

Modelagem Epidemiológica do COVID-19

Modelo SIR/SEIR global, espacial e e estratificado

Força tarefa do ITA para simulação avançada do COVID-19

Instituto Tecnológico de Aeronáutica

Departamento de Física

Prof. Marco Antonio Ridenti

Relatório para comunicação das investigações teóricas voltadas à modelagem epidemiológica do COVID-19. Versão 4.0.

23 de março de 2020

Sumário

1	Objetivo	1
2	Desenvolvimento Teórico - Modelo 0-D	1
2.1	Modelo SIR	1
2.2	Modelo SEIR	2
2.3	Modelo SEAIR	3
3	Desenvolvimento Teórico - Propagação epidêmica sobre um território	3
3.1	SIR - Espacial	5
3.2	SEIR - Espacial	5
3.3	Modelagem dos fluxos populacionais	6
3.4	Transformação para cidades/microrregiões	7
3.5	Inclusão de faixa etária	8
3.5.1	Modelo global com faixas etárias	10
3.5.2	Matriz de contato	12
3.6	Cenários	13
3.6.1	Nenhuma intervenção. Cenário 1	13
3.6.2	Fechamento de escola apenas. Cenário 2	13
3.6.3	Fechamento de escola e distanciamento social. Cenário 3	13
3.6.4	Fechamento de escola, distanciamento no trabalho e distanciamento social. Cenário 4	13
3.6.5	Isolamento dos idosos. Cenário 5.	14
3.6.6	Fechamento de escola, distanciamento no trabalho, distanciamento social e isolamento dos idosos. Cenário 6	14
3.6.7	Knock Down com isolamento de idosos. Cenário 7	14
3.6.8	Normalização	15
3.7	Cálculo de grandezas úteis: número de casualidades (C), hospitalizações (H), leitos ocupados (L) e número final total de infectados (M)	15
3.7.1	Casualidades	15

3.7.2	Hospitalizações	15
3.7.3	Leitos ocupados	16
3.7.4	Número final total de infectados M	16
3.7.5	Outras saídas importantes	16
4	Parâmetros dos modelos SIR, SEIR e SEAIR	18
4.1	Parâmetros do modelo SIR, SEIR e SEAIR	18

1 Objetivo

A presente versão deste relatório tem como objetivo a rápida comunicação da modelagem matemática para a descrição quantitativa da propagação da contaminação pelo COVID-19. Apresento dois tipos de modelo. O primeiro, mais simples, é um modelo global, pois considera a população como um todo e despreza a dinâmica local e a difusão territorial da epidemia. O segundo modelo resulta de um esforço para modelar a dinâmica espacial do vírus sobre o território de um país. A razão para considerar os dois casos está no fato de que o primeiro é de rápida implementação e permite captar as características gerais do problema. O segundo é mais realista, mas de implementação mais dispendiosa.

A abordagem “0-D” segue o padrão conhecido como SIR (susceptible-infected-recovered) e SEIR (susceptible-exposed-infected-recovered) (seção 2). Nós mostramos como essas abordagens podem ser extendidas para incluir a propagação regional das doenças infecciosas (seção 3).

2 Desenvolvimento Teórico - Modelo 0-D

2.1 Modelo SIR

No modelo SIR considera-se a evolução temporal da população susceptível (S), infectada (I) e recuperada (R). As equações diferenciais que regem a variação dessas populações é dada por

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I(t) \quad (1)$$

$$\frac{dI}{dt} = \beta \frac{S}{N}I(t) - \mu_{cov}I(t - d_C) - \mu_{eq}I - \gamma I(t - d_I) \quad (2)$$

$$\frac{dR}{dt} = \gamma I(t - d_I) - \mu_{eq}R \quad (3)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I(t - d_C) \quad (4)$$

$$N = S + I + R \quad (5)$$

onde μ_{eq} é a taxa de mortalidade em condições normais e μ_{cov} é a taxa de mortalidade do COVID-19. Os outros parâmetros são Λ , a taxa natalidade típica da população,

β , coeficiente de taxa de contaminação, γ , a taxa típica de recuperação do COVID-19 e d_I e d_C são, respectivamente, os tempos típicos de recuperação e óbito. Note que a equação (4) resulta da soma das equações (1) a (3), não sendo portanto linearmente independente delas. Trata-se da equação para a população total em sua forma mais geral, admitindo aumento da população por conta da taxa de natalidade e diminuição por conta da taxa de mortalidade. As simplificações mais usuais habitualmente consideram uma situação de equilíbrio; no entanto, pelo conhecimento prévio sobre a doença, acredito ser necessário avaliar o impacto sobre a taxa de mortalidade.

2.2 Modelo SEIR

No modelo SEIR considera-se a evolução temporal da população susceptível (S), exposta(E), infectada (I) e recuperada (R). As equações diferenciais que regem a variação dessas populações é dada por

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I(t) \quad (6)$$

$$\frac{dE}{dt} = \beta \frac{S}{N}I(t) - \mu_{eq}E - E(t - d_L) \quad (7)$$

$$\frac{dI}{dt} = E(t - d_L) - \mu_{cov}I(t - d_C) - \mu_{eq}I(t) + \gamma I(t - d_I) \quad (8)$$

$$\frac{dR}{dt} = \gamma I(t - d_I) - \mu_{eq}R \quad (9)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I(t - d_C) \quad (10)$$

$$N = S + E + I + R \quad (11)$$

onde μ_{eq} é a taxa de mortalidade em condições normais e μ_{cov} é a taxa de mortalidade do COVID-19. Os outros parâmetros são Λ , a taxa natalidade típica da população, β , coeficiente de taxa de contaminação, γ , a taxa típica de recuperação do COVID-19, e d_L , d_I e d_C são, respectivamente, os tempos típicos de incubação, recuperação e óbito. Note que a equação (11) resulta da soma das equações (12) a (16), não sendo portanto linearmente independente delas. As mesmas considerações feitas anteriormente se aplicam a este caso.

2.3 Modelo SEAIR

No modelo SEAIR tem sido utilizado frequentemente na literatura recente sobre o COVID-19 [1], e consiste basicamente na introdução de uma nova variável dinâmica, a dos assintomáticos ou sub-clínicos. Os assintomáticos também podem infectar e também provêm dos expostos, mas tem uma taxa de infecção menor, nenhuma mortalidade e em tese podem ter uma taxa de recuperação menor. Nesse caso, as equações dinâmicas podem ser escritas como:

$$\frac{dS}{dt} = \Lambda N - \mu_{eq}S - \beta \frac{S}{N}I(t) - \alpha\beta \frac{S}{N}A(t) \quad (12)$$

$$\frac{dE}{dt} = \beta \frac{S}{N}I(t) + \alpha\beta \frac{S}{N}A(t) - \mu_{eq}E(t) - E(t - d_L) \quad (13)$$

$$\frac{dI}{dt} = \rho E(t - d_L) - \mu_{eq}(t)I(t) - \mu_{cov}I(t - d_C) - \gamma I(t - d_I) \quad (14)$$

$$\frac{dA}{dt} = (1 - \rho)E(t - d_L) - \mu_{eq}A(t) - A(t - d_I) \quad (15)$$

$$\frac{dR}{dt} = \gamma I(t - d_I) + A(t - d_I) - \mu_{eq}R \quad (16)$$

$$\frac{dN}{dt} = \Lambda N - \mu_{eq}N - \mu_{cov}I(t - d_C) \quad (17)$$

$$N = S + E + A + I + R \quad (18)$$

nesse modelo, é introduzido o fator α , que representa o quanto a infectividade do assintomático é menor do que a do simptomático (α deve ser um valor positivo menor ou igual a 1), e a probabilidade ρ , que é a probabilidade de um exposto se tornar sintomático.

3 Desenvolvimento Teórico - Propagação epidêmica sobre um território

A generalização para considerar a propagação espacial (ou territorial) se baseia na equação da continuidade. Nesse caso, vamos trabalhar com a área total A_T de um território (digamos, o território brasileiro), e discretizá-lo em partes de igual área, A_i , que chamaremos de elemento de área. A forma das áreas é arbitrária, contanto que tenham sempre a mesma área, mas por simplificação podemos imaginá-la como

sendo um quadrado. Cada elemento área A_i pode ser caracterizada pela coordenada geográfica central (latitude e longitude).

Em poucas palavras, o problema se resumirá a calcular a evolução das populações dentro dos elementos de área, levando-se em conta o fluxo da população para dentro e para fora dela. Antes disso, vamos introduzir alguma notação e definições importantes:

$A_T :=$ ‘Área total do território.’

$A_i :=$ ‘ i -ésimo elemento de área. Seu valor é simplesmente A .’

$$A_T = \sum_i A_i$$

$s_i = \frac{S}{A_i} :=$ ‘densidade de pessoas susceptíveis no i -ésimo elemento de área.’

$e_i = \frac{E}{A_i} :=$ ‘densidade de pessoas expostas no i -ésimo elemento de área.’

$y_i = \frac{I}{A_i} :=$ ‘densidade de pessoas infectadas no i -ésimo elemento de área.’

$r_i = \frac{r}{A_i} :=$ ‘densidade de pessoas recuperadas no i -ésimo elemento de área.’

$n_i = \frac{N}{A_i} :=$ ‘densidade total de pessoas no i -ésimo elemento de área.’

$\Gamma_n(A_i) :=$ ‘fluxo total de indivíduos através do i -ésimo elemento de área.’

$\Gamma_s(A_i) :=$ ‘fluxo de indivíduos susceptíveis através do i -ésimo elemento de área.’

$\Gamma_e(A_i) :=$ ‘fluxo de indivíduos expostos através do i -ésimo elemento de área.’

$\Gamma_y(A_i) :=$ ‘fluxo de indivíduos infectados através do i -ésimo elemento de área.’

$\Gamma_r(A_i) :=$ ‘fluxo de indivíduos recuperados através do i -ésimo elemento de área.’

Os fluxos tem tanto um contribuição positiva, isto é, de indivíduos que se deslocam de fora para dentro, $\Gamma^{\text{in}}(A_i)$, quanto uma contribuição negativa, isto é, de indivíduos que se deslocam de dentro para fora, $-\Gamma^{\text{out}}(A_i)$. Qualquer um dos fluxos acima definidos podem ser escritos como:

$$\Gamma_k(A_i) = \Gamma_k^{\text{in}}(A_i) - \Gamma_k^{\text{out}}(A_i)$$

onde k é índice que simplesmente representa uma das variáveis populacionais, isto é, $k = n, s, e, y, r$. Chamamos a atenção também para o fato de que o fluxo da população total é igual a soma dos fluxos de suas parcelas, isto é,

$$\Gamma_n(A_i) = \Gamma_s(A_i) + \Gamma_e(A_i) + \Gamma_y(A_i) + \Gamma_r(A_i) . \quad (19)$$

No modelo SIR basta omitir o segundo termo da soma. Vamos mostrar como esses fluxos podem ser modelados nas seções seguintes.

3.1 SIR - Espacial

As equações de taxa agora devem ser resolvidas para cada elemento de área, considerando-se os fluxos de saída e entrada.

$$\frac{ds_i}{dt} = \frac{\Gamma_s(A_i)}{A} + \Lambda n_i - \mu_{eq} s_i(t) - \beta \frac{s_i(t)}{n_i} y_i \quad (20)$$

$$\frac{dy_i}{dt} = \frac{\Gamma_y(A_i)}{A} + \beta \frac{s_i(t)}{n_i} y_i - \mu_{eq} y_i(t) - \mu_{cov} y_i(t - d_C) - \gamma y_i(t - d_I) \quad (21)$$

$$\frac{dr_i}{dt} = \frac{\Gamma_r(A_i)}{A} + \gamma y_i(t - d_I) - \mu_{eq} r_i \quad (22)$$

$$\frac{dn_i}{dt} = \frac{\Gamma_n(A_i)}{A} + \Lambda n_i - \mu_{eq} n_i - \mu_{cov} y_i(t - d_C) \quad (23)$$

$$n_i = s_i + y_i + r_i \quad (24)$$

Para obter a população de uma determinada parcela em todo território brasileiro, digamos, do número total de infectados, basta somar sobre todos os elementos de área e multiplicar o resultado pelo valor do elemento de área, *i.e.*

$$s = A \sum_i s_i \quad (25)$$

O acoplamento entre os diversas elementos de área é regido pelos fluxos de entrada, que como veremos mais a frente, é função dos fluxos de saída dos outros elementos de área.

3.2 SEIR - Espacial

As equações para o modelo SEIR podem ser escritas como

$$\frac{ds_i}{dt} = \frac{\Gamma_s(A_i)}{A} + \Lambda n_i - \mu_{eq} s_i - \beta \frac{s_i}{n_i} y_i \quad (26)$$

$$\frac{de_i}{dt} = \frac{\Gamma_e(A_i)}{A} + \beta \frac{s_i}{n_i} y_i - \mu_{eq} e_i(t) - e_i(t - d_L) \quad (27)$$

$$\frac{dy_i}{dt} = \frac{\Gamma_y(A_i)}{A} + e_i(t - d_L) - \mu_{eq} y_i - \mu_{cov} y_i(t - d_C) - \gamma y_i(t - d_I) \quad (28)$$

$$\frac{dr_i}{dt} = \frac{\Gamma_r(A_i)}{A} + \gamma y_i(t - d_I) - \mu_{eq} r_i \quad (29)$$

$$\frac{dn_i}{dt} = \frac{\Gamma_n(A_i)}{A} + \Lambda n_i - \mu_{eq} n_i - \mu_{cov} y_i(t - d_C) \quad (30)$$

$$n_i = s_i + e_i + y_i + r_i \quad (31)$$

É importante considerar que os parâmetros do modelo podem agora ser considerados localmente, se necessário, levando-se em conta as peculiaridades de cada região do território nacional.

3.3 Modelagem dos fluxos populacionais

O fluxo total de saída em um elemento de área i pode ser prescrito, com base em conhecimento prévio sobre a região, ou por meio de algum tipo de modelo. É razoável considerar, por exemplo, que o fluxo de saída é proporcional à densidade populacional:

$$\Gamma_n^{\text{out}}(A_i) = \xi n_i . \quad (32)$$

Sabendo o fluxo total, podemos considerar que os fluxos das parcelas é proporcional às parcelas, isto é,

$$\begin{aligned} \Gamma_s^{\text{out}}(A_i) &= \frac{s_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_e^{\text{out}}(A_i) &= \frac{e_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_y^{\text{out}}(A_i) &= \frac{y_i}{n_i} \Gamma_n^{\text{out}}(A_i) \\ \Gamma_r^{\text{out}}(A_i) &= \frac{r_i}{n_i} \Gamma_n^{\text{out}}(A_i) \end{aligned}$$

Para modelar os fluxos de entrada em um determinado elemento de área A_i , nós temos que considerar que todas os outros elementos de área contribuem para o fluxo.

Podemos escrever, de modo geral, o fluxo de entrada total em um elemento de área A_i como

$$\Gamma_n^{\text{in}}(A_i) = \sum_{j \neq i} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \quad (33)$$

onde a soma é feita sobre todo os elementos de área, menos o elemento i , d_{ij} é a distância entre o elemento de área j e i e $f(d_{ij})$ é uma função decrescente de d_{ij} . Essa função deve ter a seguinte propriedade

$$\sum_{i \neq j} f(d_{ij}) = 1 \quad (34)$$

A função $f(d_{ij})$ deve ser interpretada como a fração do fluxo de população em A_j que se desloca para A_i . Os fluxos parciais também podem ser considerados proporcionais às parcelas populacionais, ou seja,

$$\begin{aligned} \Gamma_s^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{s_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_e^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{e_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_y^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{y_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \\ \Gamma_r^{\text{in}}(A_i) &= \sum_{j \neq i} \frac{r_j}{n_j} f(d_{ij}) \Gamma_n^{\text{out}}(A_j) \end{aligned}$$

3.4 Transformação para cidades/microrregiões

No modelo apresentado anteriormente, definimos um elemento de área arbitrário do território nacional. Note que as equações diferenciais desse modelo podem ser escritas independentemente do valor da área, bastando multiplicá-las membro a membro pela densidade A . Em geral, podemos trabalhar com o número total de municípios ou microrregiões em vez de trabalhar com elementos de área arbitrários. Tal abordagem facilita a implementação, a integração das informações do IBGE e a interpretação dos resultados. As equações, nessa abordagem, tomam a seguinte forma¹:

¹A partir desse ponto, vamos escrever as equações sempre no formalismo SEIR.

$$\frac{dS_i}{dt} = \Gamma_i^s + \Lambda N_i - \mu_{eq} S_i - \beta \frac{S_i}{N_i} I_i \quad (35)$$

$$\frac{dE_i}{dt} = \Gamma_i^e + \beta \frac{S_i}{N_i} I_i - \mu_{eq} E_i(t) - E_i(t - d_L) \quad (36)$$

$$\frac{dI_i}{dt} = \Gamma_i^y + E_i(t - d_L) - \mu_{eq} I_i - \mu_{cov} I_i(t - d_C) - \gamma I_i(t - d_I) \quad (37)$$

$$\frac{dR_i}{dt} = \Gamma_i^r + \gamma I_i(t - d_I) - \mu_{eq} R_i \quad (38)$$

$$\frac{dN_i}{dt} = \Gamma_i^n + \Lambda N_i - \mu_{eq} N_i - \mu_{cov} I_i(t - d_C) \quad (39)$$

$$N_i = S_i + E_i + I_i + R_i \quad (40)$$

Agora cada população ou parcela populacional (S_i, E_i, I_i, R_i) é indexada pela i -ésima microrregião ou i -ésimo município. A interpretação para os fluxos e para a função $f(d_{i,j})$ é análoga ao caso anterior. Note também que mudamos a notação do fluxo; o sub-índice i indica a microrregião/município e o índice superescrito indica a parcela populacional.

3.5 Inclusão de faixa etária

A inclusão da faixa etária pode ser realizada a partir da quebra da população em estratos de faixa etária. Dessa forma, uma dada variável de população, digamos, a população total $N_{i,k}$, é agora indexada pela sua i -ésima microrregião e sua k -ésima faixa etária. No Brasil, por exemplo, temos $M = 558$ microrregiões e podemos, por exemplo, trabalhar com faixas etárias espaçadas em cinco anos até 80 ou mais, totalizando $Q = 17$ faixas etárias. Considerando que temos quatro parcelas populacionais (S, E, I, R) , teremos no total $4 \times 558 \times 17 = 37944$ variáveis dependentes do tempo e $558 \times 17 = 9486$ sistemas de equações diferenciais ordinárias a resolver. Apesar de parecer muito, é um cálculo factível para os padrões computacionais atuais. Antes de escrever as equações, é necessário introduzir uma nova notação.

$S_{ik} :=$ ‘número de indivíduos susceptíveis na i -ésima região e k -ésima faixa etária.’
 $E_{ik} :=$ ‘número de indivíduos expostos na i -ésima região e k -ésima faixa etária.’
 $I_{ik} :=$ ‘número de indivíduos infectados na i -ésima região e k -ésima faixa etária.’
 $R_{ik} :=$ ‘número de indivíduos recuperados na i -ésima região e k -ésima faixa etária.’
 $N_{ik} :=$ ‘número total de indivíduos na i -ésima região e k -ésima faixa etária.’

$\Gamma_{ik}^n :=$ ‘fluxo total de indivíduos através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^s :=$ ‘fluxo total de susceptíveis através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^e :=$ ‘fluxo total de expostos através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^y :=$ ‘fluxo total de infectados através da i -ésima região, na k -ésima faixa etária.’
 $\Gamma_{ik}^r :=$ ‘fluxo total de recuperados através da i -ésima região, na k -ésima faixa etária.’

$$\begin{aligned}
 S_i &= \sum_k S_{ik} & ; & & \Gamma_i^s &= \sum_k \Gamma_{ik}^s \\
 E_i &= \sum_k E_{ik} & ; & & \Gamma_i^e &= \sum_k \Gamma_{ik}^e \\
 I_i &= \sum_k I_{ik} & ; & & \Gamma_i^y &= \sum_k \Gamma_{ik}^y \\
 R_i &= \sum_k R_{ik} & ; & & \Gamma_i^r &= \sum_k \Gamma_{ik}^r \\
 N_i &= \sum_k N_{ik} & ; & & \Gamma_i^n &= \sum_k \Gamma_{ik}^n
 \end{aligned}$$

As equações podem então ser escritas como:

$$\frac{d}{dt}S_{ik} = \Gamma_{ik}^s + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}S_{ik} - \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij}(t) \quad (41)$$

$$\frac{d}{dt}E_{ik} = \Gamma_{ik}^e + \sum_j \beta_{i,kj} \frac{S_{ik}}{N_i} I_{ij}(t) - \mu_{eq,ik}E_{ik} + E_{ik}(t - d_L) \quad (42)$$

$$\begin{aligned} \frac{d}{dt}I_{ik} = & \Gamma_{ik}^y + E_{ik}(t - d_L) - \mu_{ik,eq}I_{ik}(t) - \mu_{ik,cov}I_{ik}(t - d_C) \\ & - \gamma_{ik}I_{ik}(t - d_I) \end{aligned} \quad (43)$$

$$\frac{d}{dt}R_{ik} = \Gamma_{ik}^r + \gamma_{ik}I_{ik}(t - d_I) - \mu_{eq,ik}R_{ik} \quad (44)$$

$$\frac{d}{dt}N_{ik} = \Gamma_{ik}^n + \delta_{1k}\Lambda_{ik}N_{ik} - \mu_{eq,ik}N_{ik} - \mu_{ik,cov}I_{ik}(t - d_C) \quad (45)$$

$$N_{ik} = S_{ik} + E_{ik} + I_{ik} + R_{ik} \quad (46)$$

onde δ_{1k} é a função delta de Kronecker (1 se $k = j$ e 0 noutro caso) e $\beta_{i,kj}$ é o coeficiente de taxa de exposição que um indivíduo susceptível da faixa etária k tem a um indivíduo infectado da faixa etária j , na região i . Note na notação aqui adotada a dependência dos parâmetros com a região e faixa etária estão explicitadas por meio dos índices subscritos, por exemplo, $\mu_{eq,ik}$ é o coeficiente de taxa de mortalidade típica na faixa etária k na região i .

Em uma escala temporal maior do que a escala típica das epidemias, o envelhecimento da população produz um fluxo entre as faixas etárias, apenas positivo. Nas equação acima, não levamos em consideração esse fluxo contínuo, desprezível contanto que as faixas etárias não sejam mais estreitas do que um ano.

3.5.1 Modelo global com faixas etárias

É possível escrever as equações SEIR considerando as faixas etárias, desconsiderando no entanto sua distribuição espacial. Isso pode ser útil em modelos globais em que a informação sobre os coeficientes de taxa de infecção por faixa etária é importante. Nesse caso, os índices subscritos indicam apenas a faixa etária, e não mais a região.

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \beta_{kj} \frac{S_k}{N} I_j(t) \quad (47)$$

$$\frac{d}{dt}E_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j(t) - \mu_{eq,k}E_k(t) - E_k(t - d_L) \quad (48)$$

$$\frac{d}{dt}I_k = E_k(t - d_L) - \mu_{k,eq}I(t) - \mu_{k,cov}I(t - d_C) - \gamma_k I_k(t - d_I) \quad (49)$$

$$\frac{d}{dt}R_k = \gamma_k I_k(t - d_I) - \mu_{eq,k}R_k \quad (50)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,ik}N_k - \mu_{cov,k}I_k(t - d_C) \quad (51)$$

$$N_k = S_k + E_k + I_k + R_k \quad (52)$$

O modelo SIR pode ser escrito como:

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \beta_{kj} \frac{S_k}{N} I_j(t) \quad (53)$$

$$\frac{d}{dt}I_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j(t) - \mu_{k,eq}I_k(t) - \mu_{k,cov}I_k(t - d_C) - \gamma_k I_k(t - d_I) \quad (54)$$

$$\frac{d}{dt}R_k = \gamma_k I_k(t - d_I) - \mu_{eq,k}R_k \quad (55)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,ik}N_k - \mu_{k,cov}I_k(t - d_C) \quad (56)$$

$$N_k = S_k + I_k + R_k \quad (57)$$

Por último, o modelo SEAIR:

$$\frac{d}{dt}S_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}S_k - \sum_j \beta_{kj} \frac{S_k}{N} A_j(t) - \sum_j \alpha_j \beta_{kj} \frac{S_k}{N} I_j(t) \quad (58)$$

$$\frac{d}{dt}E_k = \sum_j \beta_{kj} \frac{S_k}{N} I_j(t) + \sum_j \alpha_j \beta_{kj} \frac{S_k}{N} A_j(t) - \mu_{eq,k}E_k - E_k(t - d_L) \quad (59)$$

$$\frac{d}{dt}I_k = \rho_k E_k(t - d_L) - \mu_{k,eq}I_k(t) - \mu_{k,cov}I_k(t - d_C) - \gamma_k I_k(t - d_I) \quad (60)$$

$$\frac{d}{dt}A_k = (1 - \rho_k)E_k(t - d_L) - \mu_{k,eq}A_k(t) - A_k(t - d_I) \quad (61)$$

$$\frac{d}{dt}R_k = \gamma_k I_k(t - d_I) + A_k(t - d_I) - \mu_{eq,k}R_k \quad (62)$$

$$\frac{d}{dt}N_k = \delta_{1k}\Lambda_k N_k - \mu_{eq,k}N_k - \mu_{cov,k}I_k(t - d_C) \quad (63)$$

$$N_k = S_k + E_k + A_k + I_k + R_k \quad (64)$$

3.5.2 Matriz de contato

Um dos grandes problemas está em calcular a matriz de β_{kj} . Essa matriz pode ser escrita como a probabilidade p de um contato resultar na infecção vezes a chamada matriz de contato C_{kj} ($\beta_{kj} = pC_{kj}$). Essa matriz contabiliza o número de contatos que uma população da faixa k faz com uma população da faixa j . Felizmente, encontramos na literatura trabalhos recentes que calculam a matriz de contato para vários países, inclusive o Brasil (Prem *et al* [2]). Nesses trabalhos, a matriz de contato total resulta da soma de matriz de contato de casa C_{kj}^H , matriz de contato do trabalho C_{kj}^W , matriz de contato de escola C_{kj}^S e matriz de contato de outras origens C_{kj}^O

$$C_{kj} = C_{kj}^H + C_{kj}^W + C_{kj}^S + C_{kj}^O \quad (65)$$

Como não sabemos o valor de p , vamos trabalhar com o valor de β , estimado ou pela regressão linear no início da curva ou estimado pelo valor de R_0 . Nesse caso, nós precisamos escalar a matriz, primeiro normalizando-a

$$\tilde{C} = \frac{C}{\det C} \quad (66)$$

para então multiplicá-la por β ($\beta_{kj} = \beta \tilde{C}_{kj}$). Se o β foi obtido a partir do modelo não estratificado, ele pode ser posteriormente refinado para reajustar a curva.

3.6 Cenários

3.6.1 Nenhuma intervenção. Cenário 1

Deve-se utilizar a matriz de contato total dada pela equação 65.

3.6.2 Fechamento de escola apenas. Cenário 2

Nesse caso, considera-se que não há mais contato entre indivíduos no ambiente escolar. A matriz escolar é portanto zerada.

$$C_{kj} = C_{kj}^H + C_{kj}^W + 0 \cdot C_{kj}^S + C_{kj}^O \quad (67)$$

3.6.3 Fechamento de escola e distanciamento social. Cenário 3

Nesse caso, considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + C_{kj}^W + 0 \cdot C_{kj}^S + \mathbf{B} \cdot C_{kj}^O \quad (68)$$

onde \mathbf{A} é uma matriz diagonal, $A_{ii} = 1.5$ para $i = 1$ a 4 e $A_{ii} = 1.1$ para os outros valores de i , e \mathbf{B} é uma matriz diagonal, $B_{ii} = 0.5$ para $i = 1$ a 4 e $B_{ii} = 1.0$.

3.6.4 Fechamento de escola, distanciamento no trabalho e distanciamento social. Cenário 4

Nesse caso, considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade. Considera-se também uma diminuição de 50% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot C_{kj}^W + 0 \cdot C_{kj}^S + \mathbf{B} \cdot C_{kj}^O \quad (69)$$

onde ζ_t é um escalar, com $\zeta = 0.5$ qualquer que seja o $i = 1$. Ela representa uma diminuição de 50% dos contatos no trabalho. Para fechamento do trabalho, basta multiplicar a matriz de trabalho por 0.

3.6.5 Isolamento dos idosos. Cenário 5.

Nesse cenário, os contatos dos idosos ficam restritos à casa. Nesse caso deve-se realizar a seguinte transformação de simetria sobre as matrizes:

$$C_{kj} = C_{kj}^H + (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (70)$$

onde $\tilde{\mathbf{I}}$ é uma matriz diagonal tal que $\tilde{I}_{kk} = 1$ se $k < i_{th}$ e $\tilde{I}_{kk} = 0$ se $k \geq i_{th}$. O índice fixo i_{th} é limiar de idade. Caso o isolamento comece na faixa de 55 anos para cima, $i_{th} = 12$.

3.6.6 Fechamento de escola, distanciamento no trabalho, distanciamento social e isolamento dos idosos. Cenário 6

Nesse caso, os contatos dos idosos ficam restritor à casa. Considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 50% nos contatos dos jovens na comunidade. Considera-se também uma diminuição de 50% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + 0 \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + \mathbf{B} \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (71)$$

3.6.7 Knock Down com isolamento de idosos. Cenário 7

Nesse cenário, os contatos dos idosos ficam restritor à casa. Considera-se um aumento de 50% no contato dos jovens (0 a 20 anos) e um aumento de 10% das outras faixas etárias na matriz casa. A matriz escolar é zerada. Considera-se uma diminuição de 25% nos contatos dos jovens e 50% nos contatos do adultos na comunidade. Considera-se também uma diminuição de 25% nos contatos de trabalho.

$$C_{kj} = \mathbf{A} \cdot C_{kj}^H + \zeta_t \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^W \cdot \tilde{\mathbf{I}}) + 0 \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^S \cdot \tilde{\mathbf{I}}) + \zeta_c \mathbf{B} \cdot (\tilde{\mathbf{I}} \cdot C_{kj}^O \cdot \tilde{\mathbf{I}}) \quad (72)$$

onde $\zeta_t = 0.25$ e $\zeta_c = 0.5$.

3.6.8 Normalização

A normalização ou procedimento de escalar a matriz deve ser feito após somar das matrizes não escalonadas no cenário sem intervenção. O determinante obtido é o fator de normalização para todos os outros cenários.

3.7 Cálculo de grandezas úteis: número de casualidades (C), hospitalizações (H), leitos ocupados (L) e número final total de infectados (M)

Há certas grandezas importantes que não entram no modelo como grandezas dinâmicas em um primeiro momento, mas que podem ser calculadas com base nos parâmetros e nas variáveis dinâmicas do modelo. A seguir vamos mostrar como calcular essas variáveis nos modelos (SIR, SEIR, SEAIR) estratificados. A generalização para o caso territorial é imediata.

3.7.1 Casualidades

A taxa de mortalidade é simplesmente dada por:

$$\frac{dC_k}{dt} = \mu_{k,cov} I_k(t - d_C) \quad (73)$$

onde $\mu_{k,cov}$ é a taxa de mortalidade típica do COVID-19 por faixa etária. Para obter o número $C_k(t)$ em função do tempo, a equação acima deve ser integrada numericamente a partir da função $I(t)$ determinada numericamente. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR. O número total é obtido fazendo a soma $C(t) = \sum_k C_k(t)$

3.7.2 Hospitalizações

A taxa de hospitalização é dada pela equação

$$\frac{dH_k}{dt} = \phi_k (I_k(t - d_H) - H_k(t - d_A)) + \eta_k L_k(t - d_{AU}) \quad (74)$$

onde ϕ_k é a porcentagem média de infectados que precisarão de hospitalização, d_H o tempo típico do início da infecção até a hospitalização e d_A é o tempo típico de alta.

Para obter o número $H_k(t)$ em função do tempo, a equação acima deve ser integrada numericamente, após a integração da equação de leitos em UTI, a partir da função $I(t)$ determinada numericamente. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR. O número total é obtido fazendo a soma $H(t) = \sum_k H_k(t)$.

3.7.3 Leitos ocupados

A taxa de ocupação de leitos é dada pela equação

$$\frac{dL_k}{dt} = \eta_k I_k(t - d_U) - L_k(t - d_{AU}) - \psi_k L_k(t - d_O) \quad (75)$$

onde η_k é a porcentagem média de infectados que precisarão de leitos em UTI, d_U o tempo típico do início da infecção até a hospitalização em UTI, d_{AU} é o tempo típico de alta da UTI, d_O é o tempo típico de óbito na UTI e ψ_k é a porcentagem típica de internações que levam a óbito. Para obter o número $L_k(t)$ em função do tempo, a equação acima deve ser integrada numericamente a partir da função $I_k(t)$ determinada numericamente. Note que a equação acima vale para os modelos SIR, SEIR e SEAIR. O número total é obtido fazendo a soma $L(t) = \sum_k L_k(t)$.

Em nosso modelo, consideramos que um infectado, quando entra no hospital, ou é levado ao leito de UTI ou hospitalizado. Consideramos também que parte dos internados em UTI podem voltar a ser hospitalizados.

3.7.4 Número final total de infectados M

Definimos como número final total de infectados $M(t)$ como o número de indivíduos que pegaram a doença e que ou se recuperaram ou morreram até o instante t . Nos modelos SIR, SEIR e SEAIR esse número é dado simplesmente como

$$M_k(t) = R_k(t) + C_k(t) \quad (76)$$

O número total é obtido fazendo a soma $M(t) = \sum_k M_k(t)$.

3.7.5 Outras saídas importantes

Outras saídas importantes do código são:

- $I_{\max}(t)$: número máximo de infecções;

- $t_{max,I}$: instante em que $I_{\max}(t)$ ocorre;
- $L_{\max}(t)$: número máximo de leitos ocupados
- $t_{max,L}$: instante em que $L_{\max}(t)$ ocorre;
- $H_{\max}(t)$: número máximo de hospitalizações;
- $t_{max,H}$: instante em que $H_{\max}(t)$ ocorre;

4 Parâmetros dos modelos SIR, SEIR e SEAIR

4.1 Parâmetros do modelo SIR, SEIR e SEAIR

Um problema chave na implementação dos códigos é o conhecimento sobre os parâmetros do modelo. Como a quantidade de parâmetros é muito grande, é impossível obter todos eles por ajuste pelo método dos mínimos quadrados ou outra metodologia utilizando apenas informação de infectados notificados no início da epidemia. Por isso, é importante que alguns parâmetros, supostos intrínsecos à natureza da doença, possam ser fornecidos ou estimados a partir de outros estudos. Nós vamos usar essa abordagem, e fazer uso dos dados da curva brasileira para verificar se o conjunto de dados da literatura é consistente. No entanto, para ajustar os dados experimentais, se faz necessário ajustar algum parâmetro. Isso ocorre pois os parâmetros são muito sensíveis ao número inicial de infectados e esse número é quase sempre subnotificado. Por essa razão, é conveniente estabelecer como parâmetro livre para o ajuste o número inicial de infectados, I_0 , variando-os dentro de um limite razoável, que não o exceda em mais de 100%.

Na tabela 1 apresentamos os valores de referência dos parâmetros e algumas observações sobre sua aplicação. São parâmetros globais, baseados no conhecimento mais atualizado sobre a COVID-19. Os valores regionalizados e estratificados estão apresentados em uma planilha anexada como material suplementar.

Tabela 1: Tabelas de parâmetros para modelos SIR, SEIR e SAEIR

Parâmetro	Definição	Valor	Observação	Referência
β	coeficiente de taxa de infecção	$\beta \sim R0/d_I$	A ser inferido a partir de valor de $R0$ e d_I	
$R0$	número de reprodução	2.2 (1.6-3.0)		
C_{ij}	matriz de contato			
Λ	taxa de natalidade	$3.9 \cdot 10^{-5} \text{ dia}^{-1}$	Estimativa para 2015. A multiplicar pela população total.	IBGE, 2013
μ_{eq}	taxa de mortalidade típica			
γ	porcentagem de recuperação			
μ_{cov}	porcentagem de casualidades			
d_I	tempo médio de infecção			
d_L	tempo médio de incubação			
d_C	tempo médio desde infecção até óbito			
d_H	tempo desde início da infecção até hospitalização			
d_U	tempo desde início da infecção até internação			
d_A	tempo típico em hospitalização até a alta			
d_{AU}	tempo típico em UTI até a alta			
Continuação...				

Tabela 1: Continuação.

Parâmetro	Definição	Valor	Observação	Referência
d_O	tempo típico até paciente vir a óbito			
ϕ_k	porcentagem de infectados que necessitam de hospitalização			
η_k	porcentagem de infectados que necessitam de internação em UTI			
ψk	porcentagem típica de internações que levam a óbito			
α	fator de correção para coeficiente de infecção de subclínico			
ρ_i	probabilidade de caso ser clínico			
I_0	Número inicial de infectados			
E_0	Número inicial de expostos			

Referências

- [1] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, M. Jit, P. Klepac *et al.*, “The effect of control strategies that reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china,” *Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working and Jit, Mark and Klepac, Petra, The Effect of Control Strategies that Reduce Social Mixing on Outcomes of the COVID-19 Epidemic in Wuhan, China (3/9/2020)*, 2020.
- [2] K. Prem, A. R. Cook, and M. Jit, “Projecting social contact matrices in 152 countries using contact surveys and demographic data,” *PLoS computational biology*, vol. 13, no. 9, p. e1005697, 2017.