



UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
DEPARTAMENTO DE MATEMÁTICA



RAFAEL GUSTAVO ALVES

O Modelo *Time-dependent SIR* e a Covid-19:
estudos preliminares, teoria e simulações numéricas

Ouro Preto, Minas Gerais, Brasil
Março, 2023

RAFAEL GUSTAVO ALVES

**O Modelo *Time-dependent SIR* e a Covid-19:
estudos preliminares, teoria e simulações numéricas**

Trabalho de Conclusão de Curso apresentado ao Curso de Licenciatura em Matemática da Universidade Federal de Ouro Preto como parte dos requisitos para a obtenção do Grau de Licenciado em Matemática.

Orientador: Prof. Dr. Felipe Rogério Pimentel

Ouro Preto, Minas Gerais, Brasil
Março, 2023

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

A474o Alves, Rafael Gustavo.

O modelo Time-dependent SIR e a Covid-19 [manuscrito]: estudos preliminares, teoria e simulações numéricas. / Rafael Gustavo Alves. - 2023.

49 f.: il.: color., gráf..

Orientador: Prof. Dr. Felipe Rogério Pimentel.

Monografia (Licenciatura). Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Graduação em Matemática .

1. Modelos Epidemiológicos. 2. modelagem matemática. 3. epidemiologia. 4. Covid-19. I. Pimentel, Felipe Rogério. II. Universidade Federal de Ouro Preto. III. Título.

CDU 51

Bibliotecário(a) Responsável: Elton Ferreira de Mattos - SIAPE: 1.754.007



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
COLEGIADO DO CURSO DE LICENCIATURA EM
MATEMÁTICA



FOLHA DE APROVAÇÃO

Rafael Gustavo Alves

Monografia apresentada ao Curso de Licenciatura em Matemática da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Licenciado em Matemática

Aprovada em 30 de março de 2023

Membros da banca

Dr. Felipe Rogério Pimentel - Orientador - Universidade Federal de Ouro Preto
Dr. Edson Alberto Coayla Teran - Universidade Federal da Bahia
Ma. Monique Rafaella Anuniação de Oliveira - Universidade Federal de Ouro Preto
Dr. Ricardo Tavares - Universidade Federal de Ouro Preto

Felipe Rogério Pimentel, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 05/04/2023



Documento assinado eletronicamente por **Felipe Rogério Pimentel, PROFESSOR DE MAGISTERIO SUPERIOR**, em 05/04/2023, às 13:47, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0483644** e o código CRC **8F847E83**.

RESUMO

A Covid-19 é uma doença infectocontagiosa causada pelo vírus SARS-CoV-2, que teve seu primeiro caso registrado na cidade de Wuhan (China) e, após sua rápida proliferação em nível global, recebeu o *status* de pandemia, determinado pela Organização Mundial da Saúde (OMS) em março de 2020. Nesse cenário pandêmico, diversos estudos e modelos matemáticos epidemiológicos começaram a ser propostos para a doença. Seguindo essa linha, apresentamos uma variação do modelo epidemiológico clássico SIR (proposto por Kermack e McKendrick), conhecida como *Time-dependent SIR*, para descrever a dinâmica de transmissão do vírus SARS-CoV-2. Temos como objetivo central o uso desse modelo adaptado para obtermos previsões das taxas de transmissão e de recuperação, além das previsões dos números de infectados e recuperados da doença no estado de Minas Gerais, num contexto inicial da pandemia, em que fatores como vacina, reinfeção, dentre outros, ainda eram desconhecidos. Além do tema central, apresentamos também importantes conteúdos da literatura, a título de pré-requisitos, tais como modelos matemáticos bidimensionais descritos por sistemas de equações diferenciais, regressão linear e mínimos quadrados lineares.

Palavras-chave: *Time-dependent SIR*, modelagem matemática, epidemiologia, Covid-19, *ridge regression*, otimização com regularização

ABSTRACT

Covid-19 is an infectious disease caused by the SARS-CoV-2 virus which first case was registered in Wuhan (China). After its fast proliferation at a global level it received the pandemic status, declared by the World Health Organization (WHO) in March 2020. Several studies and mathematical models in epidemiology have been proposed for the disease in this pandemic scenario. Following this line of study, we present a variation of the classical SIR model (proposed by Kermack and McKendrick), known as Time-dependent SIR, to describe the transmission dynamics of the SARS-CoV-2 virus. Our main objective is to use this adapted model to obtain the predictions of transmission and recovery rates, in addition to the predictions of the numbers of infected and recovered from the disease in the state of Minas Gerais, in an initial context of the pandemic, when factors such as vaccine, reinfection, among others, were still unknown. In addition to the main goal, we also present important literature contents, as prerequisites, such as two-dimensional mathematical model described by systems of differential equations, linear regression and, linear least squares.

Keywords: *Time-dependent SIR*, mathematical modelling, epidemiology, Covid-19, *ridge regression*, regularized optimization

SUMÁRIO

	INTRODUÇÃO	9
	O CONTEXTO DA PANDEMIA DA Covid-19 NO ENSINO BÁSICO	11
1	REVISÃO SOBRE SISTEMAS DINÂMICOS	13
1.1	Sistemas Bidimensionais	13
1.2	Sistemas autônomos lineares bidimensionais	20
1.3	Sistemas autônomos não lineares bidimensionais	22
1.4	Modelos epidemiológicos	24
2	INTRODUÇÃO À INTERPOLAÇÃO POLINOMIAL, REGRESSÃO LINEAR E QUADRADOS MÍNIMOS LINEARES	27
2.1	Interpolação Polinomial	27
2.2	Quadrados Mínimos Lineares	29
2.2.1	Regressão Linear para o caso unidimensional	30
2.3	Regressão Linear para Dimensões Maiores	31
3	MODELO EPIDEMIOLÓGICO	34
3.1	Ridge Regression	35
3.2	Ridge Regression Aplicado ao Modelo Epidemiológico	39
4	RESULTADOS NUMÉRICOS	42
4.1	Materiais	42
4.2	Previsões	43
	CONSIDERAÇÕES FINAIS	47
	Referências	48

LISTA DE FIGURAS

Figura 1 – Exemplo do Campo $X(x, y) = \left(\frac{y}{x^2 + y^2}, -\frac{x}{x^2 + y^2} \right)$	14
Figura 2 – Órbita de um sistema autônomo bidimensional para o Exemplo 1	15
Figura 3 – Órbita de um sistema autônomo bidimensional para o Exemplo 2	16
Figura 4 – Órbita de um sistema autônomo bidimensional para o Exemplo 3	17
Figura 5 – Trajetórias de fase na vizinhança de ciclos limite: (a) um ciclo limite estável, (b) um ciclo limite instável, (c) e (d) ciclos limite semi-estáveis	18
Figura 6 – Trajetórias do sistema (1.4). O círculo $r = 1$ é um ciclo limite.	19
Figura 7 – Trajetórias do sistema (1.4) com as assíntotas radiais.	19
Figura 8 – Plano de fase do sistema linear bidimensional (1.9), em que $\lambda_1 = 3$ e $\lambda_2 = -1$	21
Figura 9 – Plano de fase do sistema linear bidimensional (1.10), em que $\lambda_1 = -1$ e $\lambda_2 = -3$	22
Figura 10 – Plano de fase do sistema linear bidimensional (1.11), em que $\lambda_{1,2} = -1 \pm 2i$	23
Figura 11 – Plano de fase para o sistema (1.16)	25
Figura 12 – Plano de fase para o sistema (1.16). A curva em preto é uma separatriz.	25
Figura 13 – Exemplo de polinômio interpolador.	27
Figura 14 –	43
Figura 15 –	43
Figura 16 –	44
Figura 17 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de maio de 2020.	44
Figura 18 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de junho de 2020.	45
Figura 19 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de julho de 2020.	45
Figura 20 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de agosto de 2020.	46

INTRODUÇÃO

A Covid-19 é uma doença infectocontagiosa causada pelo vírus SARS-CoV-2 (*Severe acute respiratory syndrome coronavirus 2*), e sua transmissão costuma ocorrer pelo ar ou por contato pessoal com secreções contaminadas. Entre os principais sintomas da doença estão a febre, tosse seca e dificuldade para respirar [9].

O primeiro caso registrado da doença foi na cidade de Wuhan, na província de Hubei, China, em dezembro de 2019. Com a rápida proliferação do vírus em diversos países do mundo, em 11 de março de 2020, a Organização Mundial da Saúde (OMS) declarou estado de pandemia da Covid-19 [23].

Do advento da pandemia até os dias de hoje, foram levantadas diversas perguntas acerca do vírus e da doença: como e o quanto o vírus se prolifera e contamina outros indivíduos, quais medidas diminuem a proliferação do vírus, quando o número de infectados atingiria o seu pico, como identificar os indivíduos infectados, a existência de infectados assintomáticos, a possibilidade de reinfeção pelo vírus, quais fármacos podem ser usados contra a doença, entre muitas outras.

Para auxiliar na busca de respostas para algumas dessas perguntas, diversos estudos e modelos matemáticos começaram a ser propostos. No caso dos modelos matemáticos epidemiológicos, que descrevem a dinâmica de transmissão da Covid-19, quando eles reproduzem essa dinâmica com eficácia, podem fornecer parâmetros de suporte e auxílio no planejamento e implementação de ações no combate à doença.

Dos trabalhos clássicos sobre os modelos epidemiológicos, o modelo SIR, proposto por Kermack e McKendrick (1927), em que a população hospedeira é subdividida em três classes: a classe S dos indivíduos suscetíveis à doença, a classe I dos indivíduos infectados e a classe R dos indivíduos recuperados, foi escolhido para estudo neste trabalho. Nosso objetivo principal será o uso de uma adaptação do modelo clássico SIR para obtermos previsões das taxas de transmissão e de recuperação, além das previsões dos números de infectados e recuperados da Covid-19 no estado de Minas Gerais (Brasil), num contexto inicial da pandemia, em que fatores como vacina, reinfeção, dentre outros, ainda eram desconhecidos. Tal método adaptado é conhecido como *Time-dependent SIR* e está amplamente detalhado em [8].

Iniciamos este trabalho com um breve relato relacionando o contexto da pandemia da Covid-19 com a educação básica e a formação de professores. No Capítulo 1, apresentamos uma revisão da teoria básica para sistemas autônomos bidimensionais com foco na análise qualitativa. Através de alguns exemplos, enunciamos alguns resultados importantes da

teoria de Sistemas Dinâmicos que podem ser vistos com mais detalhes em [11].

No Capítulo 2, estabelecemos métodos numéricos utilizados para encontrar uma função que melhor se aproxima a um conjunto de dados conhecidos, como a interpolação polinomial e o método dos quadrados mínimos lineares. O leitor que tiver interesse nesse assunto poderá consultar as referências [14, 19].

No Capítulo 3, descrevemos o modelo *Time-dependent* SIR, proposto por pesquisadores taiwaneses [8], que será utilizado para a modelagem da Covid-19. Além disso, apresentamos o método *Ridge Regression*, uma variação do método dos quadrados mínimos lineares, que será a ferramenta utilizada para a resolução de problemas de otimização necessários para a obtenção das previsões das taxas de transmissão e de recuperação da Covid-19. Para a obtenção da solução ótima utilizamos o Teorema da Decomposição em Valores Singulares (SVD). Por fim, apresentamos um algoritmo que descreve o método de previsão utilizado neste trabalho. Detalhes sobre o método Ridge Regression podem ser vistos em [19], e variações do *Time-dependent* SIR poderão ser encontradas em [10].

No Capítulo 4, apresentamos os resultados numéricos obtidos para as previsões dos números de infectados e recuperados da Covid-19, para o estado de Minas Gerais, em diferentes períodos ao longo do ano de 2020. As simulações numéricas foram realizadas através da implementação do algoritmo apresentado no Capítulo 3 em linguagem Octave [17]. Finalizamos o panorama deste trabalho com as considerações finais.

O Contexto da Pandemia da Covid-19 no Ensino Básico

A pandemia da Covid-19 produziu efeitos não apenas de ordem médica e epidemiológica em escala mundial, mas também impactou as áreas social, econômica, política, cultural, histórica e educacional [16]. Uma pesquisa realizada em dezembro de 2021 pelo Instituto DataSenado, com a finalidade de levantar informações sobre o impacto da pandemia na educação brasileira, mostrou que, além das mudanças na rotina do dia a dia e na sociabilidade das crianças e adolescentes, que impactam negativamente o processo de aprendizagem dos(as) alunos(as), a principal percepção levantada pelos participantes da pesquisa é que os anos de 2020 e 2021 foram anos perdidos para a educação, o que resulta em consequências graves a longo prazo [15].

Com o retorno das aulas presenciais, surgiu um novo cenário educacional, que busca recuperar o prejuízo no ensino causado pela pandemia que, além de provocar estragos, gerou uma quantidade enorme de dados e informações que foram (e ainda são) divulgados nas mais diversas mídias. Esse novo cenário propõe aos professores o desafio de estimular o desenvolvimento de uma análise crítica do estudante em relação a essas informações propagadas [7]. Esse desafio vai ao encontro da Lei de Diretrizes e Bases da Educação (LDB, Art. 35, inciso III), que estabelece como finalidade do Ensino Médio “o aprimoramento do educando como pessoa humana, incluindo a formação ética e o desenvolvimento da autonomia intelectual e do pensamento crítico” [5]. Nesse contexto, o artigo [7] questiona como o professor de Matemática pode beneficiar-se da diversidade de informações matemáticas, estatísticas e de contexto sobre a pandemia e utilizá-las em sala de aula, proporcionando aos alunos a capacidade de desenvolver senso crítico, compreender o mundo e tomar decisões conscientes.

De acordo com [7], o contexto da pandemia nos permite “articular conceitos de Matemática e Estatística com outros advindos da área de Ciências Humanas e Sociais, Ciências Naturais, Linguagens e temas transversais como saúde, economia, meio ambiente, trabalho, entre outros”, em consonância com os Parâmetros Curriculares Nacionais (PCN, [6]) e a Base Nacional Comum Curricular (BNCC, pag. 9, [4]), que apresenta 10 competências gerais que devem ser desenvolvidas durante a Educação Básica, dentre as quais destacamos a sétima competência:

Argumentar com base em fatos, dados e informações confiáveis, para formular, negociar e defender ideias, pontos de vista e decisões comuns que respeitem e promovam os direitos humanos, a consciência socioambiental e o consumo responsável em âmbito local, regional e global, com

posicionamento ético em relação ao cuidado de si mesmo, dos outros e do planeta.

Assim, como forma de atender algumas das demandas da BNCC, usando a pandemia da Covid-19 como pano de fundo, Cazorla propõe em [7] seis *Casos de Ensino*¹, dentre os quais destacamos: (i) Números absolutos e relativos, (ii) Informações oficiais sobre a pandemia e (iii) Crescimento exponencial em contexto, que oferecem ao professor ferramentas para abordar e propor reflexões acerca de conceitos e conteúdos matemáticos e estatísticos na pandemia.

Desse modo, a prática docente neste ambiente de pós pandemia exige do professor um amplo conhecimento envolvendo outras áreas na direção de uma organização interdisciplinar. Por isso, a interdisciplinaridade torna-se ferramenta fundamental para a quebra de paradigmas na formação do docente, permitindo visões múltiplas sobre a forma de pensar um objeto de estudo, transformando a sala de aula em um espaço propício para debates e reflexões que extrapolam o contexto da área de formação do professor [28].

¹ “Casos de Ensino são narrativas que mostram situações complexas vividas por professores durante sua atividade docente. Trazem exemplos de como é possível lidar com determinadas situações e, mais que isso, explicitam dilemas e conflitos enfrentados por docentes ao lidar com o ensino e com seus alunos.” (Extraído do portal < <http://www.portaldosprofessores.ufscar.br/caso.jsp> >.)

1 Revisão sobre Sistemas Dinâmicos

1.1 Sistemas Bidimensionais

Nos últimos anos, muito se falou sobre o estudo de modelos epidemiológicos. Esses estudos investigam a transmissão de uma doença e de seus agentes causadores e, neles, a modelagem matemática é um dos aspectos essenciais, se propondo a descrever uma situação real como um problema matemático que será objeto de estudo e, futuramente, será resolvido interpretando os resultados obtidos com estes modelos numa linguagem mais próxima da realidade [2, 11].

Neste capítulo, discutiremos sobre os modelos matemáticos bidimensionais, que são descritos por sistemas de equações diferenciais da forma:

$$\begin{cases} \frac{dx}{dt} = F(x, y) \\ \frac{dy}{dt} = G(x, y), \end{cases} \quad (1.1)$$

em que x e y são funções na variável t , com t pertencente a um intervalo aberto da reta. Consideraremos o sistema (1.1) como um *sistema autônomo*, que é caracterizado desta forma quando as funções F e G não dependem explicitamente da variável independente t , mas apenas das variáveis dependentes x e y [11].

Para melhor compreendermos o comportamento desses sistemas, o conhecimento de alguns outros requisitos é essencial. Apresentaremos a seguir, algumas definições necessárias, que podem ser encontradas em [27].

Definição 1.1.1 *Seja Δ um subconjunto aberto do \mathbb{R}^2 . Definimos como campo vetorial uma aplicação $X : \Delta \rightarrow \mathbb{R}^2$, de classe C^k , $k \geq 1$, que associa a cada ponto $\mathbf{x} \in \Delta$, um vetor $X(\mathbf{x})$ do \mathbb{R}^2 .*

Como exemplo, observe a Figura 1, na qual esboçamos alguns vetores do campo $X(x, y) = \left(\frac{y}{x^2 + y^2}, -\frac{x}{x^2 + y^2} \right)$ para $(x, y) \in \{(x, y) \in \mathbb{R}^2; -0,9 \leq x \leq 0,9 \text{ e } -0,9 \leq y \leq 0,9\}$, reduzidos em comprimento através de um fator multiplicativo de $\epsilon = 0,4$.

Definição 1.1.2 *Sejam Δ um subconjunto aberto do \mathbb{R}^2 , $X : \Delta \rightarrow \mathbb{R}^2$, um campo vetorial de classe C^k , $k \geq 1$, I um intervalo da reta e considere a equação diferencial*

$$\mathbf{x}' = X(\mathbf{x}). \quad (1.2)$$

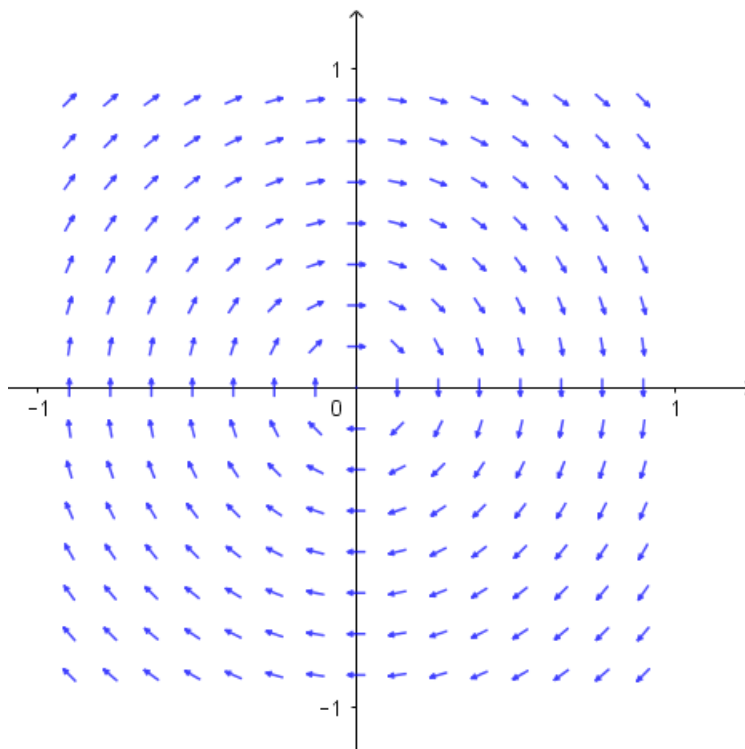


Figura 1 – Exemplo do Campo $X(x, y) = \left(\frac{y}{x^2 + y^2}, -\frac{x}{x^2 + y^2} \right)$

Fonte: o autor.

As aplicações diferenciáveis $\varphi : I \rightarrow \Delta$, que satisfazem a equação (1.2), isto é,

$$\frac{d\varphi}{dt} = X(\varphi(t)), \forall t \in I \quad (1.3)$$

são chamadas órbitas, trajetórias ou curvas integrais do sistema.

Vale destacar que, conforme [11] cita, duas órbitas nunca podem se cruzar em um ponto.

Definição 1.1.3 *Sejam Δ subconjunto aberto do \mathbb{R}^n , $X : \Delta \rightarrow \mathbb{R}^n$ um campo vetorial de classe C^k , $k \geq 1$, $\varphi(t) = \varphi(t, p)$ órbita de X passando por um ponto p , definida no seu intervalo máximo, I_p . Denotaremos os extremos do intervalo por I_p por $\omega_-(p)$ e $\omega_+(p)$, de modo que $I_p = (\omega_-(p), \omega_+(p))$.*

Se $\omega_+(p) = \infty$, definimos:

$$\omega(p) = \{q \in \Delta : \exists t_n, \text{ com } t_n \rightarrow \infty \text{ e } \lim_{n \rightarrow \infty} \varphi(t_n) = q\},$$

Se $\omega_-(p) = -\infty$, definimos:

$$\alpha(p) = \{q \in \Delta : \exists t_n, \text{ com } t_n \rightarrow -\infty \text{ e } \lim_{n \rightarrow \infty} \varphi(t_n) = q\}.$$

em que $\omega(p)$ e $\alpha(p)$ são chamados, respectivamente, de ω -limite e α -limite da órbita $\varphi(t)$.

Definição 1.1.4 Seja $\varphi(t) = \varphi(t, p)$ uma trajetória do campo X , no sentido da definição 1.1.2, que passa pelo ponto p quando $t = 0$, isto é, $\varphi(0, p) = p$. Chamamos o conjunto $\gamma_p = (\varphi(t, p); t \in I_p)$ de órbita de X pelo ponto p .

Definição 1.1.5 Dizemos que uma órbita $\gamma(t)$, dada na equação (1.3), com $\gamma(0) = \gamma_0$, é periódica, de período τ , se $X(\gamma_0) \neq 0$ e existe um valor $t_1 > 0$ tal que $\gamma(t_1) = \gamma(0) = \gamma_0$ e o período τ é o menor dos possíveis valores $t_1 > 0$.

Vejamos alguns exemplos genéricos.

Exemplo 1: A curva em vermelho na Figura 2 abaixo representa uma órbita de um sistema autônomo bidimensional, γ_p , passando por um dado ponto p , que converge para uma órbita fechada simples, γ_q , passando por um ponto q , representada na figura em cor azul. Neste exemplo, estamos considerando o intervalo máximo de γ_p como sendo $I_p = (-\infty, \infty)$. Veremos mais adiante (definição 1.1.6) que γ_q representa um ciclo limite. É fácil ver que γ_q representa o ω -limite da órbita vermelha γ_p . De fato, seja $\tilde{q} \in \gamma_q$, então existe uma sequência t_n , com $t_n \rightarrow \infty$ e $\lim_{n \rightarrow \infty} \varphi(t_n, p) = \tilde{q}$. Para exemplos de pontos \tilde{q} com essa propriedade, considere os pontos q_1 , q_2 e q_3 na figura. Então, pela definição (1.1.3), $\tilde{q} \in \omega(p)$. Dessa forma, temos que o ω -limite, para a órbita γ_p , é dado por $\omega(p) = \gamma_q$. Quanto ao α -limite, neste exemplo conceitual e arbitrário, tal conjunto vai depender do comportamento da órbita γ_p quando $t \rightarrow -\infty$.

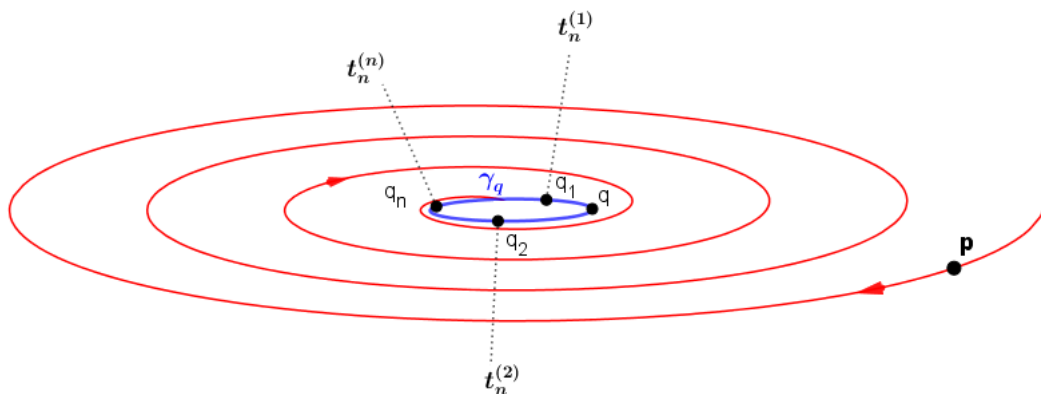


Figura 2 – Órbita de um sistema autônomo bidimensional para o Exemplo 1
Fonte: o autor.

Exemplo 2: A curva em vermelho na Figura 3 abaixo representa um outro exemplo de órbita, γ_p , de um sistema autônomo bidimensional, passando por um dado ponto p . Vemos que γ_p converge para a órbita azul da figura, representada por $\gamma_1 \cup \gamma_2 \cup \{p_0\}$. O conjunto $\{p_0\}$ simboliza a órbita constante $\varphi(t) = p_0, \forall t \in \mathbb{R}$, ou seja, p_0 é um ponto

de equilíbrio¹ do sistema autônomo associado. Aqui estamos novamente considerando o intervalo máximo de γ_p como sendo $I_p = (-\infty, \infty)$. É fácil notar que $\gamma := \gamma_1 \cup \gamma_2 \cup \{p_0\}$ representa o ω -limite da órbita vermelha γ_p , pois, novamente, observa-se geometricamente que existe uma sequência t_n , com $t_n \rightarrow \infty$ e $\lim_{n \rightarrow \infty} \varphi(t_n, p) = \tilde{q}$, para qualquer $\tilde{q} \in \gamma$ (na figura, considere \tilde{q} sendo representado, por exemplo, por q_1 ou q_2 ou p_0).

A visualização geométrica dessa convergência de pontos é simples: seja $r_{\tilde{q}}$ uma reta que passa por \tilde{q} . O ponto \tilde{q} divide essa reta em duas semirretas, $r_{\tilde{q}}^+$ e $r_{\tilde{q}}^-$, ambas com origem em \tilde{q} . Os pontos $\varphi(t_n, p)$ convergindo para \tilde{q} podem ser vistos através da interseção de uma dessas semirretas com a órbita γ_p . Diferentes semirretas geram diferentes conjuntos de pontos $\{\varphi(t_n, p)\}$, que satisfazem $\varphi(t_n, p) \rightarrow \tilde{q}$, para alguma sequência $t_n \rightarrow \infty$. Veja a figura para maiores detalhes.

Da mesma forma que foi feito o estudo para o ponto p , podemos também estabelecer os conjuntos ω -limite e α -limite de cada ponto da órbita γ . Na Figura 3, por exemplo, observa-se que ω -limite e α -limite dos pontos q_1 , q_2 e p_0 são todos iguais a $\{p_0\}$.

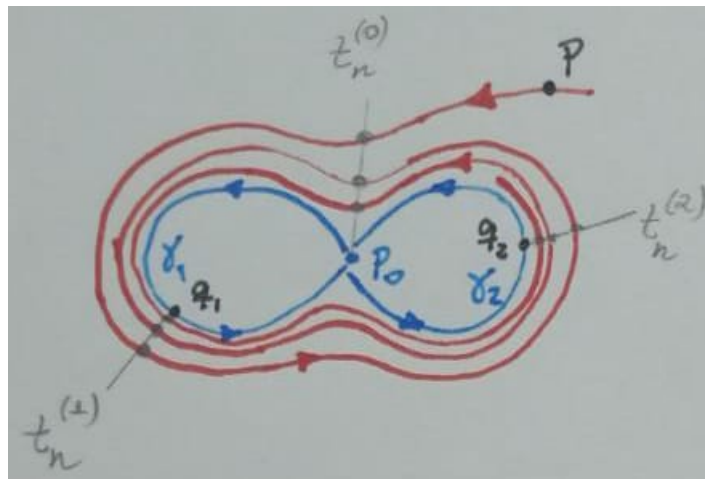


Figura 3 – Órbita de um sistema autônomo bidimensional para o Exemplo 2

Fonte: o autor.

Exemplo 3: Neste exemplo esboçamos mais um modelo de órbita associado a um sistema autônomo (Figura 4). Aqui, notamos que γ_p tem sentido anti-horário e converge assintoticamente para o ponto q . Novamente consideramos $I_p = (-\infty, \infty)$. Diferentemente dos exemplos anteriores, o $\omega(p)$ aqui é identificado com o ponto de equilíbrio q . Além disso, se $\gamma := \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4 \cup \{p_1, p_2, p_3, p_4\}$, é fácil perceber que $\alpha(p)$ é representado por γ , pois, para qualquer $\tilde{q} \in \gamma$, existe uma sequência t_n , com $t_n \rightarrow -\infty$ e $\lim_{n \rightarrow \infty} \varphi(t_n, p) = \tilde{q}$. Veja na figura abaixo, casos particulares para $\tilde{q} = q_4$ e $\tilde{q} = p_3$.

Como nos exemplos anteriores, também podemos estabelecer os conjuntos ω -limite e α -limite para os pontos de γ . Na figura, vemos que, $\omega(q_i) = \{p_i\}$, para todo $i = 1, \dots, 4$ e

¹ Um ponto $p \in \Delta$ é dito ponto de equilíbrio (ou ponto singular, ou ainda ponto crítico) de X se $X(p) = 0$ [11].

$\alpha(q_i) = \{p_{i+1}\}$, para todo $i = 1, \dots, 3$ e $\alpha(q_4) = \{p_1\}$. Para finalizar, $\omega(p_i) = \alpha(p_i) = \{p_i\}$, para todo $i = 1, \dots, 4$.

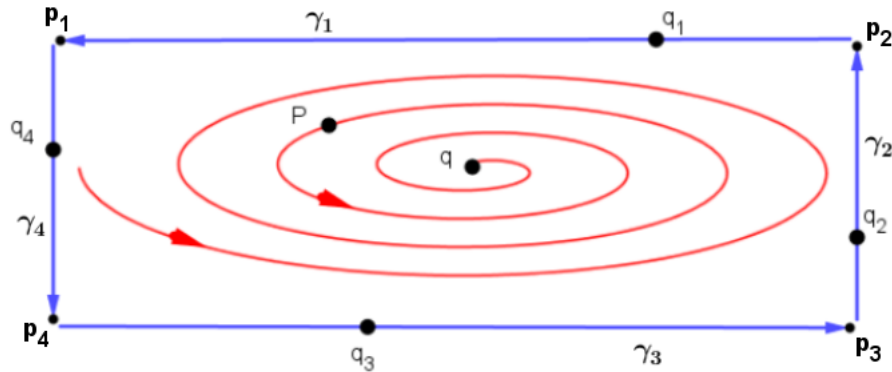


Figura 4 – Órbita de um sistema autônomo bidimensional para o Exemplo 3

Fonte: o autor.

No exemplo 1, dissemos que a órbita em azul, da Figura 2 ilustra um exemplo de um ciclo limite.

Vamos então definir o que é um ciclo-limite. Na definição abaixo $d(p, \gamma)$ representa a distância euclidiana entre o ponto p e o conjunto γ e $int(\gamma)$ e $ext(\gamma)$ representam os conjuntos dos pontos de $\Delta \subset \mathbb{R}^2$ que são interiores a γ e exteriores a γ respectivamente.

Definição 1.1.6 *Sejam Δ um subconjunto aberto do \mathbb{R}^2 e $X : \Delta \rightarrow \mathbb{R}^2$ um campo vetorial de classe C^1 . Uma órbita periódica γ de X chama-se ciclo-limite se $\exists V$ vizinhança de γ tal que γ é a única órbita fechada de X que intercepta V .*

Os ciclos-limite podem ser classificados em:

- Estável: $\lim_{t \rightarrow \infty} d(\varphi(t, q), \gamma) = 0, \forall q \in V$. Neste caso as órbitas se aproximam do ciclo;
- Instável: $\lim_{t \rightarrow -\infty} d(\varphi(t, q), \gamma) = 0, \forall q \in V$. Neste caso as órbitas se afastam do ciclo;
- Semi-estável: $\lim_{t \rightarrow \infty} d(\varphi(t, q), \gamma) = 0, \forall q \in V \cap ext(\gamma)$ e $\lim_{t \rightarrow -\infty} d(\varphi(t, q), \gamma) = 0, \forall q \in V \cap int(\gamma)$, ou o contrário. Neste caso, as órbitas externas se aproximam e as internas se afastam do ciclo, ou vice-versa.

De acordo com a definição (1.1.6), o ciclo-limite, γ_q , do exemplo 1 pode ser estável ou semi-estável.

Para um outro exemplo de ciclo-limite, consideremos o sistema

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + y - x(x^2 + y^2) \\ -x + y - y(x^2 + y^2) \end{pmatrix} \quad (1.4)$$

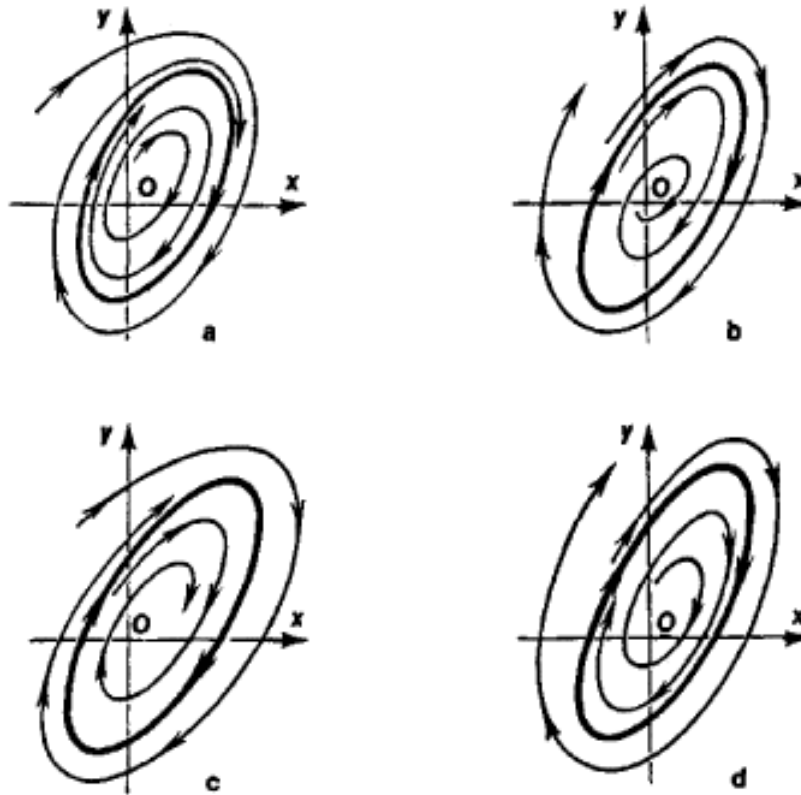


Figura 5 – Trajetórias de fase na vizinhança de ciclos limite: (a) um ciclo limite estável, (b) um ciclo limite instável, (c) e (d) ciclos limite semi-estáveis
Fonte: [24].

apresentado no Exemplo 4.1 em [11].

Conforme [3] e [11], o sistema (1.4) pode ser escrito em termos das coordenadas polares r e θ como

$$\begin{cases} r' = r(1 - r^2), & \text{para } r \neq 0 \\ \theta' = -1 \end{cases} \quad (1.5)$$

Para $r = 0$, obtemos o único ponto de equilíbrio do sistema (1.5) associado à origem, $x = 0$, $y = 0$. Os pontos iniciais (x_0, y_0) , de cada trajetória serão identificados, neste novo sistema de coordenadas por (ρ_0, α_0) , em que $x_0 = \rho_0 \cos(\alpha_0)$ e $y_0 = \rho_0 \sin(\alpha_0)$.

Verifica-se também que $r = 1$, $\theta(t) = -t + t_0$ é uma solução do sistema (1.5), o qual se identifica o círculo unitário C , orientado no sentido horário. Assim, C representa uma órbita fechada periódica do sistema (1.5). Também verifica-se que, para órbitas que se iniciam na região interior ao círculo C (isto é, $\rho_0 < 1$), elas permanecem nessa região e se aproximam deste círculo quando $t \rightarrow \infty$ e se aproximam da origem O do sistema quando $t \rightarrow -\infty$.

Para as órbitas que se iniciam na região exterior ao círculo (isto é, $\rho_0 > 1$), elas permanecem nessa região e também se aproximam de C quanto $t \rightarrow \infty$ e se aproximam de

uma reta (assíntota) radial quando $t \rightarrow -\infty$, conforme as Figuras 6 e 7.

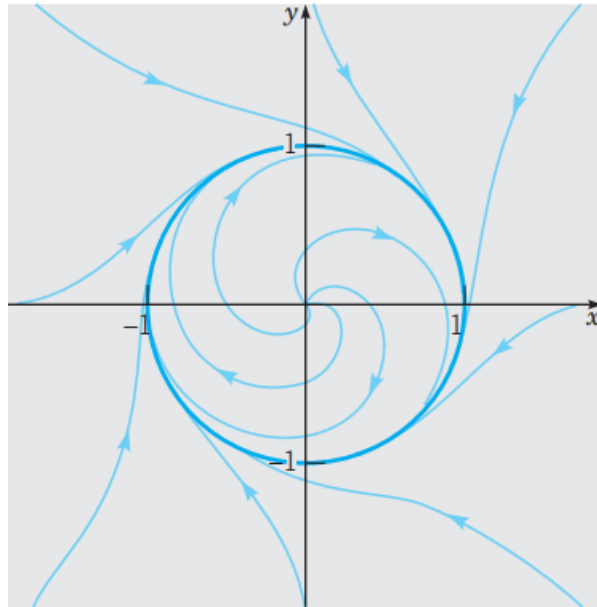


Figura 6 – Trajetórias do sistema (1.4). O círculo $r = 1$ é um ciclo limite.
Fonte: [3]

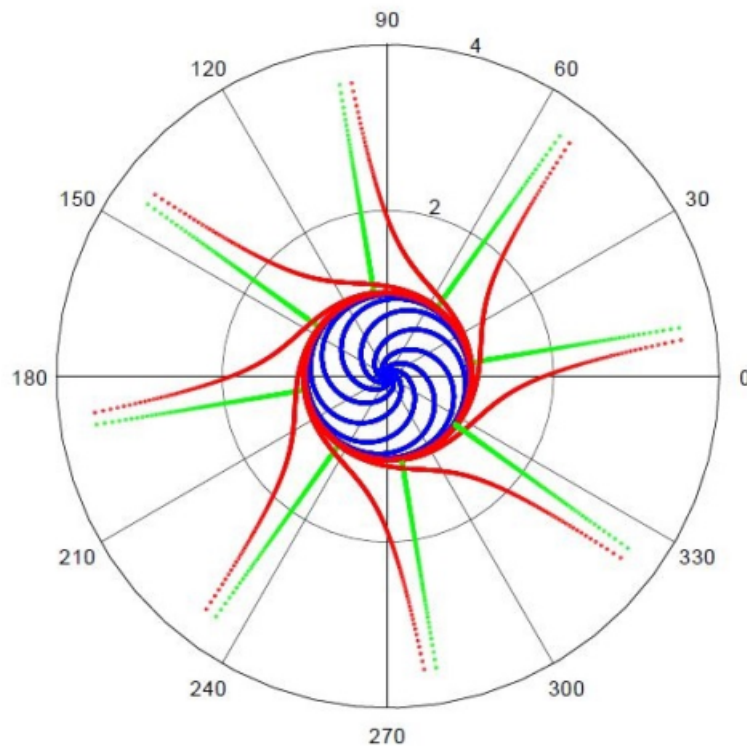


Figura 7 – Trajetórias do sistema (1.4) com as assíntotas radiais.
Fonte: Figura elaborada pelo orientador deste TCC em ambiente Octave em [11]

De acordo com a definição (1.1.6), o círculo unitário C representa um ciclo-limite estável. Por fim, podemos estabelecer os conjuntos ω -limite e α -limite para os pontos interiores, exteriores e sobre o círculo C . Na 6, vemos que, $\alpha(\tilde{q}) = \{O\}$, se $\tilde{q} \in \text{int}(C)$,

$\alpha(\tilde{q}) = \emptyset$, se $\tilde{q} \in \text{ext}(C)$, $\alpha(\tilde{q}) = C$, se $\tilde{q} \in C$. Se \tilde{q} for a origem O , então $\omega(\tilde{q}) = \alpha(\tilde{q}) = \{O\}$. Finalmente, $\omega(\tilde{q}) = C$, para qualquer \tilde{q} diferente da origem.

1.2 Sistemas autônomos lineares bidimensionais

Retornando ao sistema de equações diferenciais (1.1), vamos considerar F e G lineares do tipo

$$F(x, y) = ax + by \quad \text{e} \quad G(x, y) = cx + dy. \quad (1.6)$$

Assim, o sistema (1.1) admite a seguinte forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.7)$$

que é equivalente a

$$\mathbf{x}' = A\mathbf{x} \quad (1.8)$$

em que $\mathbf{x} = (x, y)^T$ e $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Determinando os autovalores, o traço e o determinante da matriz A , dada pelo sistema (1.8), podemos caracterizar os pontos de equilíbrio do sistema, conforme a Tabela 1. Mais detalhes podem ser encontrados em [11].

Tabela 1 – Características de Estabilidade do ponto singular de Sistemas Lineares

Autovalores	$\text{tr}^2(A) - 4\det(A)$	Singularidade	Estabilidade
$0 < \lambda_2 < \lambda_1$	< 0	Nó	Instável
$\lambda_1 < \lambda_2 < 0$	> 0	Nó	Assintoticamente estável
$\lambda_1 < 0 < \lambda_2$	> 0	Ponto de sela	Instável
$\lambda_1 = \lambda_2 > 0$	$= 0$	Nó próprio ou impróprio	Instável
$\lambda_1 = \lambda_2 < 0$	$= 0$	Nó próprio ou impróprio	Assintoticamente estável
$\lambda_{1,2} = a \pm bi, a > 0$	< 0	Ponto espiral	Instável
$\lambda_{1,2} = a \pm bi, a < 0$	< 0	Ponto espiral	Assintoticamente estável
$\lambda_{1,2} = a \pm bi, a = 0$	< 0	Centro	Estável

Fonte: [11]

Vejamos alguns exemplos, apresentados com mais detalhes em [11].

Exemplo 1: Consideremos o sistema linear

$$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \mathbf{x}, \quad (1.9)$$

cujos autovalores da matriz $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$ são $\lambda_1 = 3$ e $\lambda_2 = -1$.

O plano de fase do sistema linear bidimensional para este sistema é mostrado na Figura 8. Neste caso, como os autovalores são reais e têm sinais opostos, a origem, que é um ponto de equilíbrio do sistema, representa um ponto de sela.

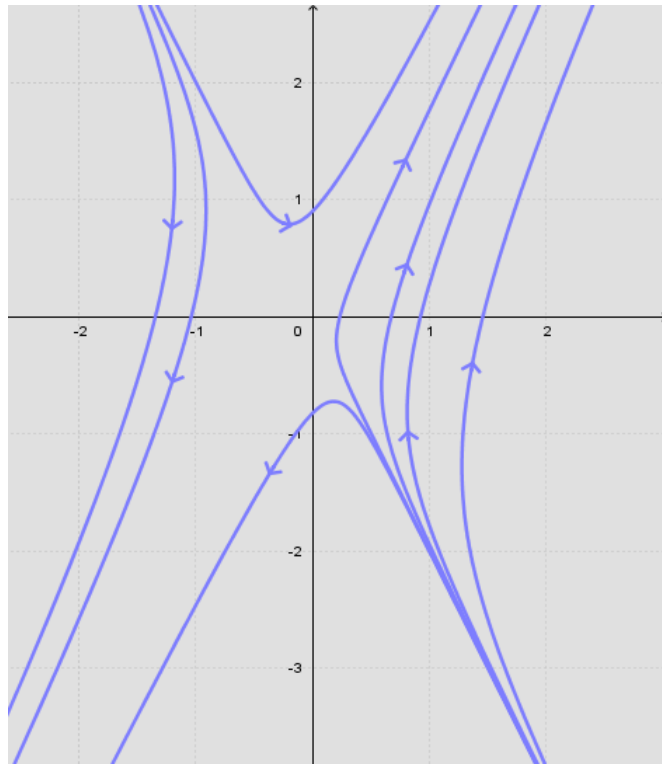


Figura 8 – Plano de fase do sistema linear bidimensional (1.9), em que $\lambda_1 = 3$ e $\lambda_2 = -1$
Fonte: o autor.

Exemplo 2: Consideremos o sistema linear

$$\mathbf{x}' = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \mathbf{x}, \quad (1.10)$$

cujos autovalores da matriz $A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ são $\lambda_1 = -1$ e $\lambda_2 = -3$.

O plano de fase do sistema linear bidimensional para este sistema é mostrado na Figura 9. Neste caso, como os autovalores são reais negativos, a origem, que é um ponto de equilíbrio do sistema, representa um nó assintoticamente estável.

Exemplo 3: Consideremos o sistema linear

$$\mathbf{x}' = \begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix} \mathbf{x}, \quad (1.11)$$

cujos autovalores da matriz $A = \begin{pmatrix} -1 & -4 \\ 1 & -1 \end{pmatrix}$ são $\lambda_{1,2} = -1 \pm 2i$.

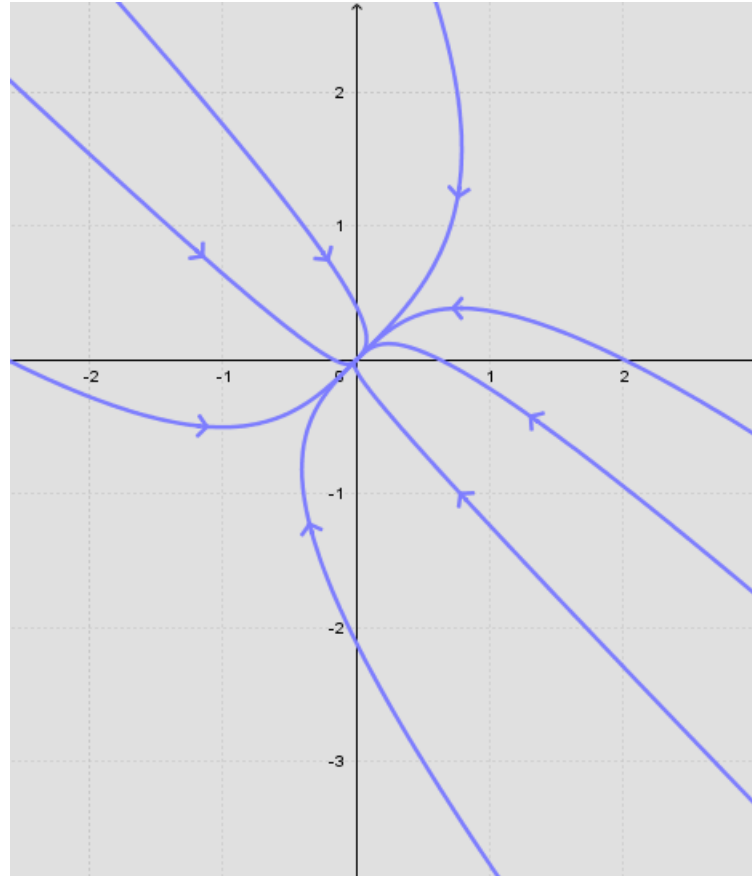


Figura 9 – Plano de fase do sistema linear bidimensional (1.10), em que $\lambda_1 = -1$ e $\lambda_2 = -3$
 Fonte: o autor.

O plano de fase do sistema linear bidimensional para este sistema é mostrado na Figura 10. Neste caso, como os autovalores são complexos e têm a parte real negativa, a origem, que é um ponto de equilíbrio do sistema, representa um ponto espiral assintoticamente estável.

1.3 Sistemas autônomos não lineares bidimensionais

Retornando ao sistema de equações diferenciais (1.1), vamos considerar F e G não lineares do tipo

$$F(x, y) = ax + by + f(x, y) \quad \text{e} \quad G(x, y) = cx + dy + g(x, y), \quad (1.12)$$

em que os termos $f(x, y)$ e $g(x, y)$ são os termos não lineares.

Assim, o sistema (1.1) admite a seguinte forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}. \quad (1.13)$$

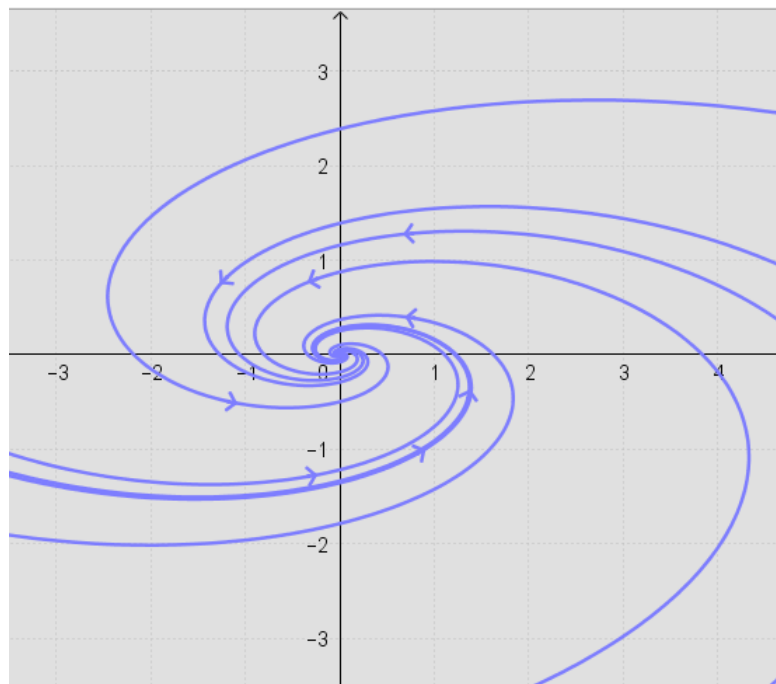


Figura 10 – Plano de fase do sistema linear bidimensional (1.11), em que $\lambda_{1,2} = -1 \pm 2i$
Fonte: o autor.

De acordo com [11], sistemas não lineares descritos por (1.13) que satisfazem

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y)}{\|(x,y)\|} = 0 \quad \text{e} \quad \lim_{(x,y) \rightarrow (0,0)} \frac{g(x,y)}{\|(x,y)\|} = 0 \quad (1.14)$$

são chamados de sistemas quase lineares, que, assim como os sistemas lineares, podem ter seus pontos de equilíbrio caracterizados. Para cada ponto de equilíbrio, devemos encontrar um sistema linear associado em sua vizinhança e caracterizá-lo a partir deste novo sistema.

A Tabela 2 traz uma comparação das propriedades de estabilidade e instabilidade dos pontos de equilíbrio de sistemas lineares e quase lineares. Mais detalhes e exemplos podem ser encontrados em [11].

O sistema visto em (1.4) é um exemplo de sistema bidimensional não linear do tipo (1.13):

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -x(x^2 + y^2) \\ -y(x^2 + y^2) \end{pmatrix}, \quad (1.15)$$

em que, $f(x,y) = -x(x^2 + y^2)$ e $g(x,y) = -y(x^2 + y^2)$ e também satisfazem as condições em (1.14).

Vejamos outro exemplo. Consideremos o sistema

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0,5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -x^2 - xy \\ -0,25y^2 - 0,75xy \end{pmatrix}, \quad (1.16)$$

que representa um modelo de espécies em competição. Mais detalhes podem ser encontrados em [3] e [11].

Tabela 2 – Características comparativas de estabilidade de pontos singulares em sistemas lineares e não lineares

Autovalores	Sistema Linear		Sistema Não Linear	
	Singularidade	Estabilidade	Singularidade	Estabilidade
$0 < \lambda_2 < \lambda_1$	Nó	Instável	Nó	Instável
$\lambda_1 < \lambda_2 < 0$	Nó	Assintoticamente estável	Nó	Assintoticamente estável
$\lambda_1 < 0 < \lambda_2$	Ponto de sela	Instável	Ponto de sela	Instável
$\lambda_1 = \lambda_2 > 0$	Nó próprio ou impróprio	Instável	Nó ou Ponto Espiral	Instável
$\lambda_1 = \lambda_2 < 0$	Nó próprio ou impróprio	Assintoticamente estável	Nó ou Ponto Espiral	Assintoticamente estável
$\lambda_{1,2} = a \pm bi, a > 0$	Ponto espiral	Instável	Ponto espiral	Instável
$\lambda_{1,2} = a \pm bi, a < 0$	Ponto espiral	Assintoticamente estável	Ponto espiral	Assintoticamente estável
$\lambda_{1,2} = a \pm bi, a = 0$	Centro	Estável	Centro ou Ponto Espiral	Indeterminado

Fonte: [11]

Neste sistema, $f(x, y) = -x^2 - xy$ e $g(x, y) = -0,25y^2 - 0,75xy$ satisfazem as condições em (1.14).

Para a origem, que é um dos pontos de equilíbrio do sistema, o sistema linear associado é dado por:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0,5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (1.17)$$

É fácil ver que os autovalores da matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 0,5 \end{pmatrix}$ são $\lambda_1 = 0,5$ e $\lambda_2 = 1$. Logo, de acordo com a Tabela (2), a origem é também um nó instável no sistema não linear (1.16). Os demais pontos de equilíbrio do sistema são $(1/2, 1/2)$, que é um ponto de sela, $(1, 0)$ e $(0, 2)$ que são nós assintoticamente estáveis. Mais detalhes, consultar [3, 11]. Para finalizar, as Figuras (11) e (12) apresentam o plano de fase do sistema (1.16).

1.4 Modelos epidemiológicos

Muitas vezes, os sistemas bidimensionais, vistos nas seções anteriores, são muito simples para descrever certos modelos epidemiológicos. O modelo SIR, proposto por Kermack e McKendrick, por exemplo, é um modelo tridimensional que relaciona a dinâmica entre as populações suscetível, infectada e recuperada de uma determinada doença

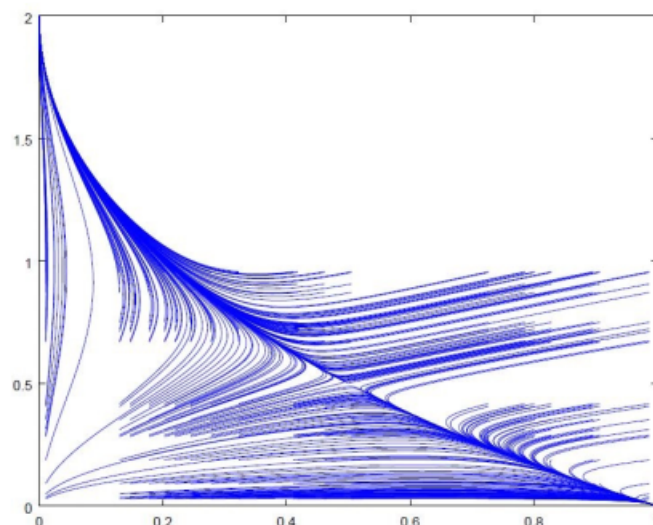


Figura 11 – Plano de fase para o sistema (1.16)

Fonte: Figura elaborada pelo orientador deste TCC em ambiente Octave em [11]

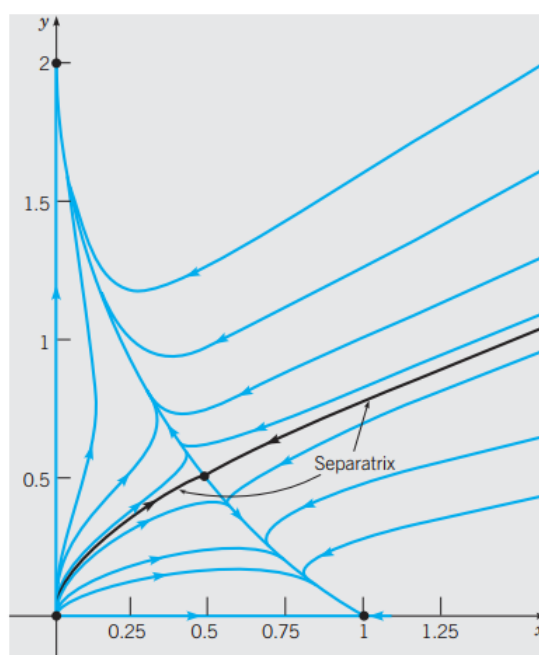


Figura 12 – Plano de fase para o sistema (1.16). A curva em preto é uma separatrix.

Fonte: [3]

infecciosa, representado no sistema (1.18).

$$\begin{aligned} S' &= -\frac{\beta SI}{n} \\ I' &= \frac{\beta SI}{n} - \gamma I, \\ R' &= \gamma I \end{aligned} \tag{1.18}$$

em que S' , I' e R' representam, respectivamente, as taxas de variação da quantidade de indivíduos suscetíveis, infectados e recuperados ao longo do tempo e n representa a população total. Neste caso, γ e β são constantes positivas, em que β representa a taxa

de contato entre indivíduos das classes suscetível e infectado e γ representa a taxa de indivíduos infectados que se recuperam da doença [1].

Em alguns casos, podemos simplificar modelos tridimensionais para um modelo bidimensional. No sistema (1.18), se considerarmos que a população total, n , será constante em qualquer instante t , isto é, $n = S(t) + I(t) + R(t)$, é possível determinar o valor de R quando S e I são conhecidos, de forma que o sistema (1.18) pode ser simplificado em um sistema bidimensional:

$$\begin{aligned} S' &= -\beta SI \\ I' &= \beta SI - \gamma I. \end{aligned} \tag{1.19}$$

Uma vez resolvido (1.19), obtém-se R através de $R = N - S - I$.

No capítulo 3 trabalharemos com uma variação do modelo tridimensional (1.18), denominado Modelo *Time-dependent* SIR, apresentado em [8], em que γ e β também são variáveis no tempo.

2 Introdução à Interpolação Polinomial, Regressão Linear e Quadrados Míni- mos Lineares

Neste capítulo, estabeleceremos métodos numéricos que permitem encontrar uma função que se ajuste, como uma aproximação, para um conjunto de dados conhecidos.

2.1 Interpolação Polinomial

Inicialmente, vamos estabelecer métodos numéricos que permitam encontrar um polinômio que passe por $n + 1$ pontos conhecidos, denotados por

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), \text{ com } x_0 < x_1 < \dots < x_n.$$

A interpolação polinomial é um método no qual conhecemos a função em pontos discretos, isto é, conhecemos $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$ e queremos encontrar um polinômio $p(x)$ tal que $p(x_0) = f(x_0), p(x_1) = f(x_1), \dots, p(x_n) = f(x_n)$. Neste caso, $p(x)$ é chamado polinômio interpolador [14].

Na Figura 13 temos uma representação de um polinômio interpolador $p(x)$ dos $n + 1$ pontos dados.

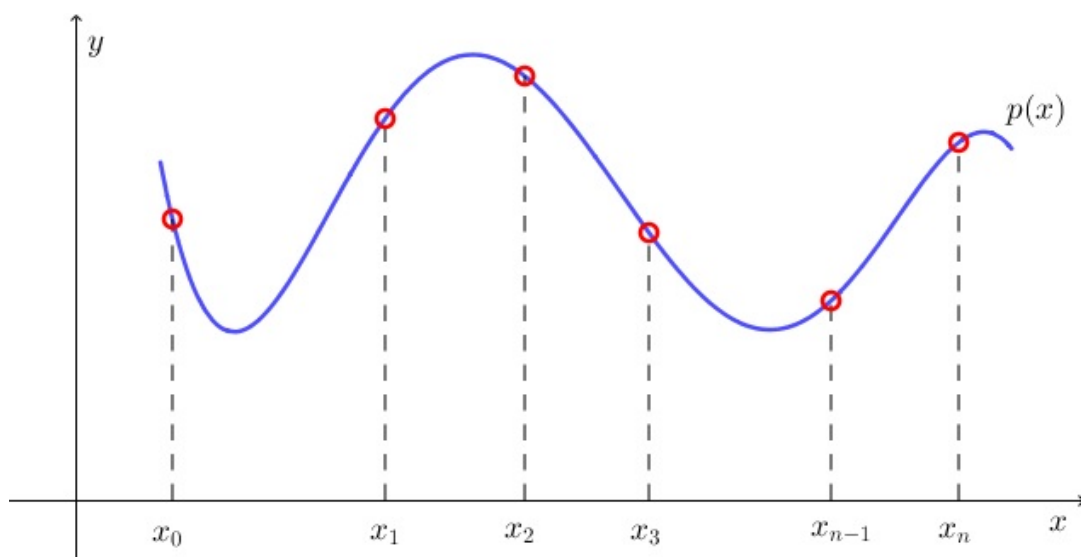


Figura 13 – Exemplo de polinômio interpolador.

Fonte: o autor.

O teorema a seguir nos garante a existência e a unicidade desse polinômio, cuja demonstração pode ser encontrada em [14]. Este problema é conhecido como *Interpolação de Lagrange*.

Teorema 2.1.1 *Suponha que $f(x)$ seja uma função conhecida nos $n + 1$ pontos, distintos dois a dois, x_0, x_1, \dots, x_n . Então existe um único polinômio $p(x)$, de grau menor ou igual a n , tal que*

$$p(x_i) = f(x_i), i = 0, \dots, n. \quad (2.1)$$

Como o polinômio interpolador se trata de uma aproximação, é natural que erros sejam introduzidos nesta aproximação. Assim, o resultado a seguir, proposto em [14], nos dá informação sobre a medida do erro do polinômio interpolador.

Teorema 2.1.2 *Seja $[a, b]$ um intervalo que contém os pontos x_0, x_1, \dots, x_n , distintos dois a dois, e $f(x)$ uma função com $n + 1$ derivadas contínuas em $[a, b]$. Pode-se mostrar que o erro E do polinômio interpolador p_n de f no ponto $x \in [a, b]$ é dado por*

$$E(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(\alpha)}{(n + 1)!} [(x - x_0)(x - x_1) \cdots (x - x_n)], \quad (2.2)$$

onde a derivada de ordem $n + 1$ é calculada num ponto $\alpha \in [a, b]$.

Visto que o Teorema (2.1.1) garante a existência do polinômio interpolador, apresentamos a seguir uma forma de encontrar este polinômio, proposta por Newton, cujo procedimento consiste de uma relação de recorrência, conhecida como *diferenças divididas*.

Primeiramente, vamos escrever o polinômio interpolador na forma

$$p_n(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \cdots + c_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}). \quad (2.3)$$

As constantes c_0, c_1, \dots, c_n serão determinadas através das condições de interpolação, dadas em (2.1), com o auxílio do método algébrico de *Diferenças Divididas*, conforme definição abaixo.

Definição 2.1.1 *Definimos inicialmente $f[x_0] = f(x_0)$. Por recorrência definimos, para qualquer k ,*

$$f[x_0, x_1, \dots, x_{k-1}, x_k, x] = \frac{f[x_0, x_1, \dots, x_{k-1}, x] - f[x_0, x_1, \dots, x_{k-1}, x_k]}{x - x_k}. \quad (2.4)$$

Assim (conforme estabelecido em [14]) os coeficientes c_k do polinômio interpolador podem ser obtidos por:

$$c_k = f[x_0, x_1, \dots, x_k], \quad k = 0, \dots, n. \quad (2.5)$$

Existe um meio prático de calcular as diferenças divididas obtido através de um pequeno algoritmo (veja algoritmo 5.1 de [14]).

2.2 Quadrados Mínimos Lineares

O método dos mínimos quadrados lineares (ou mínimos quadrados ordinários) é uma técnica no qual queremos aproximar uma função $f(x)$ por outra função $g(x)$, tomando como hipóteses algumas informações sobre a forma de $g(x)$, como por exemplo, se $g(x)$ é uma reta, uma parábola, uma função exponencial, entre outras.

Seja $f(x)$ uma função supostamente conhecida em um conjunto finito de pontos, x_1, x_2, \dots, x_m , também chamado de malha do intervalo $[a, b]$. Queremos aproximar $f(x)$ por uma função $g(x)$ da forma:

$$f(x) \cong g(x) = c_1\varphi_1(x) + c_2\varphi_2(x) + \dots + c_n\varphi_n(x) \quad (2.6)$$

em que as $\varphi_1, \varphi_2, \dots, \varphi_n$ são funções predefinidas. Se quisermos aproximar $f(x)$ por polinômios, por exemplo, as funções φ_i podem ser os monômios x^i .

O objetivo do método dos quadrados mínimos é encontrar os coeficientes c_1, c_2, \dots, c_n que nos dê a melhor aproximação de $f(x)$ por $g(x)$. Para atingir esse objetivo devemos estabelecer, em cada ponto $x_i, i = 1, \dots, m$, o resíduo (ou desvio), $r(x_i)$, entre as funções f e g , definido por:

$$\begin{aligned} r(x_i) &= f(x_i) - g(x_i) \\ &= f(x_i) - [c_1\varphi_1(x_i) + c_2\varphi_2(x_i) + \dots + c_n\varphi_n(x_i)]. \end{aligned} \quad (2.7)$$

Para a determinação da melhor aproximação, utilizamos a minimização da soma dos quadrados dos resíduos em (2.7). Para tanto, suponha $f(x_i), i = 1, \dots, m$, os dados conhecidos e queremos estabelecer os coeficientes c_1, c_2, \dots, c_n que nos levarão à melhor aproximação. Para atingir esse objetivo, devemos então minimizar

$$\sum_{i=1}^m r^2(x_i) = \sum_{i=1}^m (f(x_i) - g(x_i))^2. \quad (2.8)$$

O caso visto acima representa o caso discreto. Para o caso contínuo, em que $f(x)$ é conhecida em todo $x \in [a, b]$, é natural que minimizemos

$$\int_a^b r^2(x)dx = \int_a^b (f(x) - g(x))^2 dx, \quad (2.9)$$

onde, agora, $f(x)$ e $g(x)$ são supostamente integráveis em $[a, b]$.

As expressões (2.8) e (2.9) representam, na verdade, o produto escalar de $f - g$ por ela mesma para os casos discreto e contínuo respectivamente, conforme se pode verificar pela definição abaixo.

Definição 2.2.1 *Sejam as funções $f(x)$ e $g(x)$. O produto escalar entre elas é denotado e definido por*

$$\langle f, g \rangle = \langle f(x), g(x) \rangle = \begin{cases} \sum_{i=1}^m f(x_i)g(x_i), & \text{no caso discreto} \\ \int_a^b f(x)g(x), & \text{no caso contínuo.} \end{cases} \quad (2.10)$$

Assim, o problema dos mínimos quadrados pode ser transformado em um problema de minimização do produto escalar

$$\langle r(x; c), r(x; c) \rangle = \langle r, r \rangle = \langle f(x) - \sum_{i=1}^n c_i \varphi_i(x), f(x) - \sum_{i=1}^n c_i \varphi_i(x) \rangle, \quad (2.11)$$

visto, agora, como uma função dos coeficientes $c := (c_1, c_2, \dots, c_n)$.

A determinação das constantes c_1, c_2, \dots, c_n da aproximação definida por (2.11) equivale a encontrar o ponto crítico da função

$$c \mapsto \langle r(x; c), r(x; c) \rangle$$

que por sua vez, equivale a resolver, para c_1, c_2, \dots, c_n o *Sistema Normal*

$$\begin{cases} \langle \varphi_1, \varphi_1 \rangle c_1 + \langle \varphi_2, \varphi_1 \rangle c_2 + \dots + \langle \varphi_n, \varphi_1 \rangle c_n = \langle f, \varphi_1 \rangle \\ \langle \varphi_1, \varphi_2 \rangle c_1 + \langle \varphi_2, \varphi_2 \rangle c_2 + \dots + \langle \varphi_n, \varphi_2 \rangle c_n = \langle f, \varphi_2 \rangle \\ \vdots \\ \langle \varphi_1, \varphi_n \rangle c_1 + \langle \varphi_2, \varphi_n \rangle c_2 + \dots + \langle \varphi_n, \varphi_n \rangle c_n = \langle f, \varphi_n \rangle. \end{cases} \quad (2.12)$$

2.2.1 Regressão Linear para o caso unidimensional

Para exemplificar a teoria vista, apresentamos a seguir um exemplo clássico, conhecido como regressão linear. Na regressão linear, queremos encontrar a reta que melhor se ajusta a um conjunto de m pontos conhecidos, $(x_i, f(x_i))$, $i = 1, \dots, m$.

Como queremos a aproximação de $f(x)$ por uma reta, devemos ter $\varphi_1(x) = 1$ e $\varphi_2(x) = x$, de modo que

$$f(x) \cong g(x) = c_1 + c_2x.$$

Para cada ponto x_i conhecido, o resíduo pode ser facilmente determinado por $r(x_i) = f(x_i) - c_1 - c_2x_i$. Assim, para obtermos a melhor aproximação, devemos minimizar a função

$$\langle r, r \rangle = \sum_{i=1}^m (f(x_i) - c_1 - c_2x_i)^2.$$

Para determinarmos as constantes c_1 e c_2 , precisamos resolver o Sistema Normal dado por (2.12). Para isso, primeiramente devemos determinar os seguintes produtos

escalares, conforme a definição (2.2.1):

$$\begin{aligned}\langle \varphi_1, \varphi_1 \rangle &= \sum_{i=1}^m \varphi_1 \varphi_1 = m \\ \langle \varphi_1, \varphi_2 \rangle = \langle \varphi_2, \varphi_1 \rangle &= \sum_{i=1}^m \varphi_2 \varphi_1 = \sum_{i=1}^m x_i \\ \langle \varphi_2, \varphi_2 \rangle &= \sum_{i=1}^m \varphi_2 \varphi_2 = \sum_{i=1}^m x_i^2 \\ \langle f, \varphi_1 \rangle &= \sum_{i=1}^m f(x_i) \varphi_1 = \sum_{i=1}^m f(x_i) \\ \langle f, \varphi_2 \rangle &= \sum_{i=1}^m f(x_i) \varphi_2 = \sum_{i=1}^m x_i f(x_i).\end{aligned}$$

Assim, obtemos o seguinte Sistema Normal:

$$\begin{cases} mc_1 + c_2 \sum_{i=1}^m x_i = \sum_{i=1}^m f(x_i) \\ c_1 \sum_{i=1}^m x_i + c_2 \sum_{i=1}^m x_i^2 = \sum_{i=1}^m x_i f(x_i). \end{cases}$$

Para ilustrarmos numericamente a regressão linear, vamos fazer o ajuste linear para os dados da tabela abaixo, em que $m = 6$.

x	0	0,15	0,31	0,5	0,6	0,75
f(x)	1,0	1,004	1,031	1,117	1,223	1,422

Usando os valores apresentados na tabela temos:

$$\begin{aligned}\sum_{i=1}^6 x_i &= 2,3100 & \sum_{i=1}^6 x_i^2 &= 1,2911 \\ \sum_{i=1}^6 f(x_i) &= 6,7970 & \sum_{i=1}^6 x_i f(x_i) &= 2,8290.\end{aligned}$$

Dessa forma, resolvendo o sistema:

$$\begin{cases} 6c_1 + 2,31c_2 = 6,797 \\ 2,31c_1 + 1,2911c_2 = 2,829 \end{cases}$$

obtemos $c_1 = 0,9295$ e $c_2 = 0,5281$, ou seja, a reta que melhor se ajusta aos pontos da tabela acima é dada por

$$g(x) = 0,9295 + 0,5281x.$$

Para mais exemplos consulte a seção 5.4 de [14].

2.3 Regressão Linear para Dimensões Maiores

Na seção anterior, apresentamos os conceitos de regressão linear e mínimos quadrados para o caso unidimensional em que pontos de uma malha de um intervalo são

conhecidos. Nesta seção, apresentaremos este conceito para dimensões maiores, em que os pontos da malha são, agora, vetores.

Considere um conjunto de dados conhecidos $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, em que cada $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T \in \mathbb{R}^p$ e $y_i \in \mathbb{R}$, $i = 1, \dots, m$. Nosso objetivo é obter constantes c_0, c_1, \dots, c_p que nos dê a melhor aproximação de y_i por $g(x_i)$, $i = 1, \dots, m$ de forma que:

$$y_i \cong g(x_i) = c_0 + c_1x_{i1} + c_2x_{i2} + \dots + c_px_{ip} \quad (2.13)$$

onde g é uma função de \mathbb{R}^p em \mathbb{R} . Observe que o exemplo de regressão linear visto na seção (2.2.1) corresponde, aqui, ao caso em que $p = 1$.

Para atingir esse objetivo, estenderemos aqui o caso discreto do método visto na seção (2.2). Para tanto, definimos os vetores

$$y = (y_1, y_2, \dots, y_m)^T \quad \text{e} \quad \mathbf{g} = (g(x_1), g(x_2), \dots, g(x_m))^T.$$

Assim, de maneira similar, o resíduo entre y e \mathbf{g} é definido por

$$r = r(c) = y - \mathbf{g} = (y_1 - g(x_1), y_2 - g(x_2), \dots, y_m - g(x_m))^T.$$

De acordo com o método dos mínimos quadrados, devemos encontrar constantes c_0, c_1, \dots, c_p que minimizem

$$\|r(c)\|_2^2 = \langle r, r \rangle = \sum_{i=1}^m (y_i - g(x_i))^2, \quad (2.14)$$

em que $\|\cdot\|_2$ representa a norma euclidiana.

Para simplificar nossa abordagem, vamos particularizar o problema para o caso em que $p = 2$ e $m = 4$. Assim, cada vetor $x_i = (x_{i1}, x_{i2})^T \in \mathbb{R}^2$, $i = 1, \dots, 4$ e o conjunto de dados conhecidos é $(x_1, y_1) = (x_{11}, x_{12}, y_1)$, $(x_2, y_2) = (x_{21}, x_{22}, y_2)$, $(x_3, y_3) = (x_{31}, x_{32}, y_3)$ e $(x_4, y_4) = (x_{41}, x_{42}, y_4)$.

A aproximação que queremos obter é da forma

$$y_i \cong g(x_i) = c_0 + c_1x_{i1} + c_2x_{i2}, \quad i = 1, \dots, 4. \quad (2.15)$$

Em termos matriciais:

$$\mathbf{g} = \begin{bmatrix} g(x_1) \\ g(x_2) \\ g(x_3) \\ g(x_4) \end{bmatrix} = \begin{bmatrix} c_0 + c_1x_{11} + c_2x_{12} \\ c_0 + c_1x_{21} + c_2x_{22} \\ c_0 + c_1x_{31} + c_2x_{32} \\ c_0 + c_1x_{41} + c_2x_{42} \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}. \quad (2.16)$$

Definindo

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \\ 1 & x_{41} & x_{42} \end{bmatrix} \quad \text{e} \quad c = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}, \quad (2.17)$$

o vetor \mathbf{g} pode ser reescrito como

$$\mathbf{g} = Xc \quad (2.18)$$

e o resíduo

$$r(c) = y - Xc. \quad (2.19)$$

As fórmulas (2.16), (2.17), (2.18) e (2.19) foram desenvolvidas para o caso particular em que $p = 2$ e $m = 4$, porém, elas podem ser estendidas para o caso genérico em que p e m são naturais arbitrários. Assim, voltando ao caso genérico, para encontrarmos a melhor aproximação para (2.13), devemos obter constantes c_0, c_1, \dots, c_p que minimizem

$$\|y - Xc\|_2^2 = (y - Xc)^T(y - Xc) \quad (2.20)$$

em que X é, agora, uma matriz $m \times (p + 1)$ e o vetor c é de tamanho $(p + 1) \times 1$.

Encontrar o vetor c que minimiza (2.20) equivale a encontrar o ponto crítico da função

$$RSS(c) := (y - Xc)^T(y - Xc). \quad (2.21)$$

A função em (2.21) é uma função quadrática em $p + 1$ parâmetros. Diferenciando (2.21) duas vezes com relação a c obtemos

$$\frac{\partial RSS}{\partial c} = -2X^T(y - Xc) \quad (2.22)$$

$$\frac{\partial^2 RSS}{\partial c \partial c^T} = 2X^T X. \quad (2.23)$$

Igualando (2.22) a zero, obtemos

$$X^T(y - Xc) = 0. \quad (2.24)$$

Assumindo que X tem posto coluna cheio, $X^T X$ é positiva definida. Consequentemente, obtemos a solução única de (2.24)

$$\hat{c} = (X^T X)^{-1} X^T y, \quad (2.25)$$

que, por sua vez, representa o vetor c procurado que minimiza a norma em (2.20). Para mais detalhes, consultar seção 3.2 de [19].

Levando \hat{c} em (2.18) obtemos a aproximação procurada

$$\hat{\mathbf{g}} = X\hat{c}.$$

Note que as coordenadas de $\hat{\mathbf{g}}$ representam os valores da expressão do lado direito de (2.13) com os coeficientes dados pelas coordenadas do vetor \hat{c} .

3 Modelo Epidemiológico

Neste capítulo, apresentaremos uma variação do modelo epidemiológico SIR visto na seção 1.4, o modelo *Time-dependent SIR*, que será utilizado para a modelagem da epidemia do vírus SARS-CoV-2, causador da Covid-19.

Como vimos anteriormente, o modelo SIR pode ser representado pelo sistema (1.18):

$$\begin{aligned} S' &= -\frac{\beta SI}{n} \\ I' &= \frac{\beta SI}{n} - \gamma I, \\ R' &= \gamma I \end{aligned} \tag{3.1}$$

em que $S(t) + I(t) + R(t) = n$, onde n representa a população total.

Neste modelo clássico, temos duas variáveis invariantes no tempo, as constantes positivas γ e β que, como explicado anteriormente, representam a taxa de indivíduos infectados que se recuperam da doença e a taxa de contato entre indivíduos das classes suscetível e infectado respectivamente.

Por outro lado, em [8] vemos que, ao negligenciarmos as propriedades de que γ e β variam ao longo do tempo, torna-se mais difícil prever com precisão e eficácia o comportamento da doença. Assim, um novo modelo, no qual as taxas γ e β são funções do tempo, é apresentado no sistema (3.2) a seguir:

$$\begin{aligned} S' &= -\frac{\beta(t)S(t)I(t)}{n} \\ I' &= \frac{\beta(t)S(t)I(t)}{n} - \gamma(t)I(t). \\ R' &= \gamma(t)I(t) \end{aligned} \tag{3.2}$$

Neste sistema, S' , I' e R' representam as taxas de variação de pessoas suscetíveis, infectadas e recuperadas no tempo t . Considerando que a população total é n , a probabilidade de que uma pessoa, escolhida aleatoriamente, seja suscetível é $S(t)/n$. Dessa forma, um indivíduo infectado contamina, em média, $\beta(t)S(t)/n$ pessoas suscetíveis, por unidade de tempo t , ou seja, a cada unidade de tempo t , temos $\beta(t)S(t)I(t)/n$ novos indivíduos infectados. Isso significa que a população suscetível irá diminuir nessa mesma quantidade de indivíduos. Além disso, os indivíduos infectados irão se recuperar da doença a uma taxa $\gamma(t)$. Assim, $\gamma(t)I(t)$ indivíduos se recuperam no tempo t , o que significa que, nessa mesma quantidade, a população de infectados irá diminuir.

Para melhor trabalharmos com o sistema que acabamos de descrever, apresentaremos o modelo em sua forma discreta e forneceremos uma aproximação das derivadas presentes no sistema de equações diferenciais do modelo (3.2), utilizando o Método das Diferenças Finitas. Assim, a versão discreta para o modelo (3.2) é apresentada nas equações (3.3), (3.4) e (3.5):

$$S(t+1) - S(t) = -\frac{\beta(t)S(t)I(t)}{n} \quad (3.3)$$

$$I(t+1) - I(t) = \frac{\beta(t)S(t)I(t)}{n} - \gamma(t)I(t) \quad (3.4)$$

$$R(t+1) - R(t) = \gamma(t)I(t). \quad (3.5)$$

Nestes modelos, levando-se em consideração o início da pandemia, a população total é supostamente mantida constante, sendo igual a soma das populações suscetível, infectada e recuperada. Além disso, consideramos que toda a população é suscetível, isto é, $S(t) \approx n$. Dessa forma, podemos reescrever a equação (3.4) como:

$$I(t+1) - I(t) = \beta(t)I(t) - \gamma(t)I(t). \quad (3.6)$$

Das equações (3.5) e (3.6), podemos obter expressões para determinar $\gamma(t)$ e $\beta(t)$, apresentadas a seguir.

$$\gamma(t) = \frac{R(t+1) - R(t)}{I(t)}, \quad (3.7)$$

$$\beta(t) = \frac{[I(t+1) - I(t)] + [R(t+1) - R(t)]}{I(t)}. \quad (3.8)$$

Assim, com base em um histórico dos dados de infectados e recuperados para um certo período T de tempo, $\{I(t), R(t); 0 \leq t \leq T-1\}$, podemos obter as taxas de transmissão e recuperação, $\{\gamma(t), \beta(t); 0 \leq t \leq T-2\}$, utilizando as expressões (3.7) e (3.8). Dessa forma, a partir destes novos dados obtidos, é possível utilizar técnicas para obter estimativas para as taxas $\gamma(t)$ e $\beta(t)$ para $t \geq T$, em que T representa o dia a partir do qual desejamos fazer tais previsões. Para atingir este objetivo, utilizaremos um método para resolução de problemas de quadrados mínimos lineares regularizado, chamado *Ridge Regression*, apresentado na seção 3.1.

3.1 Ridge Regression

O método *Ridge Regression*, também conhecido como Regularização de *Tikhonov*, é uma variação dos quadrados mínimos lineares vistos na seção 2.2. A principal diferença está relacionada à obtenção dos coeficientes de *ridge* que, neste caso, são estimados minimizando

quantidades ligeiramente diferentes [19, 20]. Em particular, os coeficientes de *ridge* c^{ridge} são os valores que minimizam o problema:

$$c^{ridge} = \underset{c}{\operatorname{argmin}} \left\| y - Xc \right\|_2^2 + \lambda \left\| c \right\|_2^2 \quad (3.9)$$

em que $X \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^{m \times 1}$, $c \in \mathbb{R}^{n \times 1}$ e $\lambda > 0$ é o *parâmetro de regularização*, a ser determinado convenientemente nos problemas de aplicação [20, 21].

A expressão (3.9) pode ser reescrita da seguinte forma:

$$\underset{c}{\operatorname{argmin}} \left\| y - Xc \right\|_2^2 + \lambda \left\| c \right\|_2^2 = \underset{c}{\operatorname{argmin}} \left\| \begin{pmatrix} Xc - y \\ \sqrt{\lambda}c \end{pmatrix} \right\|_2^2 = \underset{c}{\operatorname{argmin}} \left\| \begin{pmatrix} X \\ \sqrt{\lambda}I_n \end{pmatrix} c - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|_2^2. \quad (3.10)$$

Conforme vimos na equação (2.25), da seção 2.3, a solução do problema (3.10) pode ser obtida através da equação normal

$$\begin{pmatrix} X \\ \sqrt{\lambda}I_n \end{pmatrix}^T \begin{pmatrix} X \\ \sqrt{\lambda}I_n \end{pmatrix} c = \begin{pmatrix} X \\ \sqrt{\lambda}I_n \end{pmatrix}^T \begin{pmatrix} y \\ 0 \end{pmatrix}$$

ou, equivalentemente,

$$(X^T X + \lambda I_n) c = X^T y. \quad (3.11)$$

Definição 3.1.1 *Uma matriz W é dita ortogonal se é inversível e $W^{-1} = W^T$, ou seja, $WW^T = W^T W = I$, em que I é a matriz Identidade.*

O teorema a seguir nos diz que toda matriz possui uma decomposição em valores singulares.

Teorema 3.1.1 Teorema da Decomposição em Valores Singulares (SVD) *Seja a matriz $X \in \mathbb{R}^{m \times n}$. Então existem matrizes ortogonais $U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m}$ e $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$, tais que $U^T X V = D$, em que D é uma matriz retangular diagonal $m \times n$ dada em blocos na forma:*

$$D = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \quad (3.12)$$

em que

$$\Sigma_r = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \quad (3.13)$$

com $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$, em que $p = \min(m, n)$ e r é o posto da matriz X .

Os números $\sigma_1, \sigma_2, \dots, \sigma_p$ são chamados *valores singulares* de X e os vetores coluna u_i e v_j são respectivamente o *i-ésimo vetor singular à esquerda* e o *j-ésimo vetor singular à direita*.

Comparando as colunas nas equações $XV = UD$ e $X^T U = VD^T$, é fácil concluir que

$$\begin{cases} Xv_i = \sigma_i u_i \\ X^T u_i = \sigma_i v_i \end{cases} \quad \forall i = 1 : p.$$

Outras observações interessantes que a decomposição SVD de uma matriz X revela no Teorema 3.1.1, além de $r = \text{rank}(X)$, é que $\text{Ker}(X) = \text{span}\{v_{r+1}, \dots, v_n\}$ e $\text{Im}(X) = \text{span}\{u_1, \dots, u_r\}$.

Finalmente, se denotarmos $U_r = U(:, 1 : r)^1$, $\Sigma_r = D(1 : r, 1 : r)^2$ e $V_r = V(:, 1 : r)$, então a expansão SVD de X será

$$X = U_r \Sigma_r V_r^T = \sum_{i=1}^r \sigma_i u_i v_i^T.$$

A demonstração do Teorema 3.1.1 pode ser encontrada em [18].

Assim, substituindo X pela sua decomposição SVD [19], a equação (3.11) se reescreve como

$$[(UDV^T)^T(UDV^T) + \lambda I_n]c = (UDV^T)^T y$$

ou, equivalentemente,

$$(VD^T U^T U D V^T + \lambda I_n)c = VD^T U^T y, \quad (3.14)$$

em que I_n é a identidade de ordem n , $U \in \mathbb{R}^{m \times m}$ e $V \in \mathbb{R}^{n \times n}$ são as matrizes ortogonais e $D \in \mathbb{R}^{m \times n}$ é a matriz diagonal, especificadas conforme o Teorema 3.1.1.

Substituindo $U^T U$ por I_m e I_n por $V V^T$ em (3.14), resulta

$$\begin{aligned} (VD^T D V^T + \lambda V V^T)c &= VD^T U^T y \\ V(D^T D + \lambda I_n)V^T c &= VD^T U^T y. \end{aligned} \quad (3.15)$$

Multiplicando a equação (3.15) à esquerda por V^T , obtemos:

$$(D^T D + \lambda I_n)V^T c = D^T U^T y. \quad (3.16)$$

Agora, fazendo a substituição $\xi = V^T c$, a equação (3.16) pode ser reescrita como

$$(D^T D + \lambda I_n)\xi = D^T U^T y. \quad (3.17)$$

¹ $X(:, 1 : r)$ representa a submatriz de X dada pelas suas r primeiras colunas e $Y(1 : s, :)$ representa a submatriz de Y dada pelas suas s primeiras linhas.

² $X(1 : s, 1 : r)$ representa a submatriz de X dada pelas suas s primeiras linhas e r primeiras colunas.

Note que $D^T D + \lambda I_n$ é a matriz diagonal $\text{diag}(\sigma_1^2 + \lambda, \dots, \sigma_r^2 + \lambda, \lambda, \dots, \lambda) \in \mathbb{R}^{n \times n}$, enquanto $D^T U^T$ é a matriz cujas linhas são os vetores $\begin{cases} \sigma_i u_i^T, & i = 1, \dots, r \\ 0, & i = r + 1, \dots, n \end{cases}$. Assim, a solução para a equação matricial (3.17) é dada por [21]:

$$\xi_i = \begin{cases} \frac{\sigma_i(u_i^T y)}{\sigma_i^2 + \lambda}, & i = 1 : r \\ 0, & i = r + 1 : n. \end{cases} \quad (3.18)$$

Como $\xi = V^T c$, então, a solução para a equação matricial (3.11) é obtida por

$$c = V\xi = \begin{bmatrix} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sum_{i=1}^r \xi_i v_i = \sum_{i=1}^r \frac{\sigma_i(u_i^T y)}{\sigma_i^2 + \lambda} v_i. \quad (3.19)$$

Portanto, a solução do problema de mínimos quadrados lineares regularizado (3.9) é dada por:

$$c^{\text{ridge}} = \sum_{i=1}^r \frac{\sigma_i(u_i^T y)}{\sigma_i^2 + \lambda} v_i. \quad (3.20)$$

Observemos que

$$\lim_{\lambda \rightarrow 0} c^{\text{ridge}} = \lim_{\lambda \rightarrow 0} \sum_{i=1}^r \frac{\sigma_i(u_i^T y)}{\sigma_i^2 + \lambda} v_i = \sum_{i=1}^r \frac{(u_i^T y)}{\sigma_i} v_i. \quad (3.21)$$

Além disso, é imediata a verificação de que a somatória do lado direito de (3.21) representa a solução do problema³

$$c^* = \underset{c}{\operatorname{argmin}} \left\| y - Xc \right\|_2^2, \quad (3.22)$$

ou seja, a solução do problema regularizado (3.9) converge para a solução do problema de quadrados mínimos lineares não regularizado (3.22) quando $\lambda \rightarrow 0$.

Sobre a expressão da solução obtida, vale destacar que ela revela como a adição da parcela de regularização, $\lambda \left\| c \right\|_2^2$, ao problema de quadrados mínimos lineares atua como um filtro [21]. De fato, da expressão (3.20), temos que:

$$\frac{\sigma_i(u_i^T y)}{\sigma_i^2 + \lambda} \approx \begin{cases} 0, & \text{se } 0 \approx \sigma_i \ll \lambda \\ \frac{u_i^T y}{\sigma_i}, & \text{se } \sigma_i \gg \lambda \end{cases}, \quad (3.23)$$

³ De fato, o problema (3.22) é um caso particular de (3.9) quando se faz $\lambda = 0$.

de forma que, as parcelas correspondentes aos valores singulares que forem muito maiores que o parâmetro de regularização λ , permanecem quase inalteradas se comparadas com as respectivas parcelas da solução do problema não regularizado. Já as parcelas correspondentes aos valores singulares que forem muito menores que λ são praticamente zeradas [21].

3.2 Ridge Regression Aplicado ao Modelo Epidemiológico

Voltemos aos comentários do parágrafo imediatamente anterior à seção 3.1. Vamos denotar por $\hat{\beta}(t)$ e $\hat{\gamma}(t)$, respectivamente, as taxas previstas de transmissão e recuperação da Covid-19. Essas taxas [8] podem ser obtidas usando-se uma técnica de filtragem de ruídos, chamada *Finite Impulse Response* (FIR), apresentados nas equações (3.24) e (3.25) a seguir:

$$\hat{\beta}(t) = a_1\beta(t-1) + a_2\beta(t-2) + \cdots + a_J\beta(t-J) + a_0 = \sum_{j=1}^J a_j\beta(t-j) + a_0, \quad (3.24)$$

$$\hat{\gamma}(t) = b_1\gamma(t-1) + b_2\gamma(t-2) + \cdots + b_K\gamma(t-K) + b_0 = \sum_{k=1}^K b_k\gamma(t-k) + b_0, \quad (3.25)$$

em que J e K representam as *ordens dos filtros FIR*, com $0 < J, K < T - 2$, e os números $a_j, j = 0, 1, \dots, J$ e $b_k, k = 0, 1, \dots, K$ representam os *coeficientes dos filtros FIR* utilizados. Mais detalhes sobre filtros FIR podem ser encontrados em [26].

Para obtermos os coeficientes a_j e b_k dos filtros, utilizaremos o método *Ridge Regression* visto na seção 3.1 aplicado ao seguinte problema com regularização:

$$\operatorname{argmin}_{a_j} \left\{ \sum_{t=J}^{T-2} (\beta(t) - \hat{\beta}(t))^2 + \alpha_1 \sum_{j=0}^J a_j^2 \right\} \quad (3.26)$$

$$\operatorname{argmin}_{b_k} \left\{ \sum_{t=K}^{T-2} (\gamma(t) - \hat{\gamma}(t))^2 + \alpha_2 \sum_{k=0}^K b_k^2 \right\} \quad (3.27)$$

em que α_1 e α_2 são os parâmetros de regularização. Substituindo as equações (3.24) e (3.25) nas expressões (3.26) e (3.27), respectivamente, obtemos:

$$\operatorname{argmin}_{a_j} \left\{ \sum_{t=J}^{T-2} \left[\beta(t) - \left(\sum_{j=1}^J a_j\beta(t-j) + a_0 \right) \right]^2 + \alpha_1 \sum_{j=0}^J a_j^2 \right\} \quad (3.28)$$

$$\operatorname{argmin}_{b_k} \left\{ \sum_{t=K}^{T-2} \left[\gamma(t) - \left(\sum_{k=1}^K b_k\gamma(t-k) + b_0 \right) \right]^2 + \alpha_2 \sum_{k=0}^K b_k^2 \right\}. \quad (3.29)$$

É fácil verificar que a somatória dos quadrados dos termos entre colchetes de (3.28) representa a norma-2 do vetor $y - Xa$, com y , X e a indicados na expressão abaixo:

$$\underbrace{\begin{pmatrix} \beta(J) \\ \beta(J+1) \\ \vdots \\ \beta(T-3) \\ \beta(T-2) \end{pmatrix}}_y - \underbrace{\begin{pmatrix} 1 & \beta(J-1) & \beta(J-2) & \cdots & \beta(0) \\ 1 & \beta(J) & \beta(J-1) & \cdots & \beta(1) \\ \vdots & & \vdots & \ddots & \vdots \\ 1 & \beta(T-4) & \beta(T-5) & \cdots & \beta(T-3-J) \\ 1 & \beta(T-3) & \beta(T-4) & \cdots & \beta(T-2-J) \end{pmatrix}}_X \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{J-1} \\ a_J \end{pmatrix}}_a \quad (3.30)$$

em que $X \in \mathbb{R}^{(T-J-1) \times (J+1)}$, $y \in \mathbb{R}^{(T-(J+1)) \times 1}$ e $a \in \mathbb{R}^{(J+1) \times 1}$.

Observação análoga vale para a expressão (3.29):

$$\underbrace{\begin{pmatrix} \gamma(K) \\ \gamma(K+1) \\ \vdots \\ \gamma(T-3) \\ \gamma(T-2) \end{pmatrix}}_z - \underbrace{\begin{pmatrix} 1 & \gamma(K-1) & \gamma(K-2) & \cdots & \gamma(0) \\ 1 & \gamma(K) & \gamma(K-1) & \cdots & \gamma(1) \\ \vdots & & \vdots & \ddots & \vdots \\ 1 & \gamma(T-4) & \gamma(T-5) & \cdots & \gamma(T-3-K) \\ 1 & \gamma(T-3) & \gamma(T-4) & \cdots & \gamma(T-2-K) \end{pmatrix}}_Y \underbrace{\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{K-1} \\ b_K \end{pmatrix}}_b \quad (3.31)$$

em que $Y \in \mathbb{R}^{(T-K-1) \times (K+1)}$, $z \in \mathbb{R}^{(T-(K+1)) \times 1}$ e $b \in \mathbb{R}^{(K+1) \times 1}$.

Conseqüentemente, para obtermos os coeficientes dos filtros FIR, apresentados em (3.24) e (3.25), devemos resolver os problemas de mínimos quadrados lineares regularizados dados na forma matricial:

$$a^{ridge} = \underset{a_j}{\operatorname{argmin}} \left\| y - Xa \right\|_2^2 + \alpha_1 \left\| a \right\|_2^2 \quad (3.32)$$

e

$$b^{ridge} = \underset{b_k}{\operatorname{argmin}} \left\| z - Yb \right\|_2^2 + \alpha_2 \left\| b \right\|_2^2, \quad (3.33)$$

cujas soluções são dadas pela expressão em (3.20), obtida pelo método *Ridge Regression*, adaptada às matrizes X e Y de (3.30) e (3.31) respectivamente.

Fazendo, então, uma breve recapitulação: temos os seguintes conjuntos de dados $\{I(0), I(1), \dots, I(T-1)\}$, $\{R(0), R(1), \dots, R(T-1)\}$, $\{\beta(0), \beta(1), \dots, \beta(T-2)\}$ e $\{\gamma(0), \gamma(1), \dots, \gamma(T-2)\}$ e desejamos obter as estimativas $\beta(t)$ e $\gamma(t)$ no tempo $t = T-1$ e $I(t)$ e $R(t)$ em $t = T$. Usaremos a notação $\hat{\beta}(T-1)$, $\hat{\gamma}(T-1)$, $\hat{I}(T)$ e $\hat{R}(T)$ para representar estas respectivas previsões.

Substituindo os coeficientes a_j e b_k em (3.24) e (3.25) respectivamente pelas componentes dos vetores a^{ridge} e b^{ridge} e, em seguida, fazendo $t = T-1$ nas expressões resultantes, obtemos as estimativas $\hat{\beta}(T-1)$ e $\hat{\gamma}(T-1)$.

Para prevermos $\hat{I}(T)$ e $\hat{R}(T)$, fazemos $t = T - 1$ em (3.6) e (3.5) e, em seguida, trocamos $\beta(t)$ e $\gamma(t)$ por $\hat{\beta}(T - 1)$ e $\hat{\gamma}(T - 1)$, de modo que

$$\hat{I}(T) = [1 + \hat{\beta}(T - 1) - \hat{\gamma}(T - 1)]I(T - 1) \quad (3.34)$$

$$\hat{R}(T) = R(T - 1) + \hat{\gamma}(T - 1)I(T - 1). \quad (3.35)$$

Para obtermos as demais previsões $\hat{I}(t)$ e $\hat{R}(t)$, para $t = T + 1, T + 2, \dots, T + W$, em que W representa uma *janela de previsão*, devemos obter primeiramente as estimativas $\hat{\beta}(t)$ e $\hat{\gamma}(t)$, para $t = T, T + 1, \dots, T + W - 1$, cujos valores são obtidos fazendo-se $t = T, T + 1, \dots, T + W - 1$ nas equações (3.24) e (3.25), após substituirmos os coeficientes a_j e b_k nestas respectivas equações pelos coeficientes dos vetores a^{ridge} e b^{ridge} . Finalmente, procedendo de maneira similar ao que foi feito para a obtenção das expressões (3.34) e (3.35), obtemos as estimativas:

$$\hat{I}(t + 1) = [1 + \hat{\beta}(t) - \hat{\gamma}(t)]\hat{I}(t), \quad t \geq T, \quad (3.36)$$

$$\hat{R}(t + 1) = \hat{R}(t) + \hat{\gamma}(t)\hat{I}(t), \quad t \geq T. \quad (3.37)$$

As etapas detalhadas para o método de previsão são descritas no Algoritmo 1 a seguir.

Algoritmo 1: Método para previsão

Entrada: $\{I(t), R(t), 0 \leq t \leq T - 1\}$, para o cálculo das taxas
 $\{\beta(t), \gamma(t), 0 \leq t \leq T - 2\}$, parâmetros de regularização α_1 e α_2 ,
ordem dos filtros J e K e janela de previsão W

Saída: $\{\hat{\beta}(t), \hat{\gamma}(t), t \geq T - 1\}$ e $\{\hat{I}(t), \hat{R}(t), t \geq T\}$.

início

Determinar $\{\beta(t), \gamma(t), 0 \leq t \leq T - 2\}$, usando (3.7) e (3.8);
Executar a ridge regression, usando (3.26) e (3.27);
Estimar $\hat{\beta}(T - 1)$ e $\hat{\gamma}(T - 1)$, usando (3.24) e (3.25);
Estimar o número de pessoas infectadas $\hat{I}(T)$ e o número de pessoas recuperadas $\hat{R}(T)$ para o dia seguinte T , usando (3.34) e (3.35);

enquanto $T \leq t \leq T + W$ **faça**

Estimar $\hat{\beta}(t)$ e $\hat{\gamma}(t)$, usando (3.24) e (3.25);
Prever $\hat{I}(t + 1)$ e $\hat{R}(t + 1)$, usando (3.36) e (3.37).

fim

fim

4 Resultados Numéricos

Neste capítulo apresentamos os resultados numéricos obtidos para as previsões dos números de infectados e recuperados da Covid-19 no estado de Minas Gerais usando o modelo (3.2), visto no Capítulo 3. Tal modelo é ideal e leva em conta apenas o contexto inicial da pandemia em que alguns fatores, como casos de reinfeção, eram desconhecidos na época e ainda não havia vacina contra a doença. Assim, tomamos como base para nossas simulações numéricas dados referentes a diferentes períodos que se estendem até por volta de agosto do ano de 2020. Além disso, podemos considerar que toda a população é suscetível à doença neste período.

A suposição de que $S(t) \approx n$ em nosso modelo é aceitável tendo-se em vista que, já no final de 2020, foi constatado, através dos dados coletados em [13], que 97,5% da população no estado de Minas Gerais ainda não tinham se infectado. De fato, de acordo com os dados extraídos de [13], até a data de 01 de janeiro de 2021, foram registrados 542.909 casos de infectados no estado o que representa 2,5% de uma população total de 21.411.923 habitantes, conforme dados do IBGE.

4.1 Materiais

Para a análise e previsão da Covid-19, utilizamos o conjunto de dados disponibilizado no repositório [13] no formato CSV. O repositório contém um conjunto completo de informações, destacando-se o número notificado de casos e óbitos de Covid-19 nas esferas municipal e estadual do país, apresentados por data de divulgação dos boletins de cada unidade federativa. Descrições mais detalhadas de acesso a esse banco de dados podem ser encontrados em [12].

Para este trabalho extraímos o número de casos confirmados, o número de pessoas recuperadas e o número de óbitos ao longo do ano de 2020, no estado de Minas Gerais. Utilizamos a linguagem R para elaborar as séries diárias, incluindo a média móvel de 7 dias (Figuras 14, 15 e 16) para estes dados [25, 29].

Para a implementação do ridge regression e do Algoritmo 1, necessários para a obtenção das previsões, utilizamos a linguagem Octave [17]. As previsões são obtidas para o número total de infectados e recuperados (os dados de recuperados e de óbitos são tratados conjuntamente), em diferentes períodos de 2020, para uma janela de previsão $W = 7$ dias. Além disso, conforme definido em [8], utilizamos as ordens dos filtros FIR em (3.24) e (3.25) como $J = 3$ e $K = 3$ e os parâmetros de regularização para o método ridge regression como $\alpha_1 = 0,03$ e $\alpha_2 = 10^{-6}$.

4.2 Previsões

As Figuras 14, 15 e 16 apresentam as séries diárias para os casos positivos de Covid-19, os óbitos e os recuperados, respectivamente, registrados no estado de Minas Gerais durante o ano de 2020. A linha sobreposta é a respectiva média móvel de 7 dias.

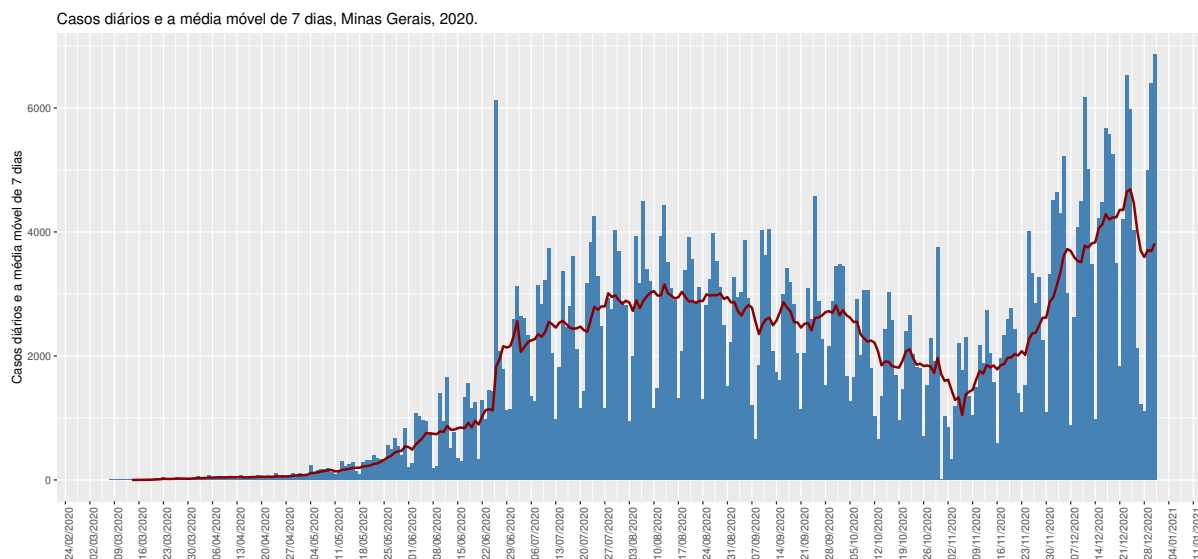


Figura 14

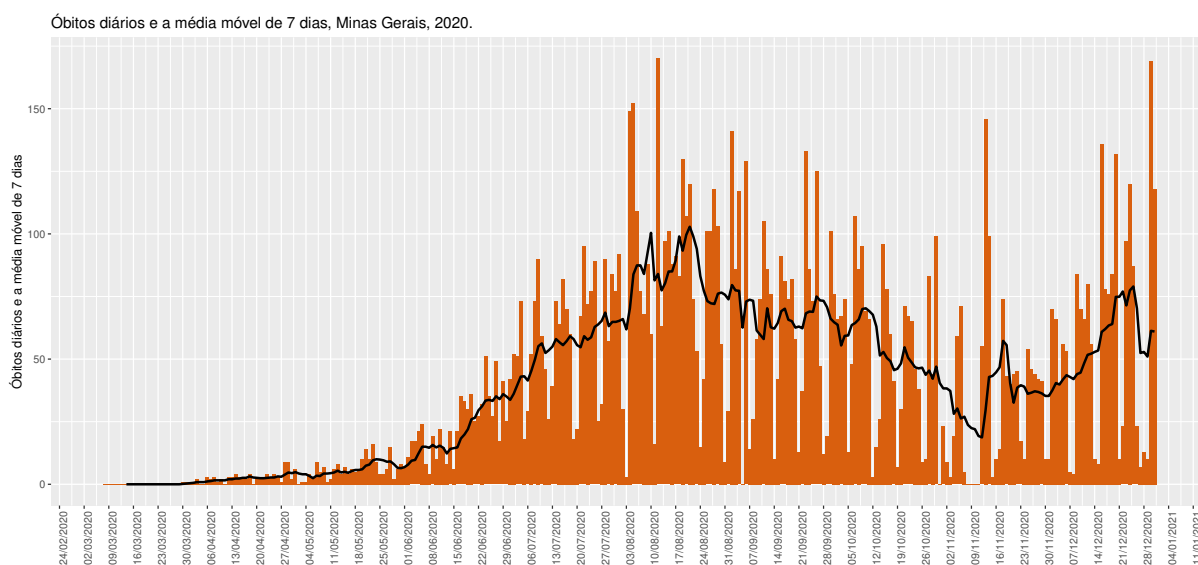


Figura 15

Nas Figuras 17, 18, 19 e 20 mostramos as evoluções temporais de infectados e de recuperados da Covid-19 para diferentes períodos, devidamente especificados em cada uma das figuras, ao longo do ano de 2020. Os pontos em vermelho representam os valores reais extraídos do conjunto de dados disponibilizados em [13]. As previsões são representadas pelos círculos azuis e foram feitas para os períodos compreendidos entre os dias 16 e 22

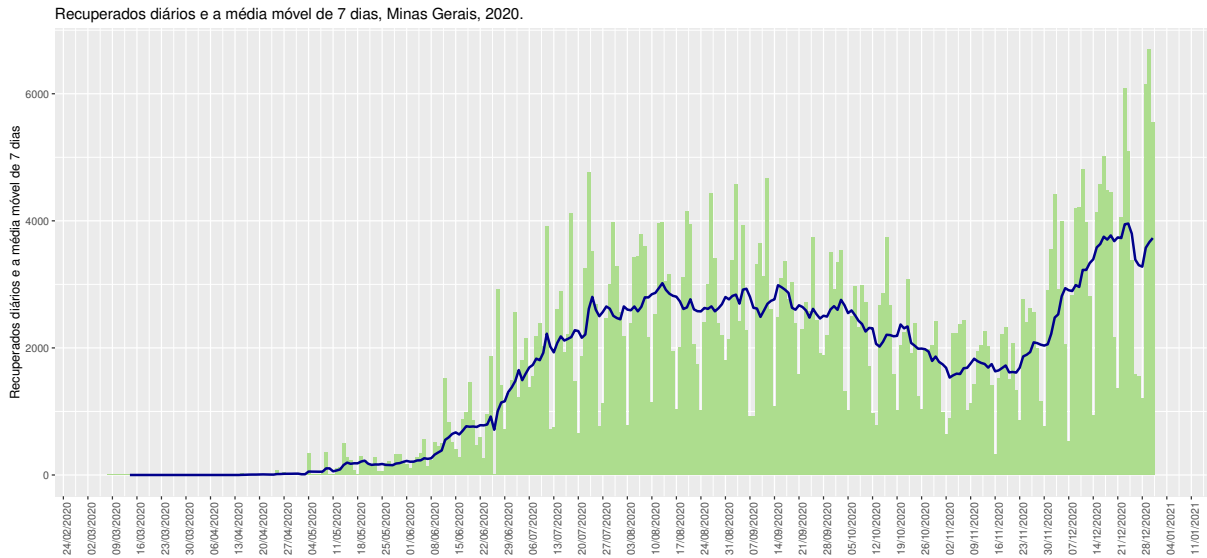


Figura 16

dos meses de maio, junho, julho e agosto, ou seja, com uma janela de previsão $W = 7$ dias em cada simulação.

Como podemos ver nas Figura 17 e 18, as previsões para os números de infectados acompanham a tendência real e são bem próximas do quantitativo real de infectados. Já as previsões para os números de recuperados acompanham a tendência real mas apresentam diferenças significativas para o quantitativo real de recuperados.

Por outro lado, nas Figura 19 e 20 observamos que essa análise se inverte: as previsões para os números de infectados apresentam diferenças significativas para o quantitativo real de infectados, enquanto as previsões para os números de recuperados se tornam mais próximas do quantitativo real de recuperados.

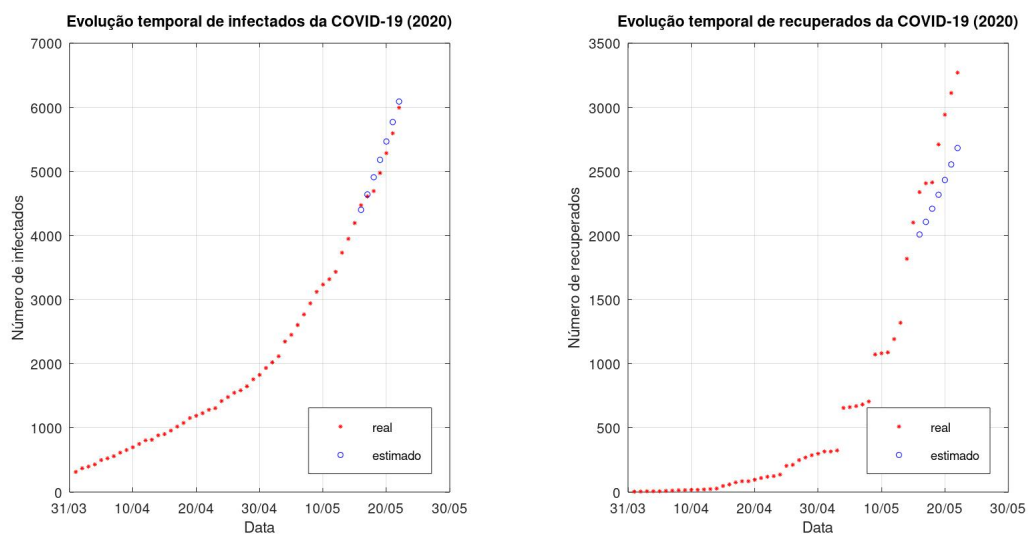


Figura 17 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de maio de 2020.

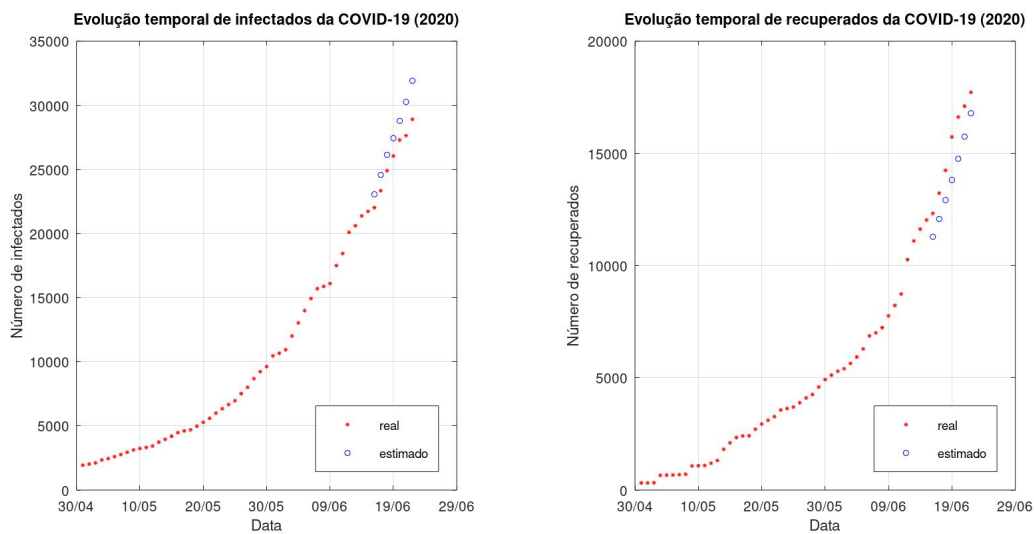


Figura 18 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de junho de 2020.

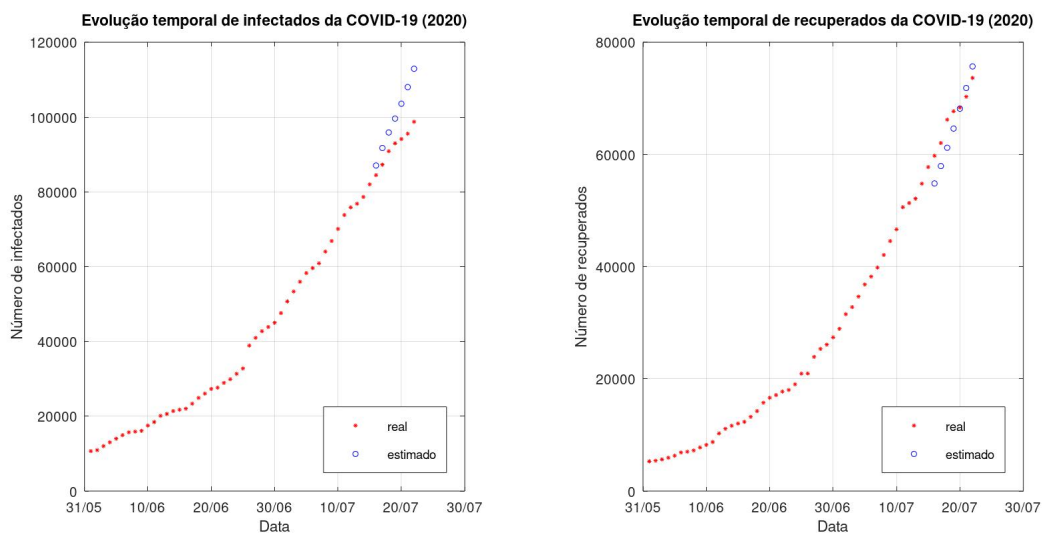


Figura 19 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de julho de 2020.

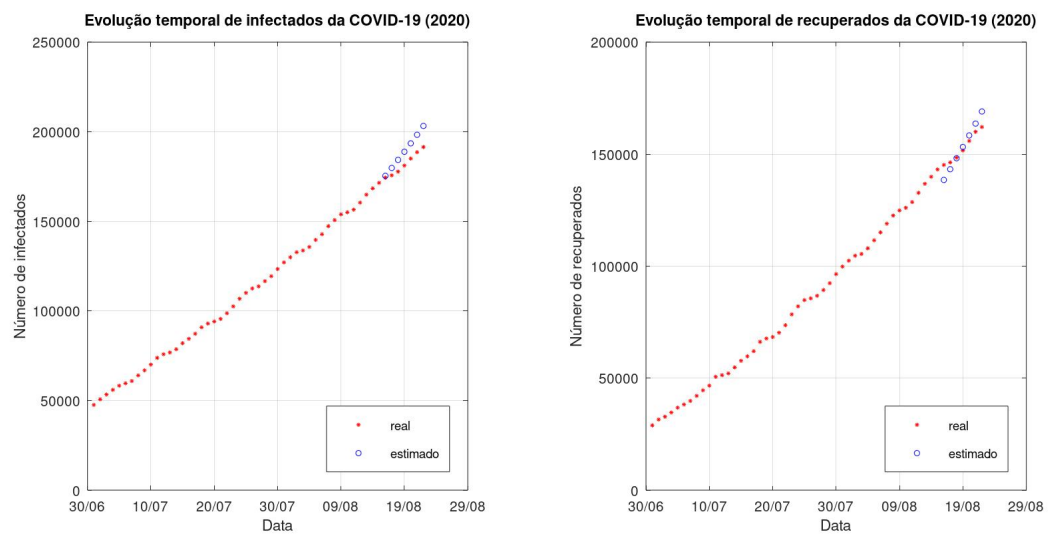


Figura 20 – Previsão para os números de infectados e recuperados para Minas Gerais de 16 a 22 de agosto de 2020.

CONSIDERAÇÕES FINAIS

O estudo apresentado neste trabalho alcançou o seu objetivo central que é o da obtenção das previsões para os números de infectados e recuperados da Covid-19 no estado de Minas Gerais, num contexto inicial da pandemia em que não havia vacina e os casos de reincidência da doença eram desconhecidos. Para atingir este objetivo, revisão de importantes conteúdos da literatura foi necessária tais como modelos matemáticos bidimensionais descritos por sistemas de equações diferenciais, regressão linear e quadrados mínimos lineares, além do próprio modelo discretizado *Time-dependent* SIR descrito no Capítulo 3.

Como observado nos resultados apresentados no Capítulo 4, as simulações realizadas com a implementação do algoritmo 1 produziram resultados numéricos estimados que, em alguns períodos, se aproximaram dos dados reais e em outros, se afastaram. Para minimizar esses resultados indesejados, estamos trabalhando na execução de algumas variações do modelo proposto pelos taiwaneses em [8] cujos resultados têm se mostrado bastante promissores, o que nos motiva na elaboração de um artigo científico para futura publicação.

Por fim, destacamos que o modelo proposto neste trabalho pode servir de inspiração para o estudo e a elaboração de modelos mais complexos num contexto em que se leve em conta o número de mortos, a vacinação, as reinfecções, entre outros. Como exemplo de variação do método proposto pelos taiwaneses, citamos Teran que, em seu artigo [10], separa a taxa de mortalidade da de recuperação usando o modelo SIRD (*Susceptibles-Infected-Recovered-Dead*) e considera que as infecções não ocorrem imediatamente após o contato entre infectados e suscetíveis, ou seja, que há um retardo para o surgimento de infectados. Interpretações similares foram feitas para suscetíveis, recuperados e mortos.

Referências

- [1] ALVES, Rafael G.; OLIVEIRA, Monique R. A. (2021). *Modelagem matemática de sistemas epidemiológicos: um modelo SEIR para o sarampo*. Revista de Matemática. RMAT, 2021. ISSN: 2237-8103.
- [2] BASSANEZI, Rodney C. *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. São Paulo, SP: Editora Contexto, 2002.
- [3] BOYCE, William E.; DIPRIMA, Richard C.; MEADE, Douglas B. *Elementary differential equations and boundary value problems*. 10ed. John Wiley & Sons, 2012.
- [4] BRASIL. 2018. *Base Nacional Comum Curricular*. Ministério de Educação. Brasília: MEC.
- [5] BRASIL. *Lei nº 9.394, de 20 de dezembro de 1996*. Estabelece as diretrizes e bases da educação nacional. Diário Oficial da União, Brasília, 23 de dezembro de 1996. Disponível em: < http://www.planalto.gov.br/ccivil_03/leis/L9394.htm >. Acesso em: 04 mar. 2023.
- [6] BRASIL. 2000. *Parâmetros Curriculares Nacionais: Ensino Médio*. Ministério de Educação. Brasília: MEC.
- [7] CAZORLA, Irene M. et al. *Reflexões sobre o papel da Educação Estatística na formação de professores no contexto da pandemia da COVID-19*. *Jornal Internacional de Estudos em Educação Matemática*, v. 13, n. 4, p. 437-449, 2020.
- [8] CHEN, Yi C.; LU, Ping E.; CHANG, Cheng S.; LIU, Tzu H. *A time-dependent SIR model for COVID-19 with undetectable infected persons*. *IEEE transactions on network science and engineering*, v. 7, n. 4, p. 3279-3294, 2020.
- [9] CIOTTI, Marco et al. *COVID-19 Outbreak: An Overview*. *Chemotherapy*. 2020. DOI 10.1159/000507423. Disponível em < <https://www.arca.fiocruz.br/handle/icict/40691> >. Acesso em 25 fev. 2022.
- [10] TERAN, Edson A. C. *A COVID-19 time-dependent SIRD model using functional differential equations*. *International Journal of Ecology & Development*, p. 1-11, 2021.
- [11] CORDEIRO, Paulo G. R. *Sistemas autônomos bidimensionais: definições, exemplos e aplicações*. 2018. Monografia (Licenciatura em Matemática), UFOP (Universidade Federal de Ouro Preto), Ouro Preto, Brasil.

- [12] COTA, Wesley. *Monitoring the number of COVID-19 cases and deaths in brazil at municipal and federative units level*, SciELOPreprints:362 (2020), DOI 10.1590/scielo-preprints.362.
- [13] COTA, Wesley. *COVID-19 data in Brazil: cases, deaths, and vaccination at municipal (city) level*. Disponível em <<https://github.com/wcota/covid19br>>. Acesso em 8 mar. 2021.
- [14] CUNHA, Maria C. *Métodos numéricos*. 2ed. Campinas, SP: Editora da Unicamp, 2000.
- [15] DATASENADO. *Impactos da pandemia na Educação no Brasil*. DataSenado: Brasília, 2022. Disponível em <<https://www12.senado.leg.br/institucional/datasenado/materias/pesquisas/impactos-da-pandemia-na-educacao-no-brasil>>. Acesso em 04 mar. 2023.
- [16] FIOCRUZ. *Impactos sociais da pandemia*. Fundação Oswaldo Cruz: Rio de Janeiro. Disponível em <<https://portal.fiocruz.br/impactos-sociais-economicos-culturais-e-politicos-da-pandemia>>. Acesso em 04 mar. 2023.
- [17] GNU General Public License. GNU Octave. Disponível em: <<https://www.gnu.org/software/octave/>>.
- [18] GOLUB, Gene H.; VAN LOAN, Charles F. *Matrix computations*. 4ed. Maryland: The Johns Hopkins University Press, 2013.
- [19] HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. *The elements of statistical learning: data mining, inference, and prediction*. 2ed. Nova York, Springer, 2009.
- [20] JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. *An introduction to statistical learning: with applications in R*. 2ed. New York: springer, 2021.
- [21] LEYKEKHMAN, Dmitriy. *Lecture 10. Regularized Linear Least Squares*. MATH 3795: Intoduction to Computational Math, Fall 2008. Disponível em <https://www2.math.uconn.edu/~leykekhman/courses/MATH3795/Lecture_slides.html>. Acesso em 11 abr. 2022.
- [22] MARGALIT, Dan; RABINOFF, Joseph; ROLEN, Larry. *Interactive linear algebra*. Georgia Institute of Technology, 2019.

- [23] OMS AFIRMA QUE COVID-19 É AGORA CARACTERIZADA COMO PANDE-
MIA. Organização Pan-Americana da Saúde - OPAS. 2020. Disponível em <[https://
www.paho.org/pt/news/11-3-2020-who-characterizes-covid-19-pandemic](https://www.paho.org/pt/news/11-3-2020-who-characterizes-covid-19-pandemic)>.
Acesso em 14 jan. 2021.
- [24] PHASE PLANE ANALYSIS. The Free Dictionary, 2022. Disponível em <[https://
encyclopedia2.thefreedictionary.com/Phase+Plane+Analysis](https://encyclopedia2.thefreedictionary.com/Phase+Plane+Analysis)>. Acesso em
21 fev. 2022.
- [25] R CORE TEAM. *R: A language and environment for statistical computing*. R Vienna,
Austria: Foundation for Statistical Computing, 2022. Disponível em <[https://www.
R-project.org/](https://www.R-project.org/)>
- [26] SILVA, Victor G. R. da; DUARTE, Dami D. N. *Filtragem de sinais cardíacos utilizando
filtros FIR*. Interfaces Científicas-Exatas e Tecnológicas, v. 3, n. 3, p. 65-74, 2019.
- [27] SOTOMAYOR, Jorge. *Lições de equações diferenciais ordinárias*. Instituto de Mate-
mática Pura e Aplicada, CNPq, 1979.
- [28] UMBELINO, Moacir; ZABINI, Franciele O. *A importância da interdisciplinaridade
na formação do docente*. Seminário Internacional de Educação Superior, v. 1, 2014.
- [29] WICKHAM, Hadley. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag
New York, 2016. Disponível em <<https://ggplot2.tidyverse.org>>.