

Prova 2 - EA616

Lucas Bertoloto Pereira - RA 201798

Junho de 2020

Nessa prova, os objetivos são comparar o modelo SI com o modelo usando a curva logística desenvolvido na prova anterior, além de modelar o crescimento da COVID-19 no Brasil usando o modelo SIR no tempo contínuo.

1 Comparação do Modelo SI com a curva logística

Na prova 1, eu achei a seguinte equação para modelar o crescimento de casos da COVID-19:

$$i(t) = \frac{M\alpha}{1 + (M\alpha - 1)e^{-\alpha\gamma t}} \quad (1)$$

em que i é o número acumulado de infectados no dia t da epidemia no Brasil, M é o total da população do Brasil, e α e γ são duas constantes que foram alteradas para ajustar a curva estimada a curva real. Agora, queremos analisar a solução do modelo SI, dado pela seguinte equação diferencial ordinária:

$$\frac{di(t)}{dt} = \beta \frac{(N - i(t))}{N} i(t), i(0) = i_0 \quad (2)$$

Em que, novamente, i é o número de infectados no dia t da epidemia no Brasil. Para resolver essa equação diferencial, eu fiz a seguinte substituição:

$$j(t) = \frac{1}{i(t)} \implies \frac{dj}{dt} = -\frac{1}{i^2} \frac{di}{dt}, j(0) = \frac{1}{i_0} \quad (3)$$

Substituindo (3) em (2), a equação diferencial pode ser escrita como

$$\frac{dj(t)}{dt} = \frac{\beta}{N} - \beta j(t) \quad (4)$$

Assim, achei uma equação diferencial linear de primeira ordem para o modelo. Resolvendo ela, e depois substituindo (3), achei a função de i no tempo como:

$$i(t) = \frac{N}{1 + (\frac{N}{i_0} - 1)e^{-\beta t}} \quad (5)$$

Comparando (1) e (5), observei que a solução do modelo SI também é uma curva logística. O parâmetro N do modelo SI é o $M\alpha$ do modelo anterior, e o parâmetro β é o $\alpha\gamma$ do modelo anterior. Além disso, observei um erro do modelo feito na prova anterior, que contribuiu muito para o modelo não ter sido muito próximo a curva real: o número de infectados no primeiro dia (i_0) foi considerado como 1 na primeira análise, tirando um grau de liberdade do fit da curva.

2 Modelo SIR

Agora, será analisado o modelo SIR, um modelo mais completo que o SI por, além de considerar pessoas suscetíveis e infectadas, também considera as pessoas "removidas", que são imunes a doença ou morreram dela. Ele é dado pelo seguinte sistema de EDOs:

$$\frac{ds(t)}{dt} = -\gamma(\frac{s(t)}{M})i(t), s(0) = s_0 \quad (6)$$

$$\frac{di(t)}{dt} = \gamma(\frac{s(t)}{M})i(t) - \alpha i(t), i(0) = i_0 \quad (7)$$

$$\frac{dr(t)}{dt} = \alpha i(t), r(0) = r_0 \quad (8)$$

sendo $s(t)$ o número de pessoas suscetíveis no instante t , $i(t)$ o número de pessoas infectadas, e $r(t)$ o número de pessoas removidas. Sendo M o tamanho da população, temos também que $s(t) + i(t) + r(t) = M$ para qualquer t , ou, equivalentemente, $s_0 + i_0 + r_0 = M$. Além disso, podemos calcular a taxa de variação do número acumulado de casos $a(t)$ como:

$$\frac{da(t)}{dt} = \gamma(\frac{s(t)}{M})i(t), a(0) = a_0. \quad (9)$$

A importância de encontrar $a(t)$ é que os dados publicados pelo Ministério da Saúde são de casos acumulados, e não o número de casos i naquele dia.

Então, substituindo $s(t) = M - i(t) - r(t)$ na equação (7), e usando essa equação com (8) e (9), o sistema a ser analisado será:

$$\frac{di(t)}{dt} = \gamma(1 - \frac{i(t) + r(t)}{M})i(t) - \alpha i(t), i(0) = i_0 \quad (10)$$

$$\frac{dr(t)}{dt} = \alpha i(t), r(0) = r_0 \quad (11)$$

$$\frac{da(t)}{dt} = \gamma(1 - \frac{i(t) + r(t)}{M})i(t), a(0) = a_0 \quad (12)$$

com a restrição que $i_0 + r_0 \leq M$. Assim, as variáveis que devem ser ajustadas para o fit serão α , γ , i_0 , r_0 e a_0 .

Os dados utilizados para fazer o fit da curva foram retirados de https://github.com/pdpcosta/COVID-19_Brazil, um conjunto dos dados do COVID-19 no Brasil atualizado por pesquisadores da Unicamp e UnB, consultando canais oficiais dos governos estaduais. Na data de escrita desse relatório, estavam disponíveis dados até o dia 103 da epidemia no Brasil, dia 07 de Junho de 2020. A população do Brasil considerada foi $M = 210.147.125$ pessoas.

O programa para fazer o fit foi feito em python, e está no fim da prova. Eu usei a função `odeint` para resolver o sistema de EDOs dado para vários valores dos parâmetros. O ajuste dos parâmetros foi feito pela função `curve_fit`, que permite escolher os valores limite dos parâmetros e um chute inicial para cada um deles. Os valores que escolhi para o fit foram:

- α entre 0 e 1, com chute inicial de 0.5;
- γ entre 0 e 2, com chute inicial de 1;
- i_0 entre 0 e $M * 10^{-6}$, com chute inicial de 1;
- r_0 entre 0 e $M * 10^{-6}$, com chute inicial de 0;
- a_0 entre 0 e $M * 10^{-6}$, com chute inicial de 1.

Os limites superiores das condições iniciais inicialmente eram M , mas observei que assim o α ficava sempre em 0.9999, e r_0 na ordem de 3 milhões, o que não parece razoável pelo que se conhece da epidemia. Usando os limites e chutes iniciais listados, os parâmetros encontrados para o modelo foram:

- $\alpha = 0.8261314887518343$;
- $\gamma = 0.8800607461904922$;
- $i_0 = 189.7943462722114$;
- $r_0 = 9.642906877951573$;
- $a_0 = 0.14642648700053443$.

Então, tracei uma comparação entre os dados do governo com o modelo que fiz:

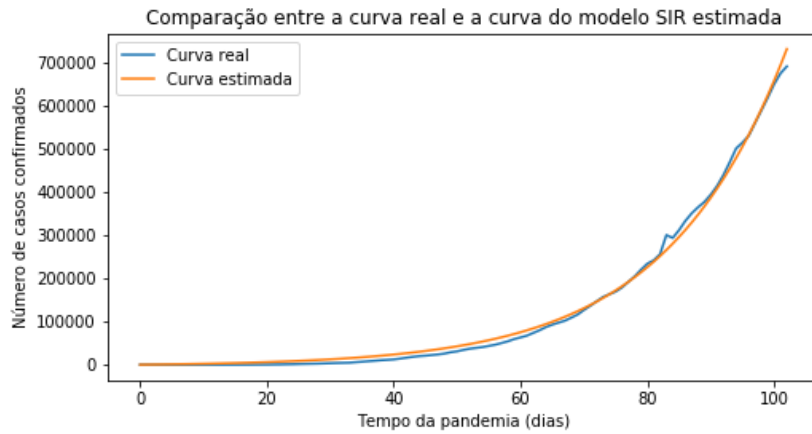


Figure 1: Comparação entre a curva real de casos acumulados de COVID-19 no Brasil com o modelo encontrado.

Observei que a curva se aproxima muito da curva real, o que mostra que o modelo é bom para o conjunto de dados utilizado. Além disso, analisando os parâmetros encontrados, observei que i_0 e r_0 seguem a restrição $i_0 + r_0 < M$ do modelo, mesmo que essa restrição não tenha sido aplicada explicitamente no programa. Esses valores também foram, comparando com a ordem de grandeza de M , muito próximos de 0, o que é esperado, já que s_0 deve ser próximo de M .

Depois, tracei a curva do modelo encontrado até o dia 300 da epidemia, para observar o que ele diz sobre o desenvolvimento futuro da epidemia no Brasil:

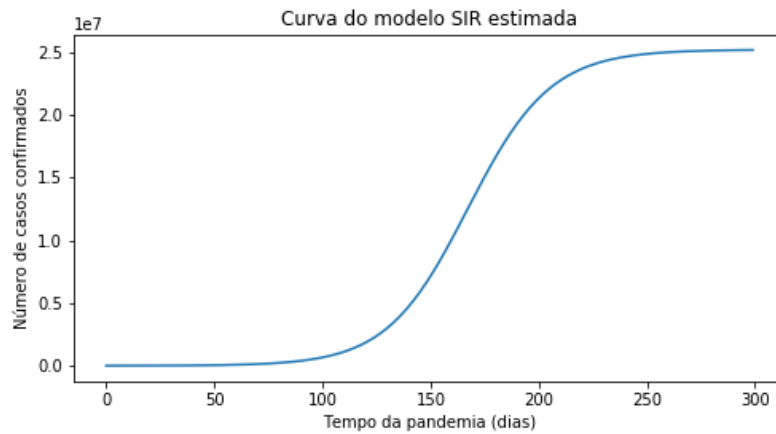


Figure 2: Curva do modelo SIR encontrado até o dia 300 da epidemia.

Observei que, de acordo com o modelo estimado, o número de casos no Brasil deve estabilizar em torno de 25 milhões de casos, próximo a Outubro de 2020.

O código em python desenvolvido para estimar o modelo foi:

```
# função que define o modelo SIR, onde i = variables[0] e r = variables[1]
def SIR(alpha, gamma, variables):
    didt = gamma * (1- (variables[0]+variables[1])/M ) * variables[0]
        - alpha * variables[0]
    drdt = alpha * variables[0]
    dadt = gamma * (1- (variables[0]+variables[1])/M ) * variables[0]
    return [didt,drdt, dadt]

# função que retorna o vetor com os números de a(t), para certos parâmetros dados
def ModelSolver(t,alpha,gamma,initialConditions):

    def LocalModel(initialConditions,t):

        return SIR(alpha,gamma,initialConditions)

    Solution=odeint(LocalModel,initialConditions,t)

    return Solution[:,2]

# função criada para ser usada pelo curve_fit para fazer a estimativa de parâmetros
def SIRModelFit(t,alpha,gamma, i0, r0, a0):

    return ModelSolver(t,alpha,gamma,[i0,r0,a0])

# achando os parâmetros que melhor encaixam o modelo nos dados que possuímos
modelParams = curve_fit(SIRModelFit,t,a, bounds = (0, [1., 2., M/10**6, M/10**6
    , M/10**6]), p0=[0.5,1,1,0,1])

alpha = modelParams[0][0]
gamma = modelParams[0][1]
i0 = modelParams[0][2]
r0 = modelParams[0][3]
a0 = modelParams[0][4]

# resolvendo o modelo para os parâmetros estimados
a2 = ModelSolver(t,alpha,gamma,[i0,r0,a0])
```