

Modelando a epidemia de Covid-19

Paulo J. S. Silva

9 de novembro de 2021

Data de entrega: veja o Moodle

1 Introdução

A Covid-19 veio para mudar a vida de todos, isso porque exigiu uma mudança brutal no nosso comportamento com as exigências de isolamento social que começaram em meados de março de 2020.

O objetivo desse exercício é entender um pouco melhor os motivos que levaram os governos a decretar as políticas de isolamento. Para isso vocês terão que fazer simulações baseadas em modelos clássicos da epidemiologia matemática buscando entender qual seria a trajetória natural da pandemia se não houvesse intervenções.

2 Modelos compartimentais

a Leitura dessa seção é opcional, o que realmente importa para fazer o trabalho é a equação SEIR no final. Coloque o texto para contextualizar o trabalho.

Para simular a propagação de surtos epidêmicos em uma população usamos modelos matemáticos baseados em equações diferenciais. Vamos começar explicando um pouco como uma doença com transmissão por contato social se comporta e entender, mesmo que superficialmente o que dizem esses modelos.

O ponto de partida é pensar na população composta por pessoas que se movimentam e interagem, seja no trabalho, na escola, no transporte público ou na praça. Doenças de contato social, como a Covid-19, se aproveitam disso. Ao interagir com um indivíduo *infeccioso*, uma pessoa doente que está no período de transmissão da doença, um indivíduo *suscetível*, que não possui defesas imunológicas capazes de impedir uma infecção, tem uma *chance* de pegar a doença.

Mas como essa chance de contágio se comporta: isso depende de vários fatores. Depende do patógeno, alguns microrganismos têm mais facilidade de se propagar que outros. Depende de como é a interação. Se for curta a chance de transmissão é menor. Se for longa e próxima maior. Se for em ambiente fechado, no caso da Covid-19 a possibilidade de transmissão aumenta muito. Ou seja, é um fenômeno muito complexo.

Podemos simular tudo isso no computador. Sim! De fato é possível fazer simulações muito complexas usando supercomputadores, em que as pessoas são modeladas individualmente cada uma com um comportamento descrevendo onde mora, como se locomove, se trabalha, se vai à escola, etc. O sistema então iria então simular a vida de toda uma simulação e, a cada encontro, entre dois indivíduos com um infeccioso tentar simular várias nuances do processo de infecção.

Mas vocês podem imaginar que um sistema tem várias dificuldades práticas para ser implementado. Como descrever os vários indivíduos? Quais são os vários parâmetros que descrevem as formas de interação naquela sociedade? Qual a capacidade do vírus de propagar? Entre vários outros.

Mas há outro caminho possível: ao invés de se tentar perseguir o estado de cada indivíduo podemos usar equações de descrevem uma espécie de comportamento médio de toda a população. Como a população é composta de uma enorme quantidade de indivíduos vemos que essas simulações muitas vezes complexas possuem comportamento médio bastante estável, pois as diferenças se combinam de variadas maneiras. Em várias situações esses modelos médios que representam a combinação do comportamento uma infinidade de agentes / partículas podem ser capturados em equações diferenciais. Essa é a origem dos modelos compartimentais de epidemiologia matemática que vamos implementar.

2.1 Modelo SIR

Desses modelos o mais simples é o modelo SIR. Nele a população é dividida em um grupo de pessoas *Suscetíveis* (S), *Infecciosas* (I) e *Recuperadas* (R). A ideia é que essas categorias descrevem os estágios da doença. Uma pessoa começa Suscetível, eventualmente fica doente e passa a Infeccioso e por fim se Recupera (ou falece). Vamos usar a hipótese que ao ficar doente e se recuperar a pessoa adquire imunidade contra a doença e não volta se infectar.

Vamos então modelar primeiro o comportamento médio do processo infecção. Como vimos acima isso envolve apenas pessoas dos grupos S e I . No nosso modelo, vamos considerar que essas variáveis guardam o número de pessoas em cada estado em uma população com N indivíduos no total ($S + I + R = N$). Então se há I pessoas transmitindo a doença na população, cada encontro de uma pessoa de S com um indivíduo ao acaso tem probabilidade de

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

de ocorrer com um indivíduo infeccioso. Note que a quantidade desses encontros aumenta com a quantidade de pessoas suscetíveis e que nem todo encontro gera uma transmissão, então é preciso levar em consideração a probabilidade de uma interação levar a uma infecção. No final chegamos a um modelo simplificado na forma

$$\frac{dS}{dt} = -\beta S \frac{I}{N}.$$

Vamos entender cada um dos termos. O sinal é negativo porque o número de indivíduos em S só pode diminuir. Então sua derivada ser negativa. O parâmetro β tenta capturar vários fatores sendo maior se houve maior chance de contágio a cada interação e menor se a probabilidade for menor. Ele captura, por exemplo, o comportamento social típico (as pessoas têm muitos contatos durante o dia? esses contatos são mais seguros ou arriscados?) e características biológicas da doença (duração da janela de infecção, transmissibilidade do patógeno, etc.). Já o S parece porque os contatos típicos de cada pessoa em S colabora para infectar mais pessoas. Em outras palavras, quantas mais pessoas suscetíveis na população, maior a transmissão possível. Por fim, já vimos que o I/N é a probabilidade de um contato ser com uma pessoa infectada (e quanto maior a proporção de infectados mais chances de infecções ocorrerem).

Agora qual é a derivada da variável I ? Entram em I as pessoas que saem de S ou seja

$$\frac{dI}{dt} = \beta S \frac{I}{N} - \text{pessoas que saem de } I.$$

Para definir a segunda parcela, o número de pessoas que saem de I , precisamos modelar o processo das pessoas ficarem boas. Um modelo muito usado nesse caso é que a cada unidade de tempo a chance de alguém que está infectado ficar recuperado é constante. Esse tipo de modelo considera então que a taxa de saída é proporcional a uma contate, que chamaremos de γ . A ideia é que a cada dia uma proporção γ dos infectados transitem para o estado R . Em suma temos

$$\frac{dI}{dt} = \beta S \frac{I}{N} - \gamma I.$$

Por fim, temos a derivada de S . Passam a ser recuperados os indivíduos que saem de I , ou seja

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

Esse processo é representado graficamente na Figura 1 em que cada caixa representa um dos estados.

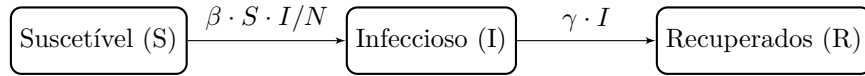


Figura 1: Modelo SIR

Já o sistema completo de equações diferenciais fica

$$\begin{aligned} \frac{dS}{dt} &= -\beta S \frac{I}{N}, \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I. \end{aligned}$$

Antes de terminar essa seção vamos fazer uma pequena mudança de variáveis que é interessante na hora de resolver o problema numericamente. Se multiplicarmos todas as equações por $1/N$ temos

$$\begin{aligned}\frac{1}{N} \frac{dS}{dt} &= \frac{d(S/N)}{dt} = -\beta \frac{S}{N} \frac{I}{N}, \\ \frac{1}{N} \frac{dI}{dt} &= \frac{d(I/N)}{dt} = \beta \frac{S}{N} \frac{I}{N} - \gamma \frac{I}{N}, \\ \frac{1}{N} \frac{dR}{dt} &= \frac{d(R/N)}{dt} = \gamma \frac{I}{N}.\end{aligned}$$

Dando nomes para os valores $\mathcal{S} = S/N$, $\mathcal{I} = I/N$ e $\mathcal{R} = R/N$, que representam proporções da população que estão em cada estado, variando entre 0 e 1, o sistema fica um pouco mais simples

$$\begin{aligned}\frac{d\mathcal{S}}{dt} &= -\beta \mathcal{S} \mathcal{I}, \\ \frac{d\mathcal{I}}{dt} &= \beta \mathcal{S} \mathcal{I} - \gamma \mathcal{I}, \\ \frac{d\mathcal{R}}{dt} &= \gamma \mathcal{I}.\end{aligned}$$

Além disso vamos definir novas parâmetros a partir dos parâmetros β e γ originais $r_0 = \frac{\beta}{\gamma}$ e $T_{\text{inf}} = \frac{1}{\gamma}$. Esses novos parâmetros tem uma interpretação mais natural. O r_0 é conhecido como número de reprodução basal e representa o número de pessoas que um doente iria infectar em média durante o curso de sua doença. Esse é um número que ouvimos muito falar nos jornais. Se for menor que 1 quer dizer que a doença “diminuindo”, já se for maior que 1 a epidemia está se alastrando. Já o T_{inf} representa o tempo médio que a pessoa fica infecciosa, ou seja a duração média que um doente fica ativamente transmitindo a doença. Em função desses parâmetros o modelo final fica:

$$\begin{aligned}\frac{d\mathcal{S}}{dt} &= -\frac{r_0}{T_{\text{inf}}} \mathcal{S} \mathcal{I}, \\ \frac{d\mathcal{I}}{dt} &= \frac{r_0}{T_{\text{inf}}} \mathcal{S} \mathcal{I} - \frac{1}{T_{\text{inf}}} \mathcal{I}, \\ \frac{d\mathcal{R}}{dt} &= \frac{1}{T_{\text{inf}}} \mathcal{I}.\end{aligned}$$

2.2 O modelo SEIR

Para modelar a Covid-19 o ideal é acrescentarmos mais um estado ao modelo, que chamaremos de Exposto. Uma pessoa exposta é alguém que já pegou a doença mas ainda não a está transmitindo ativamente (ou seja ainda não é Infecciosa). É como se a pessoa tivesse incubando a doença. Para isso precisamos também introduzir um parâmetro que é o tempo médio que a pessoa fica nesse estado. Vamos usar T_{inc} para isso. Chegamos ao esquema:

O respectivo sistema de equações diferenciais é

$$\begin{aligned}\frac{dS}{dt} &= -\frac{r_0}{T_{\text{inf}}}S\mathcal{I}, \\ \frac{d\mathcal{E}}{dt} &= \frac{r_0}{T_{\text{inf}}}S\mathcal{I} - \frac{1}{T_{\text{inc}}}\mathcal{E}, \\ \frac{d\mathcal{I}}{dt} &= \frac{1}{T_{\text{inc}}}\mathcal{E} - \frac{1}{T_{\text{inf}}}\mathcal{I}, \\ \frac{d\mathcal{R}}{dt} &= \frac{1}{T_{\text{inf}}}\mathcal{I}.\end{aligned}\tag{SEIR}$$

Essa é a EDO que você terá que resolver nesse projeto.

Se quiser mais informações sobre modelos compartimentais, leia [1].

3 O que você deve fazer

O nosso objetivo nesse projeto computacional é resolver um PVI vetorial descrito por (SEIR) com parâmetros típicos da Covid-19 no início da pandemia. O objetivo será acompanhar a curva de doentes e estimar a demanda de leitos de UTI, comparando com a disponibilidade para ver se uma ação para bloquear a transmissão era ou não necessária para evitar o colapso do sistema de saúde.

Para isso, você deve resolver o sistema de equações diferenciais ordinárias SEIR considerando os seguintes parâmetros e condições iniciais:

1. Número de reprodução basal (r_0): 2.5, sendo otimista (adimensional) [2].
2. Tempos médio de incubação e infecção: $T_{\text{inc}} = 5.2$ dias, $T_{\text{inf}} = 2.9$ dias [6].
3. Demanda de leitos de UTI: estimava-se que 3% dos doentes precisariam de leito de UTI [3, 5].
4. Mediana dos tempos de UTI: 7 dias [3, 5].
5. Disponibilidade de leitos de UTI na região Sudeste em 2019 (antes da pandemia): 27 por 100 mil habitantes [4].
6. Condição inicial: $S_0 = 0.999999$, $E_0 = 10^{-6}$, $I_0 = 0$, $R_0 = 0$. Ou seja, há uma pessoa por milhão na população que está exposta e ainda não começou a transmitir a doença. Todo o restante da população é suscetível.

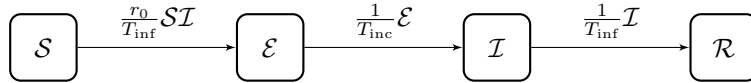


Figura 2: Modelo SEIR

O algoritmo que você deve usar é o método do ponto médio para EDO vetoriais e o horizonte de simulação deve ser 498 dias (um ano e meio). O passo no tempo para ser usado é de um dia.

Ao final você terá as trajetórias das proporções projetadas de pessoas em cada compartimento para cada dia a partir do início da simulação. De posse disso você deve escrever um relatório descrevendo o que foi feito e respondendo as seguintes perguntas:

1. Qual a maior proporção de pessoas doentes observadas em um único dia? Considerando que o estado de São Paulo possui 44,04 milhões de habitantes, essa proporção representa quantas pessoas doentes ao mesmo em um estado como o nosso?
2. Quantos dias são necessários para chegar no pico de doentes?
3. Estime o número de leitos de UTI para lidar com o pico. Note que isso precisa ser feito com cuidado. Não basta multiplicar o pico de \mathcal{I} pela necessidade de leitos. Você tem que levar em consideração a dinâmica e o tempo em UTI. O que você deve:
 - (a) Os compartimento \mathcal{I} não diz quantos novos doentes há por dia, ele diz a proporção de pessoas doentes, mas eles podem estar doentes há vários dias (o tempo médio em \mathcal{I} é de 2.9 dias). Para saber o fluxo de novos doentes, a forma mais eficaz é ver o fluxo de pessoas saindo de \mathcal{S} , que se tornaram doentes depois de T_{inc} dias de incubação, ou o número de pessoas entrando diariamente em \mathcal{R} que pararam de transmitir a doença naquele dia (seja porque se curaram, seja porque foram parar no hospital). Nesse sentido é mais natural considerar o fluxo diário de pessoas transitando de \mathcal{I} para \mathcal{R} como o fluxo de novos doentes que podem ir parar na UTI. Calcule e use esse número.
 - (b) Multiplicando esse fluxo de novos doentes pela a demanda de leitos de UTI por doente dada acima, você descobre quantos novos leitos são necessários a cada dia.
 - (c) Mas os leitos não são usados por um único dia, sim por 7 dias (veja acima). Então para estimar a necessidade total de leitos você deve somar essa demanda de novos leitos por 7 dias em todo o horizonte da simulação.
 - (d) O número máximo de leitos necessários é então o máximo dessa soma em janela deslizando de 7 dias.
4. O número de leitos disponíveis é suficiente para lidar com o pico de demanda? Qual o excesso de demanda máximo, se houver?
5. Você consegue estimar quantas pessoas *ao todo* iriam ficar sem acesso a UTIs? Justifique a forma que você fez essa estimativa.
6. Com base nessa simulação o que você diria ao governador do estado no início da pandemia? As restrições de circulação foram necessárias?

4 O que você deve entregar?

O trabalho deve ser feito em grupos de 2 ou 3 pessoas. O grupo deve entregar um único arquivo compactado em formato zip ou tar.gz contendo:

1. Um arquivo texto chamado grupo.txt com os nomes e RAs dos membros do grupo.
2. O código fonte usado na análise, incluindo a implementação do método do ponto médio. O seu programa pode ser escrito em C, Python, Matlab/Octave ou Julia.
3. Um arquivo explicando como compilar (se for o caso) e executar o código fonte para reproduzir os seus resultados.
4. Um relatório, em formato PDF descrevendo o que foi feito e as respostas para as perguntas da seção anterior.

Referências

- [1] Julie C. Blackwood and Lauren M. Childs. An introduction to compartmental modeling for the budding infectious disease modeler. *Letters in Biomathematics*, 5(1):195–221, December 2018.
- [2] Ying Liu, Albert A Gayle, Annelies Wilder-Smith, and Joacim Rocklöv. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2), March 2020.
- [3] Seyed M. Moghadas, Affan Shoukat, Meagan C. Fitzpatrick, Chad R. Wells, Pratha Sah, Abhishek Pandey, Jeffrey D. Sachs, Zheng Wang, Lauren A. Meyers, Burton H. Singer, and Alison P. Galvani. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proceedings of the National Academy of Sciences*, 117(16):9122–9126, April 2020.
- [4] Úrsula Neves. Covid-19: IBGE divulga distribuição de UTIs, respiradores e profissionais de saúde no Brasil. <https://pebmed.com.br/covid-19-ibge-divulga-distribuicao-de-utis-respiradores-e-profissionais-de-saude-no-brasil/>, May 2020.
- [5] Rachel Lane Socolovitch, Renata Rego Lins Fumis, Bruno Martins Tomazini, Laerte Pastore, Filomena Regina Barbosa Gomes Galas, Luciano Cesar Pontes de Azevedo, and Eduardo Leite Vieira Costa. Epidemiology, outcomes, and the use of intensive care unit resources of critically ill patients diagnosed with COVID-19 in Sao Paulo, Brazil: A cohort study. *PLOS ONE*, 15(12):e0243269, December 2020.
- [6] Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV

outbreak originating in Wuhan, China: A modelling study. *The Lancet*, 395(10225):689–697, February 2020.