Modelagem Matemática de uma Epidemia

André Nepomuceno[1]

[1]Universidade Federal Fluminense, Departamento de Ciências da Natureza, Rio das Ostras, RJ, Brasil.

Abstract

Este artigo investiga aspectos da modelagem matemática de doenças infecciosas, tomando como exemplo o modelo SIR (Suscetível-Infectado-Removido). Focamos principalmente na análise numérica, ilustrando como o método de Runge-Kutta pode ser aplicado a solução de um sistema de equações diferenciais não-lineares acopladas. As previsões e limitações do modelo são discutidas. Uma animação interativa, que mostra aspectos qualitativos de uma epidemia, foi desenvolvida e é apresentada no texto.
Palavras-chave: SIR, Runge-Kutta, epidemia.

This article investigates the mathematical modeling of infectious diseases, in the framework of the SIR (Susceptible-Infectious-Removed) model. We focus on the numerical analysis to illustrate how the Runge-Kutta method can be applied to solve a set of coupled nonlinear ordinary differential equations. The model predictions and their limitations are discussed. An interactive animation, that shows qualitative aspects of an epidemic, was developed.
Key-words: SIR, Runge-Kutta, epidemic

1 Introdução

A modelagem matemática da propagação de doenças infecciosas é uma interessante área interdisciplinar que envolve a interação entre epidemiologista, matemáticos, físicos, biólogos e cientista da computação. Seu papel tem se mostrado de grande importância ao oferecer direções as medidas de saúde pública frente as várias epidemias que assolaram a humanidade. A recente pandemia de COVID-19, por exemplo, despertou considerável atenção a essa área, e diferentes modelos matemáticos estão sendo estudados e aplicados para o entendimento dos efeitos das medidas de contenção e da propagação da doença [1, 2, 3, 4, 5, 6].

Uma das primeiras tentativas de modelar matematicamente doenças infecciosas foi feita pelo físico e matemático Daniel Bernoulli. Mais conhecido entre físicos por seus trabalhos sobre hidrodinâmica, Bernoulli desenvolveu um modelo de propagação da varíola, doença que causou milhares de mortes por ano na Europa do século XVIII. Bernoulli utilizou seus resultados para defender uma técnica de vacinação rudimentar conhecida como variolação, levando-o a um intenso debate com o matemático francês Jean-Baptiste d'Alembert, que se opunha a técnica e questionava seus benefícios [7]. A partir do início do século XX, diversos modelos epidemiológicos foram propostos para tentar entender a dinâmica de epidemias de sarampo, malária, influenza, tuber-

culose, AIDS, dentre outras [8].

Os modelos epidemiológicos modernos englobam modelos determinísticos e estocásticos, e empregam uma variedade de técnicas como simulações computacionais e cadeias de Markov. Muitos desses modelos incluem aspectos importantes na descrição da epidemia, como a dependência das infecções com a idade, o efeito das campanhas de vacinação e imunidade dos indivíduos limitada a um determinado intervalo de tempo. Para o leitor interessante nos detalhes matemáticos dos diversos modelos, sugerimos os artigos de revisão [9, 10].

Neste trabalho, vamos explorar, de forma didática e utilizando um dos modelos mais simples, como modelos epidemiológicos são construídos e que tipo de previsão são capazes de fazer. O método de Runge–Kutta (RK) será empregado para a solução numérica das equações diferencias do modelo. Essa técnica numérica é geralmente estudada nas disciplinas de física computacional e aplicada a solução de equações que descrevem oscilações ou sistemas sujeitos a forças dependentes da velocidade ou posição [11]. Nesta nota, vamos ilustrar como o método de RK pode ser aplicado a solução de um conjunto de equações diferenciais acopladas, num contexto interdisciplinar.

Os códigos utilizados nas análises numéricas foram escritos em Python e estão disponíveis, com comentários detalhados, em `https://github.com/aanepomuceno/sir_rk4`. Uma animação interativa, que mostra aspectos qualitativos do modelo epidemiológico estudado, também foi desenvolvida e está detalhada na última seção.

2 O Modelo SIR

Para modelar a dinâmica de uma epidemia, vamos considerar uma população formada por um grande número N de indivíduos. Suponhamos que um ou mais indivíduos contraiam uma doença infecciosa causada por um agente biológico (vírus ou bactéria), e que a doença é transmitida pelo contato entre os infectados e os inicialmente saudáveis. Podemos então dividir a população em três grupos ou compartimentos: suscetíveis a doença (S), infectados (I) e removidos (R). Os indivíduos da categoria "removidos" são aqueles que se recuperaram da doença e ganharam imunidade permanente. O número de indivíduos em cada categoria num instante t é denotado por $S(t)$, $I(t)$ e $R(t)$, respectivamente. Vamos assumir também que durante a ocorrência da epidemia não há nascimentos ou mortes, de modo que $N = S(t)+I(t)+R(t)$ é constante. O número de pares de indivíduos suscetíveis e infectados é SI. Se μ é a probabilidade por unidade de tempo do encontro do par, então teremos μSI encontros entre suscetíveis e infectados, por unidade de tempo. Se uma fração τ dos encontros resulta em contágio, então o número de indivíduos na categoria S que são infectados, na unidade de tempo, é $\tau\mu SI$. Sendo $\beta = \tau\mu$, podemos escrever a taxa de variação de S como

$$\frac{dS}{dt} = -\beta SI. \tag{1}$$

O parâmetro β pode ser interpretado como a taxa de transmissão da doença. O mesmo número de indivíduos que sai da categoria S entra na categoria I. Se γ é a taxa de recuperação, então, em média, γI indivíduos saem do compartimento I e entram na categoria R, por unidade de tempo. Portanto, as variações em I e R serão:

$$\frac{dI}{dt} = \beta SI - \gamma I \tag{2}$$

$$\frac{dR}{dt} = \gamma I. \tag{3}$$

Esse sistema de equações diferenciais não lineares acopladas é conhecido como modelo SIR (Suscetível-Infectado-Removido). Não há

solução analítica explícita para esse sistema, devendo ser resolvido numericamente.[1]

Podemos utilizar a Eq.(2) para determinar em que condições ocorrerá uma epidemia. Se os números iniciais de suscetíveis e infectados são, respectivamente, $S(0) \approx N$ e $I(0) \ll N$, então para que $I(t)$ inicialmente aumente com o tempo devemos ter

$$\beta S(0)I(0) - \gamma I(0) > 0 \tag{4}$$

$$\frac{\beta}{\gamma} S(0) > 1. \tag{5}$$

O termo $R_0 = (\beta/\gamma)S(0)$ é chamado número básico de reprodução, e indica quantos indivíduos serão contagiados por cada indivíduo infectado no início da epidemia. Quanto maior o valor de R_0, mais rápido a doença se espalha. Se $R_0 < 1$, o número inicial de infectados decai com o tempo e não há epidemia. Utilizando a definição de R_0, podemos escrever a Eq.(2), em $t = 0$, como

$$\left.\frac{dI}{dt}\right|_{t=0} = \gamma(R_0 - 1)I(0), \tag{6}$$

indicando que inicialmente o número de infectados cresce (ou decresce) de forma exponencial a uma taxa $\gamma(R_0 - 1)$.

3 Solução pelo Método de Runge-Kutta

Os métodos de Runge-Kutta são métodos numéricos de solução de equações diferenciais ordinárias (EDOs), sendo o mais utilizado o método de Runge-Kutta de 4ª ordem (RK4). A solução é construída partir do cálculo de uma sequência de pequenos passos de tamanho h, obtendo uma precisão de $\mathcal{O}(h^4)$. Para um conjunto de m equações diferencias escritas na forma

$$\frac{d\vec{y}}{dt} = \vec{f}(\vec{y}, t), \tag{7}$$

com $\vec{y} = (y_1, y_2, ..., y_m)$ e $\vec{f} = (f_1, f_2, ..., f_m)$, a solução pelo método RK4 é:

$$\vec{k}_1 = h\vec{f}(\vec{y}_n, t_n), \tag{8}$$

$$\vec{k}_2 = h\vec{f}(\vec{y}_n + \frac{\vec{k}_1}{2}, t_n + \frac{h}{2}), \tag{9}$$

$$\vec{k}_3 = h\vec{f}(\vec{y}_n + \frac{\vec{k}_2}{2}, t_n + \frac{h}{2}), \tag{10}$$

$$\vec{k}_4 = h\vec{f}(\vec{y}_n + \vec{k}_3, t_n + h), \tag{11}$$

$$\vec{y}_{n+1} = \vec{y}_n + \frac{1}{6}(\vec{k}_1 + 2\vec{k}_2 + 2\vec{k}_3 + \vec{k}_4), \tag{12}$$

onde $n = 0, 1, 2, 3....$

Para aplicar o método RK4 a solução das Eqs.(1)-(3), notemos que $\vec{y} = (S, I, R)$ e $\vec{f} = (f_S, f_I, f_R) = (-\beta SI, \beta SI - \gamma I, \gamma I)$. As expressões para o cálculo de S_{n+1}, I_{n+1} e R_{n+1} estão no apêndice.

Vamos utilizar essa técnica para investigar a propagação de uma doença infecciosa numa população de $N = 100000$ indivíduos. Para o cálculo, consideramos as condições iniciais $I(0) = 2$, $S(0) = N - 2$ e $R(0) = 0$ (não confundir com R_0). O intervalo de tempo é de 120 dias com um passo de $h = 0.1$ dia. A Figura 1 mostra o gráfico das soluções $S(t)$, $I(t)$ e $R(t)$ para três valores de R_0: 5.0, 2.5, e 2.0.

No cenário com $R_0 = 5$, vemos que a doença se propaga rapidamente, com o número de infectados atingindo o valor de pico de aproximadamente metade da população em menos de vinte dias. A curva de $R(t)$ indica que, nesse cenário, toda a população foi infectada (e se recuperou ao final do período). Para $R_0 = 2.5$ e $R_0 = 2.0$, vemos que a doença se espalha de forma mais lenta, e o pico de infectados é menos acentuado. As medidas de enfrentamento a uma epidemia visam exatamente diminuir a taxa de transmissão da doença.

É interessante notar também que, dependendo do valor de R_0, a epidemia termina antes

[1] Uma solução paramétrica foi proposta em 2014 por T. Harko *et al* [12].

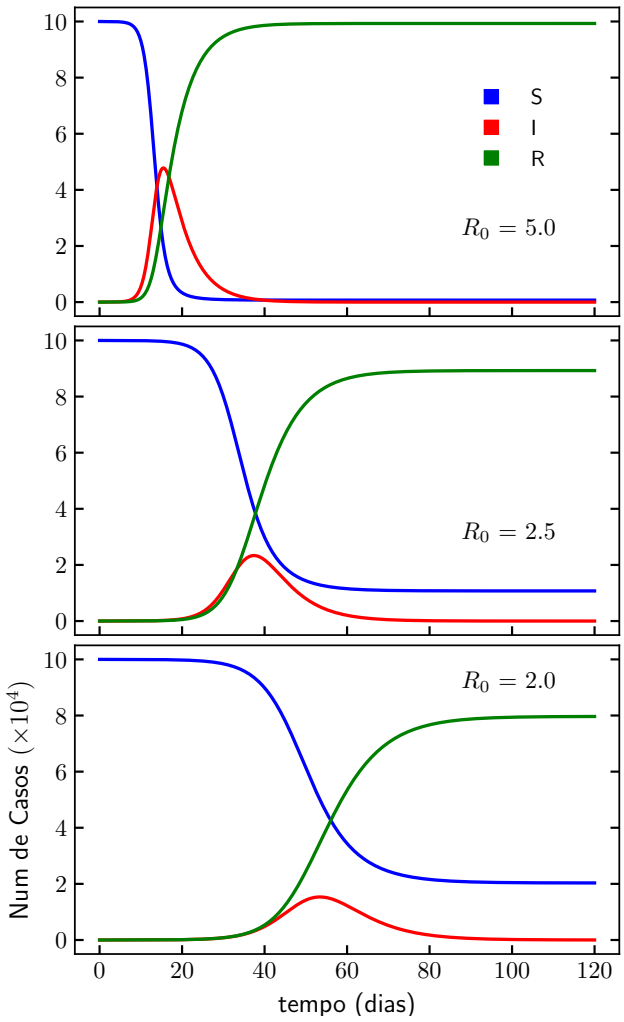

Figure 1: Evolução temporal do número de indivíduos suscetíveis (S), infectados (I) e removidos (R) numa população sujeita a uma doença infecciosa, para três valores do número básico de reprodução. Os valores de R_0 foram obtidos fixando $\gamma = 0.2$ e variando apenas o valor de β.

que todos os suscetíveis seja infectados, como ilustra a Figura 1 nos gráficos para $R_0 = 2.5$ e $R_0 = 2.0$. A Figura 2 mostra o resultado do cálculo numérico da fração de suscetíveis no final da epidemia ($S(\infty)/N$) para vários valores de R_0. Para $R_0 = 2.0$, por exemplo, vemos que 20% dos suscetíveis não são infectados. Isso ocorre porque a medida que $S(t)$ diminui, a taxa em que novas infecções ocorrem também diminui. Num determinado ponto, a taxa de recuperação supera a taxa de infeção, e $I(t)$ começa a decrescer.

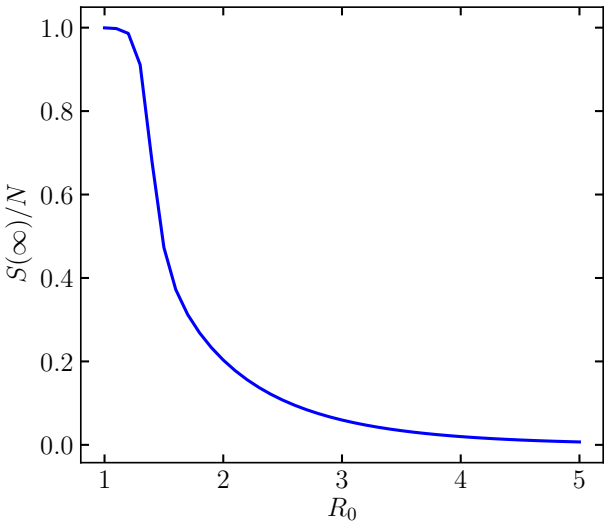

Figure 2: Fração de indivíduos suscetíveis ao final de epidemia em função de R_0.

4 Efeito da Vacinação

O modelo SIR pode ser facilmente estendido para estudarmos o efeito da vacinação. Nesse caso, vamos supor que uma vacina foi desenvolvida e que uma fração de indivíduos suscetíveis seja vacinada ao longo dos 120 dias da epidemia. Se taxa de vacinação é ν, então νS indivíduos sairão da categoria S e entrarão na nova categoria vacinados (V). Assumindo que a vacina ofereça imunização permanente, as equações para (dI/dt) e (dR/dt) permanecerem inalteradas, e as variações em S e V serão:

$$\frac{dS}{dt} = -\beta SI - \nu S, \tag{13}$$

$$\frac{dV}{dt} = \nu S, \tag{14}$$

$$N = S(t) + I(t) + R(t) + V(t). \tag{15}$$

Considerando as mesmas condições iniciais da Sec. (3), a solução numérica do novo conjunto de EDOs é mostrada na Figura 3 para $R_0 = 2.5$ e dois valores diferentes de ν (a curva $V(t)$ não é mostrada). Como esperado, o modelo prevê que a inclusão da vacinação retarda a propagação do doença e diminui consideravelmente o total de infectados.

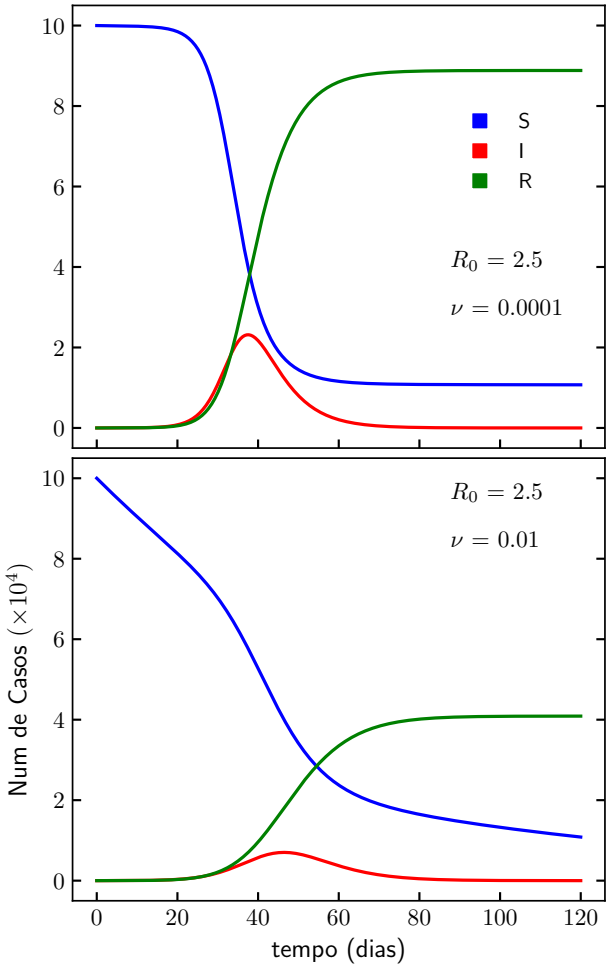

Figure 3: Soluções $S(t)$, $I(t)$ e $R(t)$ para $R_0 = 2.5$ e dois valores de ν. Em qualquer instante, o valor de $V(t)$ pode ser calculado pela Eq.(15).

A Figura 4 mostra como a fração de infectados, durante todo o período da epidemia, varia com R_0 e ν. Para pequenos valores do número básico de reprodução ($R_0 < 1.5$), vemos que a epidemia pode ser controlada com um baixo valor de ν. Para maiores valores de R_0, como por exemplo $2.5 < R_0 < 3.5$, uma taxa de vacinação de $\nu = 0.01$ produz uma redução no total de infectados de aproximadamente 30%.

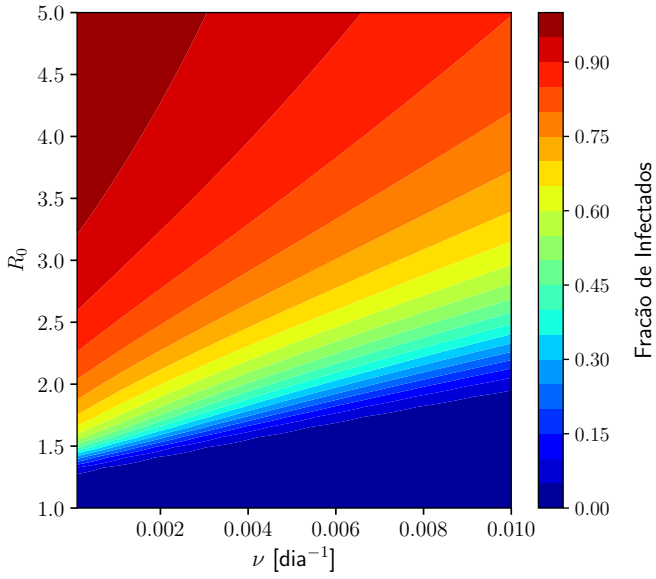

Figure 4: Fração total de infectados em diferentes combinações dos parâmetros R_0 e ν.

Como última análise, vamos investigar o cenário onde uma fração ρ da população é vacinada *antes* da epidemia, e não há vacinação durante ($\nu = 0$). O número de indivíduos na categoria vacinados é portanto $\rho S(0)$, número que se mantém constante. Então, o número inicial de indivíduos suscetíveis passa a ser $(1-\rho)S(0)$. De acordo com a Eq. (5), a epidemia não ocorrerá se:

$$\frac{\beta}{\gamma}(1-\rho)S(0) \leq 1, \tag{16}$$

ou

$$\rho \geq 1 - \frac{1}{R_0}. \tag{17}$$

Assim, para uma epidemia ser evitada, deve-se vacinar uma fração mínima de $\rho_{\text{min}} = 1 - 1/R_0$ indivíduos, o que faz com que a população atinja a chamada "imunidade de rebanho". Essa previsão interessante da mode-

lagem matemática indica que não é necessário vacinar toda a população para prevenir uma epidemia. Para $R_0 = 2.5$, esse modelo indica que 60% dos suscetíveis devem ser vacinados para um bloqueio eficiente.

Para verificar esse cenário, resolvemos novamente o sistema de EDOs pelo método RK4 para três valores de ρ: 0.3, 0.5 e 0.6. Vamos considerar também condições iniciais ligeiramente diferentes das anteriores: $I(0) = 100$, $R(0) = 0$, e $S'(0) = (1-\rho)N - I(0)$, onde $S'(0)$ é o novo número de indivíduos inicialmente suscetíveis. A solução para o número de infectados é mostrada na Figura 5, com $R_0 = 2.5$. No gráfico ampliado no canto superior direito da figura, vemos que quando $\rho = 0.6$, o número de infectados decai com o tempo e a epidemia não se desenvolve.

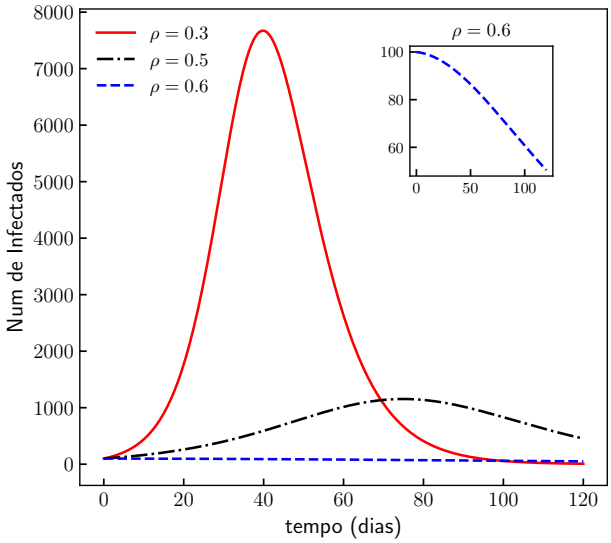

Figure 5: Evolução temporal do número de infectados com $R_0 = 2.5$ para três valores de ρ.

É interessante mencionar que o modelo SIR não leva em conta situações com longos períodos de incubação da doença, ou seja, o período em que o indivíduo está infectado mas não transmite (é o caso da catapora, por exemplo.) Para modelar essa situação, podemos introduzir uma nova categoria E (exposto), de modo que indivíduos infectados que saem da categoria S entram na categoria E, e após o período de incubação, passam para a categoria I. Essa extensão do modelo SIR é conhecida como modelo SEIR (Suscetível-Exposto-Infectado-Removido). O efeito do período incubação será o atraso no crescimento inicial do número de infectantes, mas o número básico de reprodução R_0 não é alterado se não incluímos mortes ou nascimentos. Os detalhes matemáticos desse modelo podem ser encontrados na seção 1.7.1 da referência [10] .

Antes de encerrar essa seção, vamos discutir brevemente uma outra (dentre várias) suposição do modelo SIR que não é realista. O modelo supõe que a população é distribuída de forma homogênea, o que implica que cada indivíduo tem a mesma probabilidade de encontrar qualquer outro indivíduo da população. Essa simplificação ignora o fato de que a maioria das pessoas mantém contato, na maior parte do tempo, com um pequeno grupo de indivíduos do seu círculo social e familiar. Uma possível abordagem para esse problema é considerar várias classes diferentes de suscetíveis e infectados, e supor mistura homogênea entre as subclasses [13]. A análise de modelos desse tipo é bem mais complexa, exigindo o uso de equações diferenciais parciais. O modelo SIR tem a vantagem de ser um modelo simplificado que demostra os aspectos gerais da evolução de uma epidemia, e por isso adequado para uma introdução pedagógica ao assunto.

5 Animação

A partir modelo SIR, uma animação interativa foi desenvolvida com o auxílio da biblioteca *Bokeh* [2]. Na animação, é possível variar tando o valor do número básico de reprodução (R_0) quanto a porcentagem de vacinados (ρ), e ver-

[2] https://bokeh.org/

ificar como evoluem o número de indivíduos suscetíveis, infectados e removidos. A Figura 6 mostra uma imagem do simulador.

Essa ferramenta ajuda no entendimento qualitativo da evolução de uma epidemia, ilustra a importância das vacinas, e pode ser usado em aulas do ensino fundamental e médio, ou em atividades de divulgação científica. O aplicativo está disponível online no link SirModel.

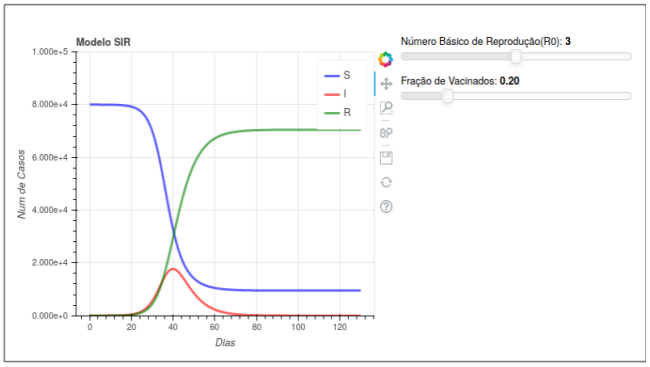

Figure 6: Imagem da ferramenta de animação.

6 Considerações finais

Neste trabalho, um modelo matemático de propagação de uma doença infecciosa foi investigado, e o método numérico de Runge-Kutta foi aplicado para resolver o conjunto de equações diferenciais do modelo. O modelo considerado contém várias simplificações, como a suposição de que o número de indivíduos da população se mantém constante, o período da epidemia é curto o suficiente para que possamos desconsiderar nascimentos ou mortes, a imunização é permanente, e suscetíveis e infectados estão distribuídos uniformemente. Modelos mais complexos lidam com cada uma dessas questões. Ainda assim, o modelo SIR faz previsões interessantes, como a necessidade de um número mínimo de suscetíveis para que a epidemia se desenvolva, e explica porque uma epidemia termina antes que toda a população contraia a doença. Esperamos que a análise aqui apresentada sirva de base para a compreensão de modelos epidemiológicos mais complexos.

Apêndice

Para aplicar o método RK4 a solução das EDOs do modelo SIR, vamos utilizar a seguinte notação: $\vec{k_j} = (s_j, i_j, r_j)$, com $j = 1, 2, 3, 4$ e $\vec{f} = (f_S, f_I, f_R)$, com f_S, f_I e f_R dados por

$$f_S(S, I, t) = -\beta SI, \tag{18}$$
$$f_I(S, I, t) = \beta SI - \gamma I, \tag{19}$$
$$f_R(I, t) = \gamma I. \tag{20}$$

Especificando o valor do passo h, os valores de β e γ, e as condições iniciais $S(0)$, $I(0)$ e $R(0)$, as soluções $S(t)$, $I(t)$ e $R(t)$ são dadas por:

$$s_1 = hf_S(S_n, I_n, t_n), \tag{21}$$
$$i_1 = hf_I(S_n, I_n, t_n), \tag{22}$$
$$r_1 = hf_R(I_n, t_n), \tag{23}$$
$$s_2 = hf_S\left(S_n + \frac{s_1}{2}, I_n + \frac{i_1}{2}, t_n + \frac{h}{2}\right), \tag{24}$$
$$i_2 = hf_I\left(S_n + \frac{s_1}{2}, I_n + \frac{i_1}{2}, t_n + \frac{h}{2}\right), \tag{25}$$
$$r_2 = hf_R\left(I_n + \frac{i_1}{2}, t_n + \frac{h}{2}\right), \tag{26}$$

$$s_3 = hf_S\left(S_n + \frac{s_2}{2}, I_n + \frac{i_2}{2}, t_n + \frac{h}{2}\right), \quad (27)$$

$$i_3 = hf_I\left(S_n + \frac{s_2}{2}, I_n + \frac{i_2}{2}, t_n + \frac{h}{2}\right), \quad (28)$$

$$r_3 = hf_R\left(I_n + \frac{i_2}{2}, t_n + \frac{h}{2}\right), \quad (29)$$

$$s_4 = hf_S(S_n + s_3, I_n + i_3, t_n + h), \quad (30)$$

$$i_4 = hf_I(S_n + s_3, I_n + i_3, t_n + h), \quad (31)$$

$$r_4 = hf_R(I_n + i_3, t_n + h), \quad (32)$$

$$S_{n+1} = S_n + \frac{1}{6}(s_1 + 2s_2 + 2s_3 + s_4), \quad (33)$$

$$I_{n+1} = I_n + \frac{1}{6}(i_1 + 2i_2 + 2i_3 + i_4), \quad (34)$$

$$R_{n+1} = R_n + \frac{1}{6}(r_1 + 2r_2 + 2r_3 + r_4). \quad (35)$$

Como o tamanho da população N é constante, a Eq. (35) é dispensável e $R(t)$ pode ser calculado pela diferença $R(t) = N - S(t) - I(t)$. Para o cenário com vacina, a única alteração ocorrerá na função f_S, que passa a ser $f_S = -\beta SI - \nu S$.

O algoritmo acima foi implementado em Python e está disponível em:
`https://github.com/aanepomuceno/sir_rk4`.

References

[1] N.G Davies, P. Klepac, Y. Liu, et al, Nature Medicine (2020).

[2] G. Giordano, F, Blanchini, R. Bruno, P. Colaneri, A. Di Fillipo, A. Di Matteo, M.Colaneri, Nature Medicine **26**, 855–860 (2020).

[3] T. Britton, F. Ball, P. Trapman1, Science (2020)

[4] N. G.Bizet, D. K. Pena, arXiv:2006.12479 (2020)

[5] N. Crokidakis, arXiv:2003.12150 (2020)

[6] S. B. Bastos, D. O. Cajueiro, arXiv:2003.14288 (2020)

[7] C. Colombo, M. Diamanti, Lettera Matematica **2**, 185–192 (2015)

[8] I. M. Foppa, *A Historical Introduction to Mathematical Modeling of Infectious Diseases* (Elsevier, Londres, 2017).

[9] A. Huppert, G. Katriel, Clinical Microbiology and Infection **19**, 999–1005 (2013).

[10] H. W. Hethcote, SIAM Review **42**(4), 599–653 (2000).

[11] C. Scherer, *Métodos Computacionais da Física* (Editora Livraria da Física, São Paulo, 2010), 2ed., p. 75.

[12] T.Harko, F.N.S. Lobo, M. K. Mak, Applied Mathematics and Computation **236**, 184-194 (2014)

[13] J. Reyes-Sylveira,A. R. Mikler, J. Zhao, A. B. Salgado, Journal of Biological Systems **19**, No. 4, 591-606 (2011)