

# Estimativa de Parâmetros do Modelo SIR Baseada na Dinâmica da População Infectada por COVID-19

Lucas S. Paiva<sup>1</sup> Marina R. B. Dias<sup>2</sup>

Departamento de Matemática - ICEx - UFF, Volta Redonda, RJ

Os modelos compartimentados SIR [2] foram amplamente utilizados durante a pandemia de COVID-19, em 2020, com objetivo de modelar a doença para auxiliar as autoridades nas tomadas de decisões. Uma das dificuldades em seu uso foi na determinação dos parâmetros envolvidos. Este trabalho propõe estimar os parâmetros do modelo SIR por meio da técnica proposta em [4], que estimou parâmetros a partir da análise do comportamento da população de infectados levando em consideração dois momentos: a fase inicial da doença (onde há crescimento da população) e a fase avançada, em que o número de infectados começa a decrescer.

As equações do modelo SIR são dadas pela Eq. (1), onde se considera a população de: suscetíveis,  $S(t)$ ; infectados,  $I(t)$ ; e removidos/recuperados,  $R(t)$ , ou seja,

$$\begin{cases} \frac{dS}{dt} = -\frac{\beta SI}{N} \\ \frac{dI}{dt} = \frac{\beta SI}{N} - \alpha I \\ \frac{dR}{dt} = \alpha I \end{cases} \quad (1)$$

onde  $\beta > 0$  e  $\alpha > 0$  são as taxas de contágio e de recuperação, respectivamente.  $S(t_0) = S_0$  e  $I(t_0) = I_0$  representam as populações de suscetíveis e infectados no instante inicial  $t_0 \in \mathbb{R}^+$ , respectivamente. Supõe-se que  $R(t_0) = 0$  e a taxa de reprodutibilidade basal é dada por  $R_0 = \frac{\beta}{\alpha}$ . Sua dinâmica é amplamente conhecida na literatura e está representada na Fig. 1a, com  $I_0 = 2$ ,  $N = 100.000$ ,  $\beta = 1,5$  e  $R_0 = 2$ . Há um total de  $N$  indivíduos, e a soma de cada população é estável, ou seja,  $S + I + R = N$ .

Os parâmetros foram estimados conforme o proposto em [4]:

- i) Inicialmente ( $N \gg I_0$ ),  $I_c(t)$  cresce até um valor máximo  $I_{max}$ . Pode-se determinar  $I_c(t) = I_0 e^{\gamma t}$ , sendo  $\gamma = \beta - \alpha$ , e o valor máximo será dado por  $I_{max} = \frac{N}{R_0}(R_0 - 1 - \log R_0)$ .
- ii) No fim da epidemia, cada população atinge um valor limite,  $S_e$ ,  $I_e$  e  $R_e$ . Supondo que a doença foi extinta, isto é,  $I_e = 0$ , e a seguinte relação é válida:  $S_e = \frac{N}{R_0} \log\left(\frac{S_e}{N}\right) + 1$ .
- iii) Como  $S + I + R = N$ , quando  $I_e = 0$ , tem-se  $\frac{R_e}{N} + \frac{S_e}{N} = 1$  e, portanto,  $\frac{R_e}{N} = -\frac{1}{R_0} \log\left(\frac{S_e}{N}\right)$ .
- iv) A expressão analítica que modela o decrescimento de  $I(t)$  é dada por  $I_d(t) = C e^{-\Gamma(t-t_0)}$ , com  $\Gamma = -\frac{\beta S_e}{N} + \alpha$  e  $C \in \mathbb{R}^+$ .
- v) Os parâmetros, portanto, podem ser estimados por meio de  $\tau = \frac{\gamma}{\Gamma} = \frac{R_0 - 1}{1 - R_0 \frac{S_e}{N}}$ ,  $\beta = \frac{\gamma}{1 - \frac{1}{R_0}}$  e das relações anteriores.

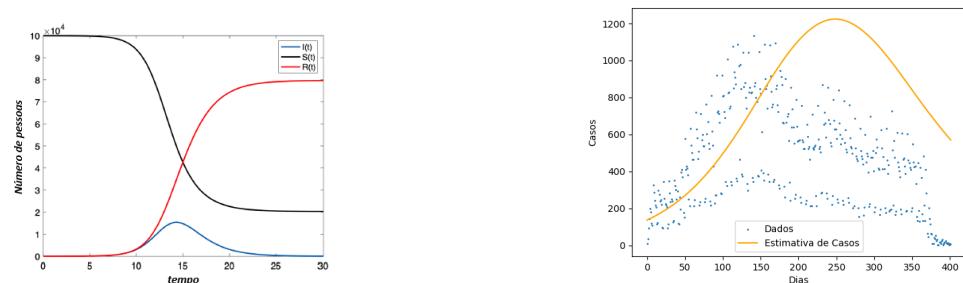
As funções  $I_c(t)$  e  $I_d(t)$  são encontradas usando uma escala logarítmica e mediante ajuste linear encontramos  $\gamma$  e  $\Gamma$ . Quanto a  $\frac{S_e}{N}$ , é necessário definir  $f(R_0) = \tau_e - \tau(R_0)$  e  $\tau(R_0) = \frac{R_0 - 1}{1 - R_0 \frac{S_e}{N}}$ , com  $\tau_e$  sendo o valor de  $\tau$  encontrado através de  $I_c(t)$  e  $I_d(t)$ . Assim pode-se calcular

<sup>1</sup>lucassilvapaiva@id.uff.br

<sup>2</sup>marinaribeiro@id.uff.br

numericamente  $\frac{S_\infty}{N}$  usando o método de Newton-Raphson [1] e minimizar  $f(R_0)$ , encontrando  $R_0$ . Já o valor de  $I_{\max}$  pode ser encontrado analisando o conjunto de dados.

Para esse trabalho os dados (dez/23 a fev/25) foram extraídos de [3]. O código-fonte foi elaborado pelos autores e está disponível como software do código aberto sob licença MIT no GitHub (<https://github.com/Lucas076993/SIR-Parameters-Estimation>). Para a simulação do modelo foi utilizado o método de Runge-Kutta de quarta ordem [1].



(a) Dinâmica do modelo SIR. Fonte: [4].

(b) Simulação de  $I(t)$ . Fonte: Autores.

Figura 1: Simulação do modelo SIR, dados reais e resultados numéricos obtidos.

Após as simulações foram obtidos:  $N = 17.490,72$ ,  $\alpha = 2,8e^{-2}$ ,  $\beta = 4,2e^{-2}$  e  $I_0 = 136,75$ . Além disso, o ajuste das curvas teve um erro quadrático médio de 0,92 para  $I_c(t)$  e 9,85 para  $I_d(t)$ . A simulação apresenta erro quadrático médio de 259.063,92. O resultado da simulação pode ser visto na Fig. 1b.

Num cômputo geral, o resultado indica a tendência do crescimento e decrescimento do número de infectados ao longo do período avaliado. Entretanto, nota-se que os dados coletados apresentam flutuação, o que pode ter contribuído negativamente para o cálculo de  $N$ . Foram obtidos, portanto,  $\alpha = 2,8e^{-2}$  e  $\beta = 4,2e^{-2}$ . Para análises mais precisas, uma possibilidade seria fazer um tratamento nos dados antes de realizar a estimativa, considerando, por exemplo, a média móvel da população.

## Agradecimentos

À Fundação Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro (FAPERJ) pelo financiamento (processo E-26/201.241/2024) e ao Instituto de Ciências Exatas, pela estrutura.

## Referências

- [1] M. C. C. Cunha. **Métodos Numéricos**. Campinas, SP: Editora da Unicamp, 2000. ISBN: 8526805215.
- [2] W. O. Kermack e A. G. McKendrick. “A Contribution to the Mathematical Theory of Epidemics”. Em: **Proc. of the Royal Society of London A**. Vol. 115. 1927, pp. 700–721. DOI: <https://doi.org/10.1098/rspa.1927.0118>.
- [3] BRASIL. MINISTÉRIO DA SAÚDE. **openDataSUS - SRAG 2021 a 2024 - Banco de Dados de Síndrome Respiratória Aguda Grave - incluindo dados da COVID-19**. Online. Acessado em 06/03/2025, <https://opendatasus.saude.gov.br/dataset/srag-2021-a-2024>.
- [4] F. G. Schmitt. “An algorithm for the direct estimation of the parameters of the SIR epidemic model from the  $I(t)$  dynamics.” Em: **Eur. Phys. J. Plus** 137, 57 (2022). DOI: <https://doi.org/10.1140/epjp/s13360-021-02237-7>.