1877) usou equações diferenciais para modelar a dinâmica da oferta e demanda em um mercado (GRAY, 2021).

Em resumo, o uso de equações diferenciais para modelagem de problemas reais é uma das aplicações mais importantes e significativas desse campo. Desde a modelagem do movimento dos corpos celestes até a modelagem de sistemas econômicos complexos, as equações diferenciais são usadas para resolver problemas práticos em várias áreas da ciência e engenharia.

## 1.2 Equações diferenciais no estudo epidemiológico

Um campo importante de aplicação das equações diferenciais é a epidemiologia, sendo a ciência que estuda a distribuição e os determinantes das doenças em populações humanas. Neste tópico, exploraremos uma breve contextualização histórica de como se deu aplicação de equações diferenciais nos estudos epidemiológicos.

Os desafios da biologia que lidam com a propagação de doenças infecciosas são considerados uns dos mais compreensíveis para aqueles que não possuem formação na área. Esses problemas seguem uma estrutura matemática estabelecida, que foi desenvolvida e aprimorada ao longo dos últimos séculos por muitos pesquisadores, aumentando sua confiabilidade (BRAUER *et al.*, 2019). Essa tendência dos modelos de se adaptarem aos problemas reais, torna este método mais atrativo para os pesquisadores.

A epidemiologia matemática tem raízes antigas, iniciando com o modelo de varíola de Daniel Bernoulli (1700 – 1782) em 1760. Com as descobertas e avanços de vários tópicos do estudo do cálculo diferencial, o uso de equações diferenciais se tornou um grande aliado da modelagem de problemas voltados a biologia e por consequência, a epidemiologia. Entre 1900 e 1935, muita teoria básica foi criada, e desde então houve progresso constante. Brauer *et al.* (2008) destacam que:

As primeiras contribuições para a epidemiologia matemática moderna são atribuídas a Pyotr Dimitrievich (1844 – 1916) entre 1873 e 1894, e as bases de toda a abordagem da epidemiologia baseada em modelos compartimentais foram estabelecidas por médicos de saúde pública como Sir Ronald Ross, R.A., William Heaton Hamer, Aderson Gray McKendrick e William Ogilvy Kermack entre 1900 e 1935, juntamente com importantes contribuições do ponto de vista estatístico por John Brownlee.

Um caso exemplo importante é o trabalho de Sir Ronald Ross, que estudou a transmissão da malária entre mosquitos e humanos. Inicialmente, suas conclusões foram rejeitadas, mas quando ele desenvolveu um modelo matemático que previa que surtos de malária poderiam ser evitados se a população de mosquitos fosse reduzida abaixo de um nível crítico, suas ideias foram comprovadas por testes de campo e levaram a sucessos brilhantes no controle da malária, o que lhe rendeu o Prêmio Nobel de Medicina em 1902 pela descoberta do método de transmissão da malária.

Brauer *et al.* (2008), também destacam que a aplicação de equações diferenciais em estudos epidemiológicos permitiu a incorporação de fatores como a imunidade e a vacinação em modelos epidemiológicos. Por exemplo, em um modelo SIR (suscetível-infectado-recuperado), do qual falaremos mais adiante, os indivíduos não podem se tornar suscetíveis novamente após a recuperação ou vacinação. Essas informações são cruciais para entender como a imunidade de grupo pode ser alcançada e como as campanhas de vacinação podem afetar a transmissão de uma doença infecciosa. Isso permite que os órgãos públicos de saúde possam fazer uma melhor análise da dinâmica de propagação de doenças e desenvolver estratégias de prevenção e controle. Estes conceitos de imunidade e vacinação foram introduzidos pelos escoceses Aderson Gray McKendrick e William Ogilvy Kermack em artigos publicados em 1927, 1932 e 1933, onde também demonstraram que, ao contrário do que se pensava antes, não é preciso uma vacinação total para erradicar uma doença por completo (KRETZSCHMAR; WALLINGA, 2010 apud FARIAS, 2017).

### 1.3 William Ogilvy Kermack e Anderson Gray McKendrick

De acordo com Smith e Kuh (2001), Kermack nasceu em Kirriemuir, um burgo de Angus, na Escócia, em 26 de abril de 1898, era filho único de William Kermack, um carteiro, e Helen Ogilvy. Teve sua educação primária na zona rural da Escócia e em 1918 se graduou em Matemática e Filosofia Natural em *Aberdeen University*, uma universidade pública tradicional e de prestígio da Escócia. Ainda durante o final de sua graduação se envolveu em estudos de

análise estatística e produção de leite, sendo apresentado às pesquisas de Karl Pearson (1857 – 1936), fundador do primeiro departamento universitário dedicado à estatística em todo o mundo e "*Student*", pseudônimo de William Sealy Gosset (1876 – 1937), conhecido pela distribuição t de *Student*.

Serviu a Força Aérea Real por um curto período, após isso, se envolveu em pesquisas com química industrial, que o levou a trabalhos de química no Laboratório do *Royal College of Physicians*, em Edimburgo, onde ficou por boa parte de sua vida, mas um acidente em 1924 o deixou cego. Com a ajuda de colegas, continuou seu trabalho, agora focado em química coloidal e medicamentos antimaláricos. Foi na *Royal College of Physicians* que conheceu Anderson Gray McKendrick, com quem teve uma colaboração proveitosa. Mesmo cego, ele demonstrou habilidades matemáticas notáveis. Suas contribuições marcaram avanços no entendimento da matemática aplicada ao estudo epidemiológico, conhecido atualmente como epidemiologia matemática.

Dietz (1997) nos traz um breve histórico sobre Anderson Gray McKendrick. Nascido em 8 de setembro de 1876 em Edimburgo da Escócia e depois em Glasgow, ainda criança, onde seu pai tinha um cargo importante na área acadêmica. Ele se formou em medicina aos 24 anos e logo se interessou por pesquisas sobre malária e matemática aplicada. Ele começou sua carreira no *Indian Medical Service (IMS)*, trabalhando em expedições para estudar a malária. Em uma colaboração de pesquisa com Ronald Ross, foi influenciado a explorar métodos matemáticos em suas pesquisas. Mais tarde, McKendrick focou em estudar a zoonose da raiva e se tornou uma autoridade reconhecida nesse campo. Além disso, ele contribuiu para a teorias matemáticas e estatísticas.

Sua trajetória o levou a trabalhar em institutos de pesquisa médica na Índia, notavelmente no Instituto Pasteur em Kasauli. Ele também atuou como superintendente do Laboratório do *Royal College of Physicians* em Edimburgo até sua aposentadoria em 1941, onde conheceu Kermack e desenvolveram juntos o modelo estudado neste trabalho.

## 1.4 Objetivos

Neste trabalho, nos atentaremos a descrever o modelo SIR (suscetível-infectado-recuperado), de McKendrick e Kermack. O modelo e teorias analisados neste trabalho segue principalmente o livro *Mathematical Epidemiology: Lecture Notes in Mathematics Book 1945*, dos autores Fred Brauer, Jianhong Wu e Pauline van den Driessche entre outros, que serão citados no decorrer do trabalho. Para o modelo SIR, primeiro o exploraremos em sua forma básica, e posteriormente acrescentaremos um exemplo simples usando a linguagem de programação R com objetivo de compreender, investigar e divulgar os conhecimentos básicos do Modelo SIR da epidemiologia matemática.

## 2 NOÇÕES DE EQUAÇÕES DIFERENCIAIS

Para que se tenha total entendimento do modelo SIR, é necessário definirmos alguns pontos importantes da teoria de equações diferenciais, para isso, este capítulo se propõe a apresentar alguns tópicos importantes de equações diferenciais ordinárias. Como referência para este capítulo, nos basearemos nas obras Equações Diferenciais Elementares e Problemas de Valores de Contorno de Boyce (2016), Equações Diferenciais Volume 1 de Zill (2016) Introdução às Equações Diferenciais Ordinárias de Santos (2011).

### 2.1 Equações diferenciais

Nesta seção apresentaremos definições básicas e resultados importantes e os métodos de resolução de equações diferenciais usados na epidemiologia matemática. Estas definições têm como principal referência a obra (BOYCE; DIPRIMA, 2010).

#### **Definição 1.** *Equação diferencial*

*Uma equação que contém as derivadas ou diferenciais de uma ou mais variáveis dependentes em relação a uma ou mais variáveis independentes, é chamada de equação diferencial (ED).*

Equações diferenciais são classificadas de acordo com tipo, ordem e linearidade.

#### **Definição 2.** *Equação diferencial ordinária*

*Uma equação que contém somente derivadas ordinárias de uma ou mais variáveis dependentes, com relação a uma única variável dependente, ela é chamada de equação diferencial ordinária (EDO).*

#### **Definição 3.** *Solução para uma equação diferencial*

*Qualquer função  $f$  definida em algum intervalo  $I$ , que, quando substituída na equação diferencial, reduz a equação a uma identidade, é chamada solução para a equação no intervalo.*

Em outras palavras, uma solução para uma equação diferencial ordinária

$$F(x, y, y', \dots, y^{(n)}) = 0 \quad (2.1)$$

é uma função que  $f$  que possui pelo menos  $n$  derivadas e satisfaz a equação, isto é,

$$F(x, f(x), f'(x), \dots, f^{(n)}(x)) = 0 \quad (2.2)$$

para todo  $x$  no intervalo  $I$ . Propositadamente, deixamos vago a forma precisa do intervalo  $I$  na Definição 3. Dependendo do contexto da discussão,  $I$  pode representar um intervalo aberto  $(a, b)$ , um intervalo fechado  $[a, b]$ , um intervalo infinito  $(0, \infty)$  e assim por diante.

### 2.1.1 Solução explícitas e implícitas de uma equação diferencial

**Definição 4.** *Solução explícita* Uma solução para uma equação diferencial que pode ser escrita na forma  $y = f(x)$  é chamada de solução explícita.

**Definição 5.** *Solução implícita* Dizemos que uma relação  $G(x, y) = 0$  é uma solução implícita de uma equação diferencial ordinária em um intervalo  $I$ , se ela define uma ou mais soluções explícitas em  $I$ .

### 2.1.2 Problema de valor inicial

Em algum intervalo  $I$  contendo  $x_0$ , o problema de resolver uma equação diferencial de  $n$ -ésima ordem sujeita a  $n$  condições de contorno especificadas em  $x_0$ :

$$\text{Resolver: } \frac{d^n y}{dx^n} = f(x, y, y', \dots, y^{n-1}) \quad (2.3)$$

$$\text{Sujeito a: } y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{n-1}(x_0) = y_{n-1} \quad (2.4)$$

onde  $y_0, y_1, \dots, y_{n-1}$  são constantes reais especificadas, é chamado de problema de valor inicial (P.V.I). Os valores de  $y(x)$  e suas  $n-1$  derivadas em um único ponto  $x_0$  :  $y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{n-1}(x_0) = y_{n-1}$  são chamados de condições iniciais.

Resolver um problema de valor inicial de ordem  $n$  como (2.3) em geral, envolve o uso de uma família a  $n$  parâmetros de soluções da equação diferencial dada e, em seguida, usando as  $n$  condições iniciais de  $x_0$  para determinar as  $n$  constantes da família. A solução particular resultante é definida em algum intervalo  $I$  contendo o ponto inicial  $x_0$ . Os casos  $n = 1$  e  $n = 2$  em (2.3),

$$\text{resolver: } \frac{dy}{dx} = f(x, y), \quad (2.5)$$

$$\text{sujeito a: } y(x_0) = y_0 \quad (2.6)$$

e

$$\text{Resolver: } \frac{dy}{dx} = f(x, y, y') \quad (2.7)$$

$$\text{Sujeito a: } y(x_0) = y_0, y'(x_0) = y_1 \quad (2.8)$$

são problemas de valor inicial de primeira e segunda ordem, respectivamente. Neste trabalho, para a modelagem e solução dos problemas relacionados aos modelos SIS e SIR, nos atentaremos em específico aos P.V.I's de primeira ordem, onde procuraremos uma solução da equação diferencial  $y' = f(x, y)$  em um intervalo  $I$  contendo  $x_0$  de tal forma que o seu gráfico passe pelo ponto  $(x_0, y_0)$ .

**Teorema 1** (Teorema de Picard-Lindelöf). *Seja  $R$  uma região retangular no plano  $xy$  definida por  $\alpha \leq x \leq \beta, \gamma \leq y \leq \delta$ , que contém o ponto  $(x_0, y_0)$  em seu interior. Se  $f(x, y)$  e  $\frac{\partial f}{\partial y}$  são*

contínuas em  $R$ , então existe um intervalo  $I$  centrado em  $x_0$  e uma única função  $y(x)$  definida em  $I$  que satisfaz o problema de valor inicial (2.3).

*Demonstração.* Parte 1: Existência

Defina a sequência de funções  $y_n(x)$  por

$$y_0(t) = y_0, \quad y_n(t) = y_0 + \int_{x_0}^x f(s, y_{n-1}(s)) ds, \text{ para } n = 1, 2, \dots$$

Como  $f(x, y)$  é contínua no retângulo  $R$ , existe uma constante positiva  $b$  tal que

$$|f(x, y)| \leq b, \text{ para } (x, y) \in R$$

Assim

$$|y_1(x) - y_0| \leq b|x - x_0|, \text{ para } \alpha < x < \beta$$

Como  $\frac{\partial f}{\partial y}$  é contínua no retângulo  $R$ , existe uma constante positiva  $a$  tal que

$$|f(x, y) - f(x, z)| \leq a|y - z| \quad (2.9)$$

para  $\alpha \leq x \leq \beta$  e  $\delta \leq y, z \leq \gamma$

Assim

$$\begin{aligned} |y_2(x) - y_1(x)| &\leq \int_{x_0}^x |f(s, y_1(s)) - f(s, y_0(s))| ds \\ &\leq a \int_{x_0}^x |y_1(s) - y_0(s)| ds \leq ab \int_{x_0}^x |s - x_0| ds = ab \frac{|x - x_0|^2}{2} \end{aligned}$$

e

$$\begin{aligned} |y_3(x) - y_2(x)| &\leq \int_{x_0}^x |f(s, y_2(s)) - f(s, y_1(s))| ds \\ &\leq a \int_{x_0}^x |y_2(s) - y_1(s)| ds \\ &\leq a^2 b \int_{x_0}^x \frac{|x - x_0|^2}{2} ds \\ &= a^2 b \frac{|x - x_0|^3}{6} \end{aligned}$$

Vamos supor, por indução, que

$$|y_{n-1}(x) - y_{n-2}(x)| \leq a^{n-2} b \frac{|x - x_0|^{n-1}}{(n-1)!}$$

Então

$$\begin{aligned} |y_n(x) - y_{n-1}(x)| &\leq \int_{x_0}^x |f(s, y_{n-1}(s)) - f(s, y_{n-2}(s))| ds \\ &\leq a \int_{x_0}^x |y_{n-1}(s) - y_{n-2}(s)| ds \leq a \int_{x_0}^x a^{n-2} b \frac{|s - x_0|^{n-1}}{(n-1)!} ds = a^{n-1} b \frac{|x - x_0|^n}{n!} \end{aligned}$$

Estas desigualdades são válidas para  $\alpha \leq \alpha' < x < \beta' \leq \beta$  em que  $\alpha'$  e  $\beta'$  são tais que  $\delta < y_n(x) < \gamma$  sempre que  $\alpha' < x < \beta'$ . Segue de (2.9) que

$$\sum_{n=1}^{\infty} |y_n(x) - y_{n-1}(x)| \leq \sum_{n=1}^{\infty} \frac{a^{n-1}(\beta - \alpha)^n}{n!}$$

que é convergente pelo critério de comparação, pois  $\sum_{n=1}^{\infty} \frac{a^{n-1}(\beta - \alpha)^n}{n!}$  é convergente. Como a soma original é menor ou igual a essa série convergente, concluímos que a soma original também é convergente.

Como

$$y_n(x) = y_0 + \sum_{k=1}^n (y_k(x) - y_{k-1}(x)),$$

então  $y_n(x)$  é convergente. Seja

$$y(x) = \lim_{n \rightarrow \infty} y_n(x)$$

Como

$$|y_m(x) - y_n(x)| \leq \sum_{k=n+1}^m |y_k(x) - y_{k-1}(x)| \leq b \sum_{k=n+1}^m \frac{a^{k-1}(\beta - \alpha)^k}{k!}$$

então, passando o limite quando  $m$  tende a infinito temos que

$$|y(x) - y_n(x)| \leq b \sum_{k=n+1}^{\infty} \frac{a^{k-1}(\beta - \alpha)^k}{k!} \quad (2.10)$$

Logo dado um  $\varepsilon > 0$ , para  $n$  suficientemente grande,  $|y(x) - y_n(x)| < \frac{\varepsilon}{3}$ , para  $\alpha' < x < \beta'$ . Daí segue que  $y(x)$  é contínua, pois dado um  $\varepsilon > 0$ , para  $s$  suficientemente próximo de  $x$ , temos que  $|y_n(x) - y_n(s)| < \frac{\varepsilon}{3}$  e para  $n$  suficientemente grande  $|y(x) - y_n(x)| < \frac{\varepsilon}{3}$  e  $|y(s) - y_n(s)| < \frac{\varepsilon}{3}$ , o que implica que

$$|y(x) - y(s)| \leq |y(x) - y_n(x)| + |y_n(x) - y_n(s)| + |y_n(s) - y(s)| < \varepsilon.$$

Além disso, para  $\alpha' < x < \beta'$ , temos que

$$\lim_{n \rightarrow \infty} \int_{x_0}^x f(s, y_n(s)) ds = \int_{x_0}^x f(s, \lim_{n \rightarrow \infty} y_n(s)) ds = \int_{x_0}^x f(s, y(s)) ds$$

pois, por (2.10), temos

$$\begin{aligned} & \left| \int_{x_0}^x f(s, y_n(s)) ds - \int_{x_0}^x f(s, y(s)) ds \right| \\ & \leq \int_{x_0}^x |f(s, y_n(s)) - f(s, y(s))| ds \\ & \leq a \int_{x_0}^x |y_n(s) - y(s)| ds \\ & \leq ab(x - x_0)ab \sum_{k=n+1}^{\infty} \frac{a^{k-1}(\beta - \alpha)^k}{k!} \end{aligned}$$

que tende a zero quando  $n$  tende ao infinito. Portanto

$$\begin{aligned} y(x) &= \lim_{n \rightarrow \infty} y_n(x) = y_n(x) = y_0 + \lim_{n \rightarrow \infty} \int_{x_0}^x f(s, y(s)) ds = \\ &= y_0 + \int_{x_0}^x f(s, \lim_{n \rightarrow \infty} y_{n-1}(s)) ds = y_0 + \int_{x_0}^x f(s, y(s)) ds \end{aligned}$$

Derivando em relação a  $x$  esta equação, veremos que  $y(x)$  é a solução do problema de valor inicial.

Parte 2, Unicidade:

par Suporemos que  $y(x)$  e  $z(x)$  sejam soluções do P.V.I. Seja

$$u(x) = \int_{x_0}^x |y(s) - z(s)| ds. \quad (2.11)$$

Assim, como

$$\begin{aligned} y(x) &= \int_{x_0}^x y'(s) ds = \int_{x_0}^x f(s, y(s)) ds \\ z(x) &= \int_{x_0}^x z'(s) ds = \int_{x_0}^x f(s, z(s)) ds \end{aligned}$$

então

$$\begin{aligned} u'(x) &= |y(x) - z(x)| \\ &\leq \int_{x_0}^x |y'(s) - z'(s)| ds = \int_{x_0}^x |f(s, y(s)) - f(s, z(s))| ds \\ &\leq a \int_{x_0}^x |y(s) - z(s)| ds \end{aligned}$$

ou seja,

$$u'(x) \leq au(x). \quad (2.12)$$

Subtraindo-se  $au(x)$  e multiplicando-se por  $e^{-ax}$  obtemos

$$\frac{d}{dx}(e^{-ax}u(x)) \leq 0, \text{ com } u(x_0) = 0$$

Isto implica que  $e^{-ax}u(x) = 0$  (lembre-se que  $u(x) \geq 0$  e portanto que  $u(x) = 0$ , para todo  $x$ . Assim,  $y(x) = z(x)$ , para todo  $x$ .  $\square$

Esta demonstração também pode ser encontrada em (SANTOS, 2011).

Acerca do teorema acima, a garantia de existência da solução não significa conseguir exibir a solução explícita do P.V.I, uma solução pode existir e não ser possível expressá-la por meio de funções elementares, apenas por métodos numéricos. As condições do Teorema da Existência de uma única Solução são suficientes, mas não necessárias. Nas próximas seções, estudaremos métodos de soluções de equações diferenciais de primeira ordem essenciais para o estudo de modelos epidemiológicos.

## 2.2 Equações separáveis

De acordo com Boyce e DiPrima (2010), uma EDO separável, ou de variáveis separáveis, é uma equação diferencial

$$\frac{dy}{dx} = f(x, y)$$

onde a função desconhecida e a variável independente podem ser separadas em fatores que dependem apenas de uma delas, podendo ser reescritas na forma

$$\frac{dy}{dx} = h(y)g(x),$$

onde  $h(y)$  depende apenas da função desconhecida  $y$  e  $g(x)$  depende apenas da variável independente  $x$ . A equação pode ser resolvida por meio da separação das variáveis, trazendo todos os termos que envolvem  $y$  para um lado da equação e todos os termos que envolvem  $x$  para o outro lado. Em seguida, se  $g(x)$  for contínua, podemos integrar ambos os lados da equação, obtendo a solução geral. De forma mais detalhada, tomemos a EDO

$$\frac{dy}{dx} = h(y)g(x). \quad (2.13)$$

Dividindo a equação (2.13) por  $h(y)$ , podemos reescrevê-la como

$$p(y)\frac{dy}{dx} = g(x), \quad (2.14)$$

onde, por conveniência, denotamos  $\frac{1}{h(y)}$  por  $p(y)$ .

Agora, se  $y = \phi(x)$  representa uma solução de (2.14), devemos ter  $p(\phi(x))\phi'(x) = g(x)$  e, portanto,

$$\int p(\phi(x))\phi'(x)dx = \int g(x)dx. \quad (2.15)$$

Mas  $dy = \phi'(x)dx$ , portanto, (2.15) é o mesmo que

$$\int p(y)dy = \int g(x)dx \text{ ou } H(y) = G(x) + c, \quad (2.16)$$

onde  $H(y)$  e  $G(x)$  são antiderivadas de  $p(y) = \frac{1}{h(y)}$  e  $g(x)$ , respectivamente.

“Não há necessidade de usar duas constantes na integração de uma equação separável, pois, se escrevermos  $H(y) + c_1 = G(x) + c_2$ , a diferença  $c_2 - c_1$  poderá ser substituída por uma única constante  $c$ , como em (2.16).” (ZILL, 2016, pg. 50).

**Exemplo 1.** Resolva  $(e^{2y} - y)\cos x \frac{dy}{dx} = e^y \sin 2x$ ,  $y(0) = 0$

*Solução:*

Dividindo a equação por  $e^y \cos x$  temos

$$\frac{e^{2y} - y}{e^y} dy = \frac{\sin 2x}{\cos x} dx.$$

Antes de integrar, usamos divisão termo a termo, no lado esquerdo, e a identidade trigonométrica  $\sin 2x = 2 \sin x \cos x$  no lado direito. Então

$$\int (e^y - ye^{-y}) dy = 2 \int \sin x dx$$

integrando por partes resulta em

$$e^y + ye^{-y} + e^{-y} = -2 \cos x + c.$$

A condição inicial  $y = 0$  quando  $x = 0$  implica  $c = 4$ . Assim, a solução do problema de valor inicial é

$$e^y + ye^{-y} + e^{-y} = 4 - 2\cos x.$$

□

## 2.3 EDO's lineares de primeira ordem e o método de fatores integrantes

Nesta seção, com base nos estudos apresentados por Boyce e DiPrima (2010), Zill (2016) e Diacu (2004), exploraremos um pouco das equações lineares de primeira ordem, as quais são um tipo específico de equações diferenciais que envolvem uma única variável independente e uma única variável dependente, juntamente com suas derivadas. Além de explorarmos também o método dos fatores integrantes.

Essas equações podem ser escritas na forma-padrão:

$$\frac{dy}{dx} + P(x)y = Q(x), \quad (2.17)$$

onde  $P$  e  $Q$  são funções dadas da variável independente  $x$ . As equações diferenciais lineares de primeira ordem são chamadas de “lineares” porque a função desconhecida e suas derivadas aparecem de forma linear na equação, ou seja,  $y$  e suas derivadas são multiplicadas por funções de  $x$ , mas não estão envolvidas em termos de produto ou divisão entre si. Se  $Q$  não é a função zero, a equação (2.17) é chamada linear não homogênea, ao passo que  $Q(x) = 0$  para todo  $x$ , ela é chamada linear homogênea. Observe que toda equação linear homogênea é separável e tem a função zero como a única solução singular. Procuramos uma solução de (2.17) em um intervalo  $I$  no qual as funções  $P$  e  $Q$  são contínuas.

### 2.3.1 Método de resolução pelos fatores integrantes

Para resolver a equação (2.17), consideremos uma função auxiliar  $\mu(x)$  de forma que ao multiplicar (2.17) por essa função a equação obtida é uma equação linear com  $P(x) = 0$ . Uma função com esta forma é chamada fator integrante da equação linear.

Seja

$$\mu(x) = e^{\int P(x)dx}.$$

Vamos mostrar que  $\mu(x) = e^{\int P(x)dx}$  é um fator integrante da equação (2.17). Observe em primeiro lugar que

$$\frac{d\mu}{dx} = e^{\int P(x)dx} \frac{d}{dx} \left( \int P(x)dx \right) = e^{\int P(x)dx} P(x) = \mu(x)P(x). \quad (2.18)$$

Assim, se multiplicarmos (2.17) por  $\mu(x)$ , obtemos

$$\mu(x) \frac{dy}{dx} + \mu(x)P(x)y = \mu(x)Q(x) \quad (2.19)$$

mas por (2.18),  $\mu(x)P(x) = \frac{d\mu}{dx}$ , então (2.19) pode ser reescrita como

$$\mu(x)\frac{dy}{dx} + \frac{d\mu}{dx}y = \mu(x)Q(x).$$

Mas o lado esquerdo dessa equação é a derivada de um produto, o que faz com que ela possa ser reescrita na forma

$$\frac{d}{dx}(\mu(x)y(x)) = \mu(x)Q(x). \quad (2.20)$$

A equação (2.20) é uma equação do tipo

$$\frac{dY}{dx} = f(x)$$

que pode facilmente ser resolvida integrando-se dos dois lados. Em que  $Y(x) = \mu(x)y(x)$  e  $f(x) = \mu(x)Q(x)$ . Assim, a solução de (2.20) é dada por

$$\mu(x)y(x) = \int \mu(x)Q(x)dx + C.$$

Como  $\mu(x) \neq 0$ , para todo  $x \in \mathbb{R}$ , dividindo-se a equação anterior por  $\mu(x)$  obtemos que a solução geral de (2.17) é dada por

$$y(x) = \frac{1}{\mu(x)} \left( \int \mu(x)Q(x)dx + C \right).$$

**Exemplo 2.** Considere a equação

$$\frac{dy}{dx} + \frac{2}{x}y = x.$$

O fator integrante é

$$\mu(x) = e^{\int \frac{2}{x}dx} = e^{2\ln x} = e^{\ln x^2} = x^2.$$

Multiplicando-se as equações acima por  $\mu(x)$  obtemos

$$x^2\frac{dy}{dx} + 2xy = x^3.$$

O lado esquerdo é igual à derivada do produto  $x^2y(x)$ . Logo a equação acima é equivalente a

$$\frac{d}{dx}(x^2y(x)) = x^3.$$

Integrando os dois lados

$$\int \frac{d}{dx}(x^2y(x))dx = \int x^3dx$$

obtemos

$$x^2y(x) = \frac{t^4}{4} + C.$$

Explicitando  $y(x)$  temos que a solução geral da ED é

$$y(x) = \frac{t^4}{4} + \frac{C}{x^2}.$$

Para avançarmos neste estudo, apenas os métodos de resolução de EDO's apresentados até este ponto são suficientes, entretanto, existem diversos métodos de resolução de EDO's que não são o foco deste trabalho.

## 2.4 Sistemas de equações diferenciais lineares de primeira ordem

A partir desta seção, denotaremos a variável independente por  $t$  e as variáveis dependentes, sendo funções de  $t$ , por  $x_1, x_2, \dots, x_n$  e suas derivadas em relação a  $t$  por  $x'$  de modo a nos aproximarmos das notações usadas pelos autores em que nos baseamos para a escrita deste trabalho. Antes de abordarmos o tema de sistemas de equações diferenciais de primeira ordem, é de suma importância lembrarmos duas definições necessárias da teoria de álgebra matricial.

### 2.4.1 Sistema de equações lineares

Um conjunto de  $m$  equações lineares com  $n$  variáveis é um conjunto de equações do tipo

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

que pode ser reduzido na forma matricial

$$Ax = b$$

onde

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \text{ e } x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

**Definição 6** (Derivada de uma matriz de funções). *Se  $A(t) = (a_{ij}(t))_{m \times n}$  for uma matriz cujos elementos são funções diferenciáveis em um intervalo comum, então*

$$\frac{dA}{dt} = \left[ \frac{d}{dt} a_{ij} \right]_{m \times n}.$$

**Definição 7** (Integral de uma matriz de funções). *Se  $A(t) = (a_{ij}(t))_{m \times n}$  for uma matriz cujas entradas são funções contínuas em um intervalo comum contendo  $t$  e  $t_0$ , então*

$$\int_{t_0}^t A(s) ds = \left[ \int_{t_0}^t a_{ij}(s) ds \right]_{m \times n}.$$

### 2.4.2 Autovalores e autovetores

De acordo com Boyce e DiPrima (2010), a equação

$$Ax = y \tag{2.21}$$

pode ser vista como uma transformação linear que leva (ou transforma) um vetor dado  $x$  em um novo vetor  $y$  com  $x$  e  $y$  in  $R^n$ . Vetores transformados em múltiplos de si são importantes

em várias aplicações, como na epidemiologia matemática. Para encontrar tais vetores, fazemos  $y = \lambda x$ , onde  $\lambda$  é um fator escalar de proporcionalidade, e procuramos soluções das equações

$$Ax = \lambda x, \quad (2.22)$$

ou

$$(A - \lambda I)x = 0. \quad (2.23)$$

A última equação tem soluções não nulas, se e somente se,  $\lambda$  for escolhido de modo que

$$\det(A - \lambda I) = 0. \quad (2.24)$$

Os valores de  $\lambda$  que satisfazem a (2.24) são chamados de autovalores da matriz  $A$ , e as soluções não nulas correspondentes das (2.22) ou (2.23), obtidas usando-se tal valor de  $\lambda$ , são chamadas de autovetores correspondentes, ou associados, àquele autovalor.

### 2.4.3 Outros resultados importantes de álgebra matricial

Considere o sistema de equações lineares

$$\begin{cases} x'_1(t) = a_{11}(t)x_1(t) + \dots + a_{1n}(t)x_n(t) + f_1(t) \\ \vdots \\ x'_n(t) = a_{n1}(t)x_1(t) + \dots + a_{nn}(t)x_n(t) + f_n(t) \end{cases} \quad (2.25)$$

que pode ser escrito na forma de uma equação diferencial matricial

$$\begin{bmatrix} x'_1(t) \\ \vdots \\ x'_n(t) \end{bmatrix} = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} + \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

ou

$$X'(t) = A(t)X(t) + F(t),$$

em que

$$A'(t) = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix}, X(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \text{ e } F(t) = \begin{bmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{bmatrix}.$$

Quando  $F(t) = 0$ , é denotado sistemas de equações lineares homogêneas, e caso contrário, não homogêneo.

Para sistemas lineares é válido o seguinte teorema sobre existência e unicidade de soluções

**Teorema 2** (Existência e unicidade). *Considere o problema de valor inicial*

$$\begin{cases} X'(t) = A(t)X(t) + F(t) \\ X(t_0) = X^{(0)} \end{cases} \quad (2.26)$$

*suponha que  $a_{ij}(t), f_i(t)$  sejam funções contínuas num intervalo  $I = [a, b]$  contendo  $t_0$ . Então (2.26) tem uma solução única no intervalo  $I$ .*

A demonstração desse teorema pode ser encontrada em (SANTOS, 2011).

Para sistemas de equações lineares homogêneos, isto é, da forma com  $F(t) = 0$ ,

$$X'(t) = A(t)X(t), \quad (2.27)$$

é válido o princípio da superposição que diz que se  $X_1(t)$  e  $X_2(t)$  são soluções de (2.27), então

$$X(t) = \alpha X_1(t) + \beta X_2(t) \quad (2.28)$$

também o é, para todas as constantes  $\alpha$  e  $\beta$ . Uma expressão da forma (2.28) é chamada de combinação linear de  $X_1(t)$  e  $X_2(t)$ . Vamos verificar que realmente  $X(t)$  dado por (2.28) é solução de (2.27).

$$\begin{aligned} X'(t) &= \alpha X_1'(t) + \beta X_2'(t) = \alpha A(t)X_1(t) + \beta A(t)X_2(t) \\ &= A(t)(\alpha X_1(t) + \beta X_2(t)) = A(t)X(t), \end{aligned}$$

pois como  $X_1(t)$  e  $X_2(t)$  são soluções de (2.27), então  $X_1'(t) = A(t)X_1(t)$  e  $X_2'(t) = A(t)X_2(t)$ . Provamos o seguinte teorema.

**Teorema 3** (Princípio da superposição). *Se  $X_1(t)$  e  $X_2(t)$  são soluções do sistema homogêneo*

$$X'(t) = A(t)X(t) \quad (2.29)$$

*então,  $X(t) = \alpha X_1(t) + \beta X_2(t)$ , também é solução para quaisquer que sejam as constantes  $\alpha$  e  $\beta$ .*

*Demonstração.* Derivando a combinação linear, para quaisquer  $\alpha$  e  $\beta$  temos

$$X'(t) = \alpha X_1'(t) + \beta X_2'(t). \quad (2.30)$$

Como hipótese  $X_1$  e  $X_2$  são soluções do sistema linear, temos que

$$X_1'(t) = A(t)X_1(t) \text{ e } X_2'(t) = A(t)X_2(t)$$

substituindo  $X_1'(t)$  e  $X_2'(t)$  na equação (2.30), obtemos

$$X'(t) = \alpha A(t)X_1(t) + \beta A(t)X_2(t).$$

Segue que,

$$X(t) = A(t)(\alpha X_1(t) + \beta X_2(t)).$$

Logo, concluímos que a combinação linear de soluções de (2.29) também é solução do sistema homogêneo para qualquer  $\alpha$  e  $\beta$ .  $\square$

Esta demonstração foi retirada de (LIMA, 2018).

**Teorema 4.** *Sejam  $X_1(t), \dots, X_n(t)$ , soluções do sistema  $X'(t) = A(t)X$  tais que*

$$\det [X_1(0) \dots X_n(0)] \neq 0.$$

*Então para toda condição inicial  $x^{(0)} \in \mathbb{R}$  o problema de valor inicial*

$$\begin{cases} X'(t) = A(t)X(t) \\ X(0) = X^{(0)} \end{cases}$$

*tem solução única da forma*

$$X(t) = c_1 X_1(t) + \dots + c_n X_n(t). \quad (2.31)$$

**Definição 8.** (a) *Sejam  $X_1 : \mathbb{R} \rightarrow \mathbb{R}^n, \dots, X_n : \mathbb{R} \rightarrow \mathbb{R}^n$  funções vetoriais. O determinante*

$$W [X_1, \dots, X_n] (t_0) = \det [X_1(t_0), \dots, X_n(t_0)]$$

*é chamado wronskiano das funções vetoriais  $X_1(t), \dots, X_n(t)$  em  $t_0$ .*

(b) *Se  $n$  soluções  $X_1(t), \dots, X_n(t)$  do sistema  $X' = A(t)'X$ , são tais que o seu wronskiano difere de zero em um ponto,  $t_0$  dizemos que elas são soluções fundamentais do sistema homogêneo*

$$X' = A(t)'X$$

(c) *Se  $X_1(t), \dots, X_n(t)$  são soluções fundamentais do sistema  $X' = A'(t)X$ , então a família de soluções*

$$X(t) = c_1 X_1(t) + \dots + c_n X_n(t), \quad (2.32)$$

*para constantes  $c_1, \dots, c_n$  é chamada solução geral de  $X' = A(t)'X$ .*

Assim, para encontrar a solução geral de um sistema homogêneo  $X' = A(t)X$ , precisamos encontrar  $n$  soluções fundamentais, ou seja, soluções  $X_1(t), \dots, X_n(t)$  tais que em um ponto  $t_0 \in \mathbb{R}$

$$W [X_1, \dots, X_n] (t_0) = \det [X_1(t_0), \dots, X_n(t_0)] \neq 0.$$

## 2.5 Estabilidade e instabilidade

Denomina-se ponto crítico ou ponto de equilíbrio do sistema, todo ponto  $(x_0, y_0)$ , tal que  $F(x_0, y_0) = G(x_0, y_0) = (0, 0)$ , ou seja, se  $(x_0, y_0)$  é ponto crítico, então

$$\begin{cases} x' = F(x_0, y_0) = 0 \\ y' = G(x_0, y_0) = 0 \end{cases}, \quad (2.33)$$

e as variações são nulas no ponto crítico.

Destaca-se que se  $(x_0, y_0)$  é um ponto crítico de (2.33), então

$$\begin{cases} x(t) = x_0 \\ y(t) = y_0 \end{cases}$$

é uma solução do sistema autônomo não linear

$$\begin{cases} x' = F(x, y) \\ y' = G(x, y) \end{cases}, \quad (2.34)$$

chamada solução estacionária ou equilíbrio. A partir disso, (MARCOLINO; SILVEIRA, 2016) traz alguns resultados importantes para a análise do modelo SIR do Capítulo 04.

**Definição 9.** Um ponto de equilíbrio  $(x_0, y_0)$  do sistema (2.34) é estável se, dado  $\varepsilon > 0$ , é possível determinar um valor  $\delta > 0$ , dependente apenas de  $\varepsilon$ , tal que se  $\|(x, y) - (x_0, y_0)\| < \delta$ , então a solução  $(x(t), y(t))$  existe para todo  $t > 0$  e  $\|(x(t), y(t)) - (x_0, y_0)\| < \delta$ .

Em outras palavras, o ponto crítico é estável se as soluções que começam perto dele não se afastam muito no futuro.

**Definição 10.** Um ponto de equilíbrio  $(x_0, y_0)$  do sistema (2.34) é assintoticamente estável se for estável e, além disso, tem-se

$$\lim_{t \rightarrow \infty} (x(t), y(t)) = (x_0, y_0). \quad (2.35)$$

Dessa forma, uma solução que começa próxima ao ponto crítico  $(x_0, y_0)$ , não somente permanece próxima a esse ponto como também tende a ele quando  $t \rightarrow \infty$ .

**Definição 11.** Um ponto crítico não estável é dito instável.

## 2.6 Aproximações lineares de sistemas não lineares

O teorema de Hartman-Grobman é uma ferramenta matemática amplamente utilizada no estudo de sistemas dinâmicos não lineares, incluindo modelos epidemiológicos como o modelo SIR. Esse teorema estabelece uma relação entre a estabilidade linear de um ponto de equilíbrio

em um sistema dinâmico não linear e a estabilidade do mesmo ponto em seu sistema linearizado correspondente.

De acordo com Perko (1991), o Teorema de Hartman-Grobman mostra que, sob certas hipóteses, numa região próxima do ponto de equilíbrio, o sistema não linear

$$x' = f(x) \quad (2.36)$$

possui uma estrutura qualitativa equivalente ao sistema linear

$$x' = Ax. \quad (2.37)$$

onde  $A = Df(p_0)$ . e  $p_0$  é uma singularidade.

Abaixo, enunciamos o Teorema de Hartman-Grobman, extraído de (PERKO, 1991).

**Teorema 5** (Teorema de Hartman-Grobman). *Seja  $E$  um subconjunto aberto de  $\mathbb{R}^n$  contendo a origem,  $f \in C^1(E)$ , e seja  $\phi_t$  o fluxo do sistema não linear (2.36). Suponha que  $f(0) = 0$  e que a matriz  $Df(0)$  autovalores com parte real nula, então existe um homeomorfismo  $H$  de um conjunto aberto  $U$  contendo a origem em um conjunto aberto  $V$  contendo a origem tal que para cada  $p_0 \in U$ , existe um intervalo aberto  $I_0 \subset \mathbb{R}$  contendo zero, tal que,  $\forall x_0 \in U$  e  $t \in I_0$*

$$H \circ \phi_t(p_0) = e^{At} H(p_0); \quad (2.38)$$

*isto é,  $H$  leva trajetórias de (2.36) próximas em trajetórias de (2.37) próximas à origem, preservando a orientação.*

A demonstração deste teorema pode ser encontrada em (PERKO, 1991).

### 3 EPIDEMIOLOGIA MATEMÁTICA

Antes de discutir a epidemiologia matemática é preciso fazer uma breve reflexão sobre a definição de modelo matemático, que pode ser encontrado em cada obra de forma diferente. Nesta pesquisa usaremos o que é proposto em (BASSANEZI, 2002), que estabelece Modelagem matemática como um conjunto de símbolos e relações matemáticas que representam de alguma forma o objeto estudado. O modelo matemático é fundamental porque nos fornece uma linguagem precisa e inequívoca para expressar nossas ideias claramente. Esta prática se define como modelagem matemática. A modelagem matemática desempenha um papel fundamental na epidemiologia matemática. A epidemiologia matemática é uma disciplina que utiliza modelos matemáticos para descrever e compreender a propagação de doenças em uma população. Esses modelos ajudam os pesquisadores a prever o curso da epidemia, avaliar o impacto de diferentes estratégias de intervenção e tomar decisões informadas para o controle e prevenção de doenças.

A modelagem matemática na epidemiologia envolve a formulação de equações e modelos que descrevem as interações entre os indivíduos em uma população, considerando fatores como a taxa de infecção, a taxa de recuperação, a transmissibilidade da doença e a suscetibilidade da população (que serão melhor apresentados mais adiante). Com base nos estudos de Ensino e Aprendizagem com modelagem matemática (BASSANEZI, 2002) e Epidemiologia e saúde: fundamentos, métodos, aplicações (FILHO; BARRETO, 2011), a proposta deste capítulo é estabelecer alguns conceitos e definições envolvidos na escolha e na interpretação de tais modelos descritos neste trabalho do ponto de vista da epidemiologia matemática.

#### 3.1 Conceitos básicos

Conceitos básicos são úteis quando se trata de modelagem matemática. Partindo da importância de definir e melhorar a linguagem do que se propõe a ser estudado, definiremos os principais conceitos de epidemiologia matemática para este trabalho. Essas definições foram retiradas de (FILHO; BARRETO, 2011, pg.283)'.

**Variável:** Grandeza de interesse, por exemplo, número de indivíduos infectados, que se modifica como uma função do tempo.

**Dinâmica:** Padrão de modificações observadas em um sistema como uma função do tempo.

**Parâmetro:** Grandeza auxiliar, geralmente um termo que alimenta um modelo, e que permanece constante no tempo. Em modelos estocásticos representa uma grandeza a ser estimada.

**Parcimônia:** Princípio que argumenta ou sugere que modelos ou explicações mais simples devam ter precedência sobre as mais complexas quando ambas são igualmente plausíveis frente às observações empíricas.

**Modelo dinâmico:** Descreve a evolução temporal de um sistema biológico.

**Modelo determinístico:** Modelo dinâmico em que a evolução temporal do sistema biológico é totalmente previsível pelas equações que o descrevem e os valores atribuídos aos parâmetros.

**Modelo estocástico:** Modelo dinâmico em que a evolução temporal do sistema biológico sofre a influência de eventos aleatórios e eventos futuros são descritos por meio de afirmações probabilísticas.

**Lei de ação das massas:** Conceito com sua origem no estudo da cinética química, que postula a taxa de formação de um composto é proporcional às concentrações dos reagentes. Em modelos de dinâmica populacional, este conceito implica que o contato entre suscetíveis e infecciosos é traduzido matematicamente pelo produto do contingente de cada uma dessas classes.

**Número básico de reprodução da doença:** descreve o número de casos secundários gerados a partir do caso primário no início do processo epidêmico.

### 3.1.1 Lei de ação das massas

Nesta seção, nos comentaremos um pouco mais sobre o conceito de lei de ação de massas definido na seção anterior, com sua origem na cinética química, ao tratar de doenças infecciosas de transmissão direta (um indivíduo infectado contamina outro indivíduo, que por ventura contamina outros).

Filho e Barreto (2011) nos trazem esta interação pode ser modelada matematicamente como

$$\lambda(t) = \beta y(t). \quad (3.1)$$

Onde:  $\lambda(t)$  representa a taxa de novos casos,  $\beta$  é um parâmetro que descreve a taxa de transmissão da doença, e  $y(t)$  é o produto das densidades de pessoas infectadas e suscetíveis.

"Esta equação diz que casos são gerados conforme pessoas infectadas e suscetíveis encontram-se a uma taxa  $\beta$  vezes o produto de suas densidades."(CODEÇO; COELHO, 2012, pg. 113).

## 3.2 Convergência entre a biologia realista e a complexidade matemática

Filho e Barreto (2011) destacam que a Lei de Ação de Massas serve apenas como aproximação do problema, pois nem sempre os contágios se dão de forma uniforme. Existe a possibilidade de criação de modelos com interações não homogêneas, já que a Equação (3.1) é pensada a partir de uma realidade matemática onde os indivíduos pertencem a mesma classe, faixa etária, e que se relacionam com seus semelhantes. Para modelos não homogêneos, por exemplo, de faixa etárias diferentes, quando o fator de estruturação é a idade, a matriz de contato

passa a ser expressa da forma  $\beta(a, a')$ , onde  $a$  e  $a'$  são indivíduos de faixa etárias diferentes. Modelos como o desse exemplo requerem utilização de ferramentas mais complexas, como os sistemas de equações diferenciais parciais, algoritmos de solução numérica, além de serem mais resistentes ao tratamento analítico, o que não é o foco deste trabalho.

Modelos determinísticos são simulados numericamente e descrevem a média do processo infeccioso na população. Modelos estocásticos consideram a aleatoriedade da transmissão da infecção e as diferenças individuais, permitindo previsões com estimativas de incertezas. Em baixa incidência, a natureza aleatória é mais evidente. Modelos baseados em indivíduos são estocásticos e detalhados, mas exigem muitos recursos computacionais e têm limitações analíticas (FILHO; BARRETO, 2011). Além disso, pelo fato dos modelos detalhados serem geralmente difíceis ou impossíveis de resolver analiticamente, sua utilidade para fins teóricos acaba sendo limitada (BRAUER *et al.*, 2008).

Sabendo da implicação dos modelos detalhados e de suas limitações teóricas, na próxima seção discutiremos brevemente a ideia de modelos compartimentais.

### 3.3 Modelos compartimentais determinísticos

Para (BRAUER *et al.*, 2008, pg. 22): "uma epidemia, que atua em uma escala temporal curta, pode ser descrita como um surto súbito de uma doença que infecta uma parcela substancial da população em uma região", os autores ainda destacam que esses surtos podem ocorrer repetidamente ao longo do tempo, com intervalos de vários anos, e podem diminuir em gravidade à medida que a imunidade se desenvolve nas populações.

Existem diversos modelos compartimentais como SEIR (suscetíveis, expostos, infectados e recuperados), SI (suscetíveis e infectados), SIS (suscetíveis, infectados e suscetíveis), SIRS (suscetíveis, infectados, removidos e suscetíveis) e SEIS (suscetíveis, expostos, infectados, e suscetíveis). Nesta pesquisa, para fins de estudo e análise do modelo proposto a ser estudado no trabalho, vamos considerar doença infecciosa para o qual ninguém tem imunidade, e uma população onde todos nascem suscetíveis a esta doença, a partir disso, consideraremos três grupos, suscetíveis (aqueles que não foram infectados pela doença), os infectados (que estão doentes e podem infectar novas pessoas) e os removidos (que não estão mais infectados e criaram imunidade após ser contaminados pela doença ou morreram), esses grupos são chamados compartimentos (BRAUER *et al.*, 2008).

Figura 1 – Diagrama do modelo de compartimentos SIR, onde  $S$  representa os suscetíveis,  $I$  os infectados e  $R$  os removidos.



(Fonte: o autor).

Estabelecida a ideia primitiva desse modelo compartimental, no próximo capítulo estudaremos o modelo SIR de forma detalhada.

## 4 O MODELO SIR

Neste capítulo, estudaremos de forma mais detalhada o modelo SIR onde uma população de tamanho  $N(t)$  composta por três categorias de indivíduos: os suscetíveis, ou seja, os não infectados, livres da doença, indicados por  $S(t)$ ; os infectados e infecciosos, indicados por  $I(t)$ ; e os removidos, ou seja, recuperados ou mortos, indicados por  $R(t)$ . Neste capítulo, usaremos a notação  $S'$  para denotar a derivada  $\frac{dS}{dt}$  e apenas  $S$  para  $S(t)$ . A respeito dos removidos, é importante destacar uma informação que não foi discutida na seção (3.2), podemos afirmar que

*a remoção é realizada através do isolamento do resto da população ou através de imunização contra a infecção ou através da recuperação da doença com total imunidade contra a reinfecção ou pela morte causada pela doença. Essas caracterizações de membros removidos diferem de uma perspectiva epidemiológica, mas muitas vezes são equivalentes de um ponto de vista da modelagem que considera apenas o estado de um indivíduo em relação à doença. (BRAUER et al., 2008, pg. 25)*

portanto, o compartimento removidos contem muitas camadas de interpretação quando é feita uma análise do ponto de vista epidemiológico, entretanto nossa preocupação neste trabalho ficará apenas no ponto de vista da modelagem.

De acordo com (BRAUER et al., 2008), o modelo proposto por Kermack e McKendrick, o qual é o ponto de partida para o nosso estudo de modelos epidêmicos, é dado pelo sistema de equações diferenciais não lineares

$$\begin{cases} S' = -\beta SI \\ I' = \beta SI - \gamma I \\ R' = \gamma I, \end{cases} \quad (4.1)$$

Consideraremos algumas premissas:

(1) Um membro médio da população faz contato suficiente para transmitir a infecção para  $\beta N$  outras pessoas por unidade de tempo, onde  $N$  será a população total. Este ponto sugere que a taxa de propagação de uma infecção é diretamente proporcional ao número de contatos feitos por um indivíduo médio e ao tamanho total da população. O parâmetro  $\beta$  representa a taxa de transmissão ou potencial de transmissão da doença, e  $\beta N$  indica que o número de contatos que um indivíduo faz é multiplicado pelo tamanho da população. Essa suposição implica que, à medida que o tamanho da população aumenta, o potencial de transmissão da doença também aumenta.

(2) Os infectados deixam de ser infectados, ou deixam o compartimento dos infectados (taxa de recuperação) a uma taxa  $\gamma I$  por unidade de tempo. Aqui, o termo “infectados” refere-se a indivíduos atualmente infectados com a doença.

(3) A população permanece constante ao longo do surto da doença, não há migração ou movimento de indivíduos para dentro ou para fora da população, exceto pela possibilidade de indivíduos morrerem devido à doença, ou seja, a população é fechada.

De acordo com a primeira premissa, para um contato feito por um indivíduo infectado, a probabilidade desse contato ser com um suscetível é

$$\frac{S}{N}.$$

Portanto, o número de novas infecções por unidade de tempo por indivíduo infectado é

$$\beta N \frac{S}{N},$$

onde  $\beta$  representa a capacidade de transmissão da doença. Isso resulta em uma taxa de novas infecções de

$$\beta N \frac{S}{N} I = \beta SI.$$

Por outro lado, a probabilidade de um contato feito por um suscetível ser com um indivíduo infectado é

$$\frac{I}{N}.$$

Portanto, a taxa de novas infecções por suscetível é

$$\beta N \frac{I}{N},$$

o que resulta em uma taxa de novas infecções de

$$\beta N \frac{I}{N} S = \beta SI.$$

Ambas as abordagens fornecem a mesma taxa de novas infecções, e existem situações em que uma será mais apropriada que a outra. A ideia principal é que a taxa de novas infecções é proporcional ao número de suscetíveis multiplicado pelo número de infectados, multiplicado pelo parâmetro de transmissão  $\beta$ . Em relação ao tamanho da população  $N$ , o modelo assume que a população é constante, a menos que as mortes sejam causadas pela doença. Para uma doença fatal, o tamanho total da população  $N$  é igual à

$$N = S + I,$$

uma vez que todos os infectados acabam morrendo. Para uma doença em que todos os infectados se recuperam com imunidade, o tamanho total da população  $N$  é

$$N = S + I + R.$$

Na terceira premissa podemos entender que a escala de tempo da doença é muito mais rápida do que a escala de tempo de nascimentos e mortes na população. Portanto, os efeitos demográficos, como nascimentos e mortes, são negligenciados no modelo. Isso ocorre porque o foco principal

está na dinâmica do surto epidêmico em si, em vez de considerar o crescimento ou decrescimento populacional devido a outros fatores.

Por último, fazendo uma melhor análise da segunda premissa, que afirma que os infectados deixam a classe de infectados a uma taxa proporcional ao número de infectados. Matematicamente, essa suposição pode ser expressa da seguinte forma: Vamos considerar um recorte de indivíduos que foram todos infectados ao mesmo tempo. Queremos acompanhar o número de indivíduos desse recorte que continuam infectados após um certo período, representado por  $s$ . Denotaremos por  $u(s)$  o número de infectados no tempo  $s$ . Sendo assim, a segunda premissa pode ser interpretada na forma da EDO

$$u'(s) = -\gamma u(s), \quad (4.2)$$

resolvendo (4.2), obtemos

$$u(s) = u(0)e^{-\gamma s}$$

onde  $u(0)$  é o número inicial de infectados no tempo  $s = 0$ .

A solução  $u(s)$  fornece o número de infectados no tempo  $s$  após terem sido infectados, assumindo uma taxa constante de remoção  $\gamma$ . Essa equação mostra que o número de infectados diminui exponencialmente ao longo do tempo, em função da taxa de remoção  $\gamma$  e do tempo decorrido  $s$ . É importante observar que essa formulação matemática assume uma taxa de remoção  $\gamma$  constante e não captura especificamente os mecanismos epidemiológicos subjacentes à recuperação ou remoção dos infectados. Ela serve como uma representação matemática básica como já foi discutido no Capítulo 3. Consequentemente, podemos afirmar que a segunda premissa assume que a duração da infecção segue uma distribuição exponencial, com média

$$\int_0^{\infty} e^{-\gamma s} ds = \frac{1}{\gamma}$$

onde  $\gamma$  é a taxa de remoção. Isso implica que o comprimento do período infectante é distribuído exponencialmente. As suposições de uma taxa de contatos proporcional ao tamanho da população  $N$  com uma constante de proporcionalidade  $\beta$  e uma taxa de recuperação distribuída exponencialmente são simplificações irreais. Embora modelos mais gerais possam ser construídos e analisados, nosso objetivo aqui é mostrar o que pode ser deduzido a partir de modelos extremamente simples. No nosso modelo, o valor de  $R$  é determinado assim que  $S$  e  $I$  são conhecidos, então podemos remover a equação de  $R$  do modelo, deixando um sistema de duas equações

$$\begin{cases} S' = -\beta SI \\ I' = (\beta S - \gamma)I \end{cases}$$

De acordo com (BRAUER *et al.*, 2008), não é possível resolver este sistema analiticamente, mas podemos aprender muito sobre o comportamento das suas soluções por meio de uma abordagem qualitativa.

Observamos que o modelo faz sentido apenas enquanto  $S$  e  $I$  permanecerem não negativos. Se algum deles atingir zero, consideramos que o sistema terminou. Observamos que  $S' < 0$  para todos os  $t$  e  $I' > 0$  se e somente se  $S > \frac{\gamma}{\beta}$ . Portanto,  $I$  aumenta enquanto  $S > \frac{\gamma}{\beta}$ , mas como  $S$  diminui para todos os  $t > 0$ ,  $I$  eventualmente diminui e se aproxima de zero. Se  $S(0) < \frac{\gamma}{\beta}$ ,  $I$  diminui para zero (sem epidemia), enquanto se  $S(0) > \frac{\gamma}{\beta}$ ,  $I$  primeiro aumenta até um máximo quando  $S = \frac{\gamma}{\beta}$  e depois diminui para zero (epidemia). A quantidade  $\frac{\beta S(0)}{\gamma}$  é um valor limite chamado número básico de reprodução e denotado por  $R_0$ , que foi apresentado de forma breve no capítulo anterior. Na próxima seção estudaremos melhor o conceito do  $R_0$  e suas consequências.

#### 4.1 Número Básico de reprodução ( $R_0$ ) e o limiar epidêmico

O número básico de reprodução, como já foi apresentado, descreve o número de casos gerados a partir do caso primário no início do processo epidêmico (FILHO; BARRETO, 2011), que também pode ser caracterizado como o instante inicial da epidemia, isto é, a razão entre a taxa de formação de novos casos,  $\beta IS$  e a taxa de recuperação  $\gamma I$ , portanto

$$R_0 = \frac{\beta IS}{\gamma I} \quad (4.3)$$

Como no instante inicial toda população é considerada suscetível, então podemos simplificar a Equação (4.3) como

$$R_0 = \frac{\beta}{\gamma} \quad (4.4)$$

Se  $R_0 > 1$ , o número de indivíduos infectados aumentará a cada geração e a doença se espalhará causando uma epidemia, se  $R_0 < 1$ , não haverá epidemia. (CODEÇO; COELHO, 2012).

Note que  $R_0$  é uma medida inicial e pode não refletir totalmente a transmissão real da doença em estágios avançados. Em geral,  $R_0$  é uma estimativa útil para entender como uma doença pode se espalhar em uma população, mas outros fatores também devem ser considerados para uma análise completa. (BRAUER *et al.*, 2008).

#### 4.2 Análise do modelo SIR

A análise do modelo SIR envolve o estudo das propriedades matemáticas e comportamentais do sistema de equações diferenciais que descrevem a propagação de uma doença em uma população. Essa análise pode ser feita utilizando diversas técnicas e abordagens, nesta seção, optamos por uma análise qualitativa dos sistemas lineares de equações diferenciais e como principal referência para este capítulo utilizaremos a dissertação Modelos Matemáticos Epidemiológicos da Tuberculose em Codó-MA de Marcolino e Silveira (2016).

No contexto do modelo SIR, o teorema de Hartman-Grobman apresentado em 2.6 é aplicado para analisar a estabilidade dos pontos de equilíbrio do sistema e obter informações sobre o comportamento dinâmico do modelo. De uma forma mais simplista, o teorema estabelece que, se o ponto de equilíbrio linearizado for estável, então o ponto de equilíbrio correspondente

no sistema não linear também será estável, e vice-versa. Uma vez linearizado o sistema SIR, o teorema de Hartman-Grobman permite efetuar uma análise mais simples da estabilidade do ponto de equilíbrio linearizado. Essa informação sobre a estabilidade do ponto de equilíbrio é fundamental no estudo epidemiológico do modelo SIR. Ela pode ser usada para determinar se uma doença se espalhará ou não em uma população, se ocorrerá uma epidemia ou apenas casos esporádicos, e para entender como diferentes fatores, como taxas de transmissão e recuperação, afetam o comportamento do sistema.

Tendo em conta os conhecimentos teóricos apresentados na sessão 2.6, consideremos o sistema (4.1) com as condições iniciais  $S(0) \geq 0$ ,  $I(0) \geq 0$ ,  $R_0 = 0$  e  $N = S_0 + I_0 + R_0$ .

Marcolino e Silveira (2017) nos dizem que a solução para esse modelo para o número de infectados em função do tempo é dado por

$$I(t) = I_0 e^{(\beta S - \lambda)t}, \quad (4.5)$$

onde  $S_0$ ,  $I_0$  e  $R_0$  representam respectivamente as proporções iniciais de indivíduos suscetíveis, infectados e recuperados no instante  $t = 0$ . É importante ressaltar que essa solução é uma simplificação e não considera a interação entre as diferentes populações (suscetíveis, infectados e removidos) ou a dinâmica temporal completa do sistema. No entanto, essa expressão pode servir para fornecer uma estimativa inicial do número de infectados em um modelo SIR simplificado, considerando as proporções iniciais de indivíduos suscetíveis e infectados.

### 4.3 Um exemplo simples de aplicação do modelo SIR em linguagem R

Nesta seção nos propomos a apresentar uma breve simulação simples do modelo SIR usando a linguagem R. A escolha dessa abordagem para este exemplo se baseia puramente no fato de que o uso da linguagem R me foi estimulado por diversas vezes na graduação. A possibilidade de reinfecção e a variabilidade na resposta imunológica da Covid-19 não serão consideradas nesse exemplo. As principais referências nesta seção serão os artigos Modelagem matemática da Covid-19 usando SIR de (ABREU, 2020) e Modelo SIR em Epidemiologia de (TAVARES, 2017).

Abreu (2020) nos apresenta alguns parâmetros necessários para a aplicação computacional do modelo SIR:

1)  $S$ : quantidade de pessoas no grupo de pessoas suscetíveis.  $S$  representa uma fração da população e, para haver propagação da infecção, deve ser maior que zero. Assim,  $0 < S \leq 1$ . Neste exemplo consideraremos a população do último censo realizado pelo Instituto Brasileiro de Geografia e Estatística — IBGE para o município de Aparecida de Goiânia, no estado de Goiás, que traz uma população de 527.550 habitantes (IBGE, 2022), ou seja  $S(0) = 1$ . Estes habitantes não são vacinados.

2)  $I$ : quantidade de pessoas no grupo dos infectados. Para termos propagação da infecção,

devemos ter  $0 < I < 1$ . No início considera-se  $I \approx 0$  (mas não igual a zero). Neste exemplo usaremos  $I(0) = \frac{1}{527550}$ .

3)  $R$ : quantidade de pessoas no grupo de pessoas removidas. Devemos ter  $0 \leq R < 1$ . Neste exemplo, usaremos  $R = 0$ , considerando que no instante inicial  $t_0 = 0$ , não haverá nenhum removido.

4)  $\gamma$ : taxa de recuperação. Para determinar  $\gamma$  é necessário saber o tempo médio de recuperação de uma pessoa infectada. Neste exemplo usaremos como base assumida por (ABREU, 2020) para a Covid-19,  $\gamma = \frac{1}{10}$ .

5)  $\beta$ : taxa de infecção. Para este exemplo consideraremos  $\beta = 0,2969558$  com base no estudo de Abreu (2020). Sabendo que  $R_0 = \frac{\beta}{\gamma}$ , temos que  $R_0 \approx 3$  para a Covid-19.

A partir destes parâmetros é possível realizar uma simulação em R usando os seguintes pacotes: **ggplot2**, usado para gerar os gráficos, **deSolve**, para resolver o sistema de equações e **reshape2** para reestruturar os dados. O próximo passo é definir os parâmetros estabelecidos anteriormente.

Figura 2 – Linha do código de definição de parâmetros em R.

```
5
6 # PARAMETROS DO MODELO
7
8 initial_state_values=c(S=527550,I=1,R=0)
9 parameters=c(gamma=1/10,beta=0.3)
10
```

(Fonte: o autor).

O tempo estabelecido para este exemplo foi de 365 dias, sendo contados dia-a-dia como mostra a linha de código da Figura (3).

Figura 3 – Tempo em que ocorreu a pandemia do Exemplo, do dia 1 ao dia 200, contado um-a-um.

```
10
11 # TEMPO
12
13 time=seq(from=1,to=200,by=1)
14
```

(Fonte: o autor)

Em seguida foi armazenado na função **sir model** todos os dados do modelo SIR, incluindo os parâmetros fixados na Figura (2).

Figura 4 – Descrição do modelo SIR e os parâmetros do exemplo.

```

15 |
16 > sir_model<- function(time,state, parameters) {
17 >   with(as.list(c(state,parameters)),{
18     N<- S+I+R
19     lambda<-beta*(I/N)
20     dS<- -lambda*S
21     dI<- lambda*S-gamma*I
22     dR<- gamma*I
23     return(list(c(dS,dI,dR)))
24 >   })
25 > }
26

```

(Fonte: o autor).

O próximo passo foi resolver a equação diferencial usando o seguinte código da Figura (5):

Figura 5 – Código para resolver as equações do sistema.

```

33 output<-as.data.frame(ode(y=initial_state_values,func = sir_model,parms=paramet
34
35
36 out_long=melt(output,id="time")
37

```

(Fonte: o autor).

Em seguida, foi gerado o gráfico da Figura(7), que, a partir de uma análise qualitativa, analisando os picos, taxas de crescimento, duração da pandemia entre outros resultados.

Figura 6 – Código usado para plotar o gráfico.

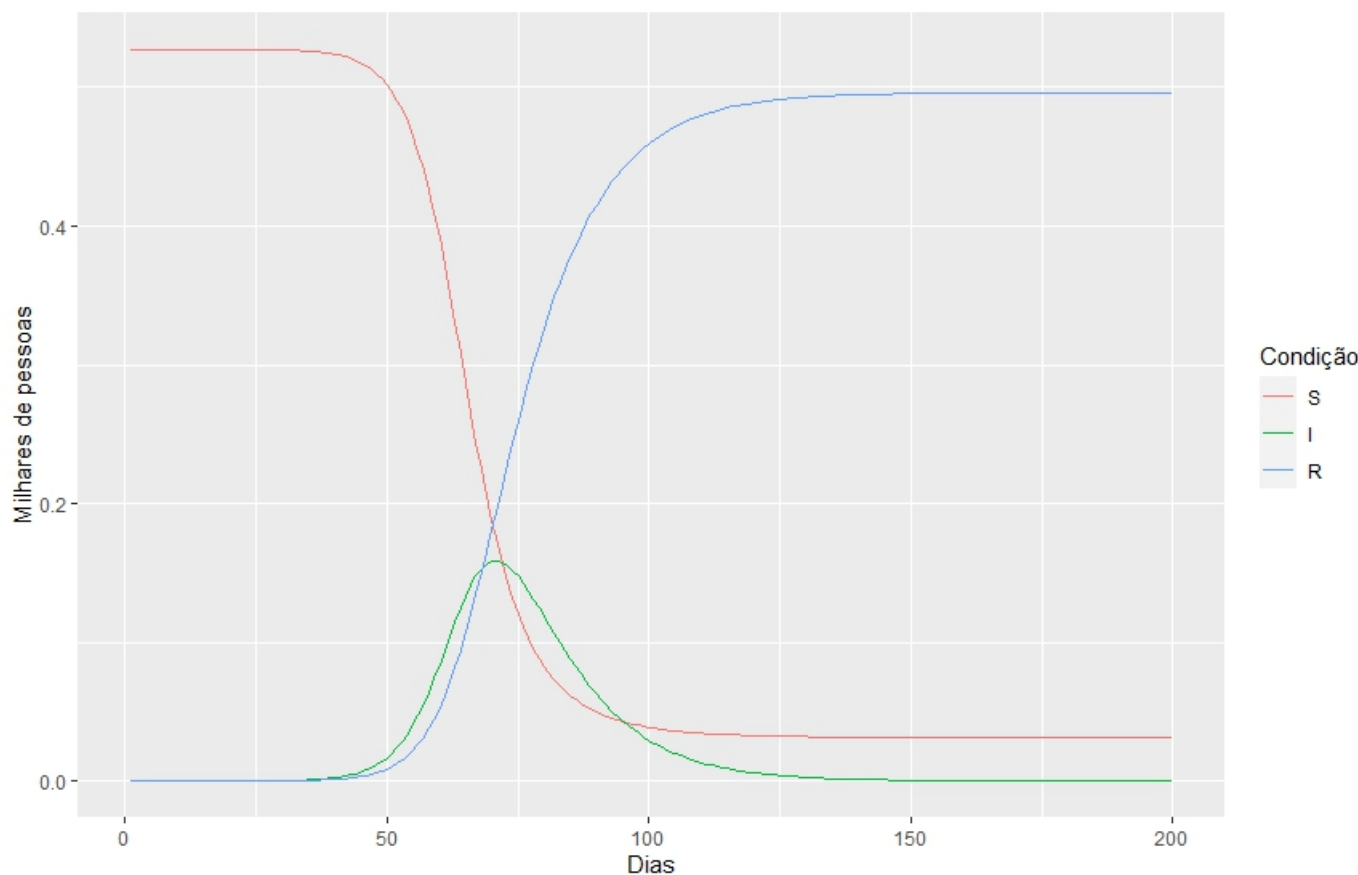
```

38 ggplot(data = out_long,
39       aes(x = time, y = value/1000000, colour = variable, group = variable)) +
40   geom_line() +xlab("Dias")+ylab("Milhares de pessoas")+scale_color_discrete(na

```

(Fonte: o autor).

Figura 7 – A solução numérica obtida pela linguagem de programação R para um surto de Covid-19 na cidade de Aparecida de Goiânia-Goiás, com uma taxa de infecção  $\beta = 0,2969558$ . Note que pouco antes do quinquagésimo dia o número de infectados começa a aumentar e atinge seu pico por volta do septuagésimo quinto dia.



(Fonte: o autor).

## 5 CONSIDERAÇÕES FINAIS

Ao decorrer desta pesquisa, percebemos que além dos conhecimentos teóricos de equações diferenciais, álgebra linear e matricial, outros assuntos correlacionados a estes que não são comuns nas ementas dos cursos de matemática também foram necessários. É importante destacar que durante o processo de pesquisa, além do modelo SIR, também foram estudados outros modelos que apresentam variadas possibilidades de pesquisa. Nossos resultados indicam que o modelo SIR consegue capturar várias dinâmicas observadas em epidemias reais, como o aumento inicial de casos, o pico da epidemia e a subsequente diminuição do número de infectados. Além disso, o modelo nos permite examinar a importância de parâmetros-chave, como a taxa de transmissão e a taxa de recuperação, no controle da propagação da doença.

No entanto, é importante destacar as limitações do modelo SIR. Primeiro, ele pressupõe uma população homogênea e não considera fatores demográficos, geográficos e comportamentais que podem influenciar a propagação da doença. Além disso, o modelo não considera a imunidade adquirida ou a possibilidade de variações no patógeno ao longo do tempo.

Apesar dessas limitações, o modelo SIR continua sendo uma ferramenta valiosa para os epidemiologistas, fornecendo *insights* importantes sobre a dinâmica das doenças infecciosas. É importante ressaltar que esse modelo deve ser usado em conjunto com outras abordagens, como dados epidemiológicos reais e técnicas estatísticas mais avançadas, para uma compreensão mais completa e precisa do problema. Além disso, pesquisas adicionais podem explorar a aplicação de outros modelos epidemiológicos, como o modelo SEIR, que considera uma fase latente da doença.

## REFERÊNCIAS

- ABREU, M. O. R. de. Modelagem matemática da covid-19 usando sir. **Jornal Eletrônico de Ensino e Pesquisa de Matemática - UEM**, v. 4, n. 1, p. 27–48, 2020.
- BASSANEZI, R. C. **Ensino-aprendizagem com modelagem matemática: uma nova estratégia**. São Paulo, SP: Editora Contexto, 2002.
- BOYCE, W. E.; DIPRIMA, R. C. **Equações diferenciais elementares e problemas de valores de contorno**. [S.l.]: LTC Rio de Janeiro, 2010. v. 10.
- BRAUER, F. *et al.* **Mathematical models in epidemiology**. New York: Springer, 2019. v. 32.
- BRAUER, F. *et al.* **Mathematical epidemiology**. [S.l.]: Springer, 2008. v. 1945.
- ÇENGEL, Y. A.; III, W. J. P. **Equações diferenciais**. [S.l.]: AMGH Editora, 2014.
- CODEÇO, C. T.; COELHO, F. C. Modelagem de doenças transmissíveis. **Oecologia Australis**, v. 16, n. 1, p. 110–116, 2012.
- DIACU, F. **Introdução a Equações Diferenciais: teoria e aplicações**. Rio de Janeiro, RJ: LTC, 2004.
- DIETZ, K. Introduction to mckendrick (1926) applications of mathematics to medical problems. In: \_\_\_\_\_. **Breakthroughs in Statistics**. New York, NY: Springer New York, 1997. p. 17–57. ISBN 978-1-4612-0667-5. Disponível em: <[https://doi.org/10.1007/978-1-4612-0667-5\\_2](https://doi.org/10.1007/978-1-4612-0667-5_2)>.
- FARIAS, A. V. Um estudo da modelagem epidemiológica sir usando conceitos de derivadas de ordem inteira e fracionária. **Monografia, UFRG**, 2017.
- FILHO, N. d. A.; BARRETO, M. L. Epidemiologia & saúde: fundamentos, métodos, aplicações. In: **Epidemiologia & saúde: fundamentos, métodos, aplicações**. Rio de Janeiro, RJ: Editora Guanabara Koogon, 2011. p. 699–699.
- GRAY, J. **Change and variations: a history of differential equations to 1900**. Milton Keynes: Springer Nature, 2021.
- IBGE. Instituto brasileiro de geografia e estatística - censo 2022. **IBGE**, Rio de Janeiro, RJ, 2022. Disponível em: <<https://censo2022.ibge.gov.br/panorama/>>.
- KRETZSCHMAR, M.; WALLINGA, J. Mathematical models in infectious disease epidemiology. **Modern infectious disease epidemiology: Concepts, methods, mathematical models, and public health**, Springer, p. 209–221, 2010.
- LIMA, J. I. d. **Um estudo sobre equações diferenciais ordinárias aplicado á epidemiologia**. — Universidade Federal de Campina Grande, Cuité, 2018.
- MARCOLINO, R. dos S.; SILVEIRA, G. P. Um modelo epidemiologico da tuberculose em codo maranhao. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 5, n. 1, 2017.

MARCOLINO, R. S.; SILVEIRA, G. P. Modelos matemáticos epidemiológicos da tuberculose em codó-ma. **Dis-sertação de Mestrado, Unicamp**, Campinas, SP, 2016.

PERKO, L. **Differential equations and dynamical systems**. New York: Springer Verlag, 1991.

SANTOS, R. J. Introduçãoas equações diferenciais ordinárias. **Imprensa Universitaria da UFMG**, Belo Horizonte, 2011.

SMITH, G. D.; KUH, D. Commentary: William Ogilvy Kermack and the childhood origins of adult health and disease. **International Journal of Epidemiology**, v. 30, n. 4, p. 696–703, 08 2001. ISSN 0300-5771. Disponível em: <<https://doi.org/10.1093/ije/30.4.696>>.

TAVARES, J. N. Modelo sir em epidemiologia. **Revista de Ciência Elementar - Universidade do Porto**, v. 5, n. 2, 2017.

ZILL, D. G. **Equações diferenciais com aplicações em modelagem**. [S.l.]: Cengage Learning, 2016.