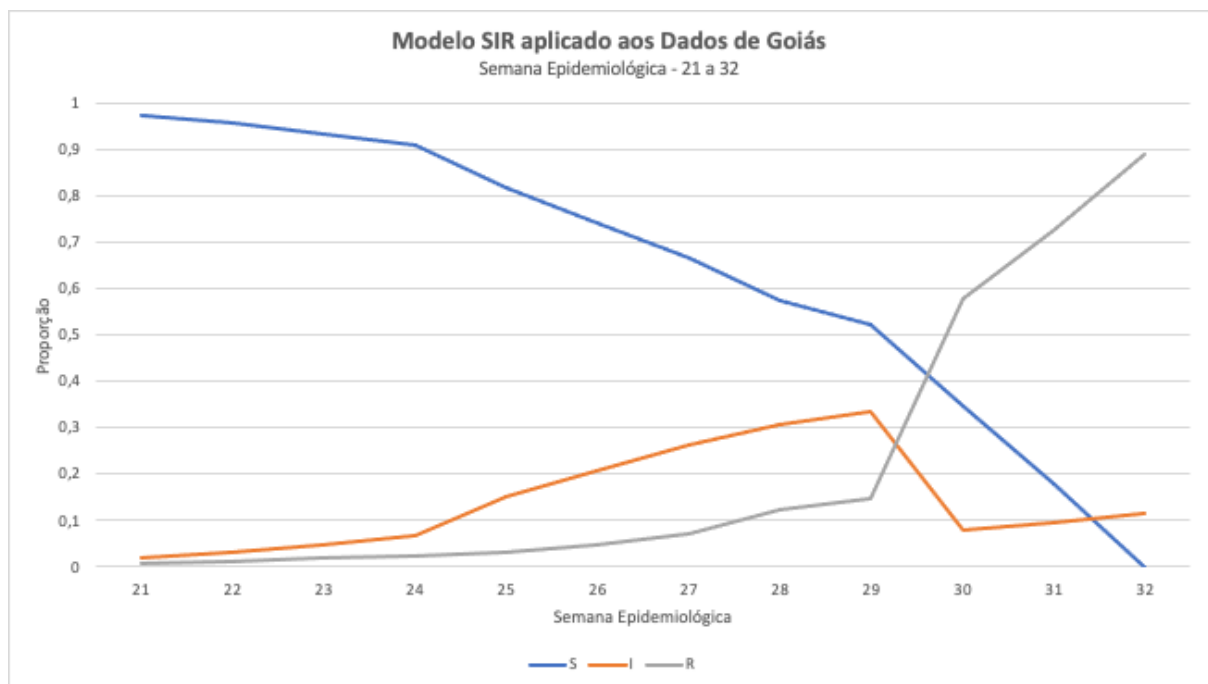


Integrantes: Bruno Kinsch de Lara Campos e Diego Magalhães Rodrigues
Números USP: 12557376 e 12704028

Gráfico 1:



Gráfico 2:



Referência 1:

<https://coronavirus.saude.mg.gov.br/blog/164-taxa-de-transmissao-covid-19>

Referência 2:

<https://raw.githubusercontent.com/wcota/covid19br/master/cases-brazil-states.csv>

Referência 3:

<https://coronavirus.saude.mg.gov.br/blog/81-taxa-de-mortalidade-da-covid-19>

Objetivos

Utilizando dados de infectados e recuperados, referentes à covid 19, do Estado de Goiás até o primeiro pico, criar um programa, baseado no método de Euler e modelo SIR, que modela as informações deste período para então prever o desenvolvimento da doença nas semanas seguintes. E então comparar o modelo com a realidade dos dados para checar sua eficácia, se o modelo não corresponder aos dados factuais então deve-se teorizar o por que de isto acontecer.

Experimento

Primeiro obtemos os dados na referência 2, para facilitar o trabalho com eles, as informações foram suavizadas em uma média móvel de 7 dias, uma semana, ilustrado pela gráfico 1. Determinamos que o período que deveríamos trabalhar, de 17/5/2020 a 8/8/2020. Então utilizando as anotações das aulas do curso de MAP 0131 como base, escrevemos um programa que utiliza o modelo SIR para trabalhar os dados.

O programa depende de valores k e b que são a taxa de recuperação dos infectados e a taxa de infecção da doença, respectivamente. Além disso, ele usa valores S , I e R que são, respectivamente, os números absolutos de suscetíveis, de infectados e de recuperados, estes dados foram obtidos também da referência 2. A partir deles calculou-se os valores s , i e r , que são os números absolutos de suscetíveis/população total, infectados/população total e recuperados/população total, respectivamente. Dessa forma obtemos a porcentagem de cada um desses grupos em determinado momento. Porém conforme a infecção avança estes valores mudam, logo precisam ser constantemente atualizados. Para fazer isso colocamos os dados suavizados em uma média móvel de uma semana em uma tabela na qual o programa busca os valores de s , i e r . O programa usa esses dados e o modelo SIR para gerar uma tabela com os três valores s , i e r em relação ao tempo. Comparamos os gráficos obtidos com um gráfico feito utilizando os dados reais e buscamos adequar nosso modelo para que ele se torne o mais realista possível e também utilizamos o gráfico da realidade como um forma de detectar possíveis problemas, já que os gráficos não poderiam ser muito diferentes.

Nesse sentido, tivemos que achar os valores de k e b que tornassem o nosso modelo mais próximo da realidade. Para fazer isso utilizamos o método dos mínimos quadrados para calcular o menor erro possível enquanto variamos os valores de k e b nos intervalos $[1/15, \frac{1}{6}]$ e $[0, 5]$, respectivamente. Dessa forma achamos os valores que causavam o menor erro possível. Por curiosidade pesquisamos qual seriam os valores reais de k e b , ambos ficaram diferentes da realidade.

Porém mesmo após fazermos isso o nosso gráfico, gráfico 2, ainda tinha diferenças significativas com o gráfico dos dados reais. Estudamos o código buscando possíveis erros, mas não achamos nenhum, tentamos mudar algumas coisas para analisar como o nosso gráfico variava, todavia não obtivemos nenhum gráfico mais acurado. Ao analisar os dados percebemos que em uma determinada semana o número de recuperados apresentava um brusco aumento, então suspeitamos que isso estava afetando nosso modelo.

Semana EPI	Data de Início	Casos Totais	Número de Recuperados	Aumento
25	14/06/2020	15450	2686	129,51%
26	21/06/2020	21865	4170	155,25%
27	28/06/2020	28526	6073	145,64%
28	05/07/2020	36273	10407	171,37%
29	12/07/2020	40795	12516	120,27%
30	19/07/2020	55567	49001	391,51%
31	26/07/2020	69820	61694	125,90%
32	02/08/2020	84963	75350	122,14%
33	09/08/2020	101290	91394	121,29%
34	16/08/2020	116058	106070	116,06%

Por fim decidimos fazer algo diferente, fizemos o programa calcular um período de tempo menor e a partir disso ele previa o resto do período, nesta tentativa obtivemos um resultado muito mais parecido com o gráfico dos dados reais. Apesar de que mesmo assim o gráfico obtido ainda apresentava algumas diferenças em relação a realidade.

Problemas enfrentados

Um dos principais problemas encontrados na criação do modelo foi a grande diferença entre a realidade dos dados e o que o modelo previa. Suspeitamos que um súbito aumento no número de recuperados que ocorreu na semana de 23/7/2020 foi o causador, neste semana os recuperados triplicaram em relação à semana anterior. Já que um aumento tão brusco não foi presenciado antes e nem depois, o que nos levou a crer que este dado destoante do resto causou algum tipo de influência no modelo.

Outro grande problema foi que inicialmente escolhemos um período para análise muito grande. Como consequência a proporção de infectados ficava muito pequena, o que resultava em um gráfico muito pequeno e pouco interessante para analisar em relação à realidade.

Entre outros problemas enfrentados tivemos alguns erros de programação que acabaram por atrasar o experimento, mas não apresentaram grande empecilho no geral.

Conclusões

Nós concluimos que o método de Euler e o modelo SIR não são eficazes para estimar o desenvolvimento da covid 19. Apesar de muitos esforços para tornar os nossos gráficos o mais próximo possível da realidade, ainda assim eles apresentaram defeitos. Por exemplo, os valores de k e b que minimizam o erro quadrático são significativamente diferentes dos valores reais. Portanto, chegamos a conclusão de que ou o modelo utilizado, Euler e o modelo SIR, é muito simples para acompanhar a progressão desta doença ou que por fatores além do nosso controle, problemas com os dados, situações reais complexas que não foram levadas em contas e outros, afetaram o modelo.

Gráfico 3:

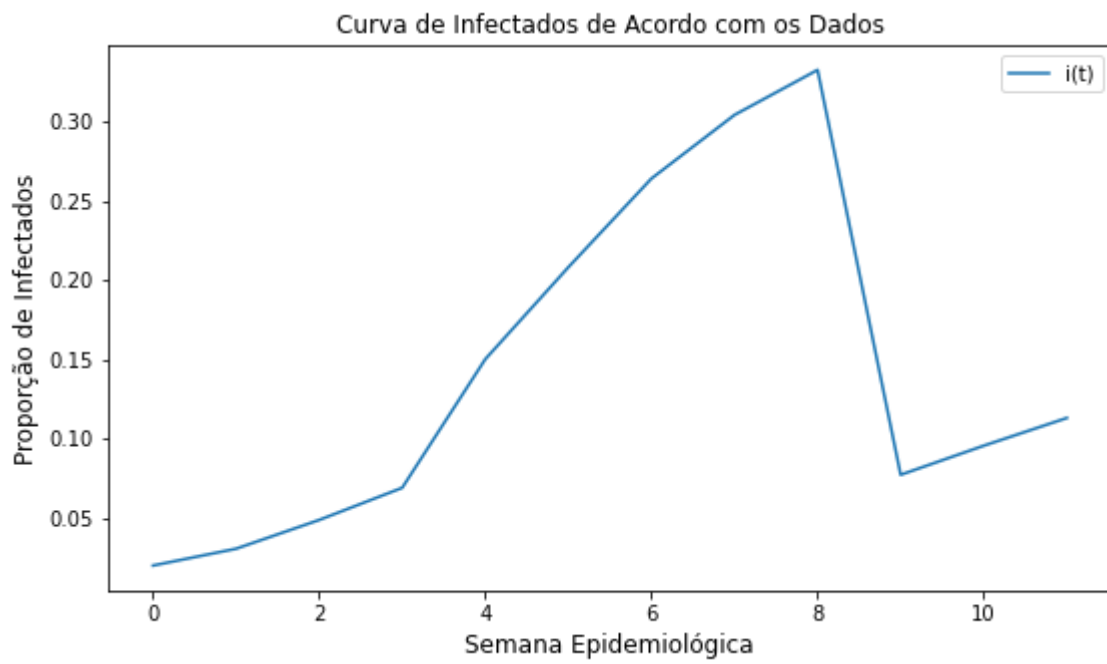


Gráfico 4:

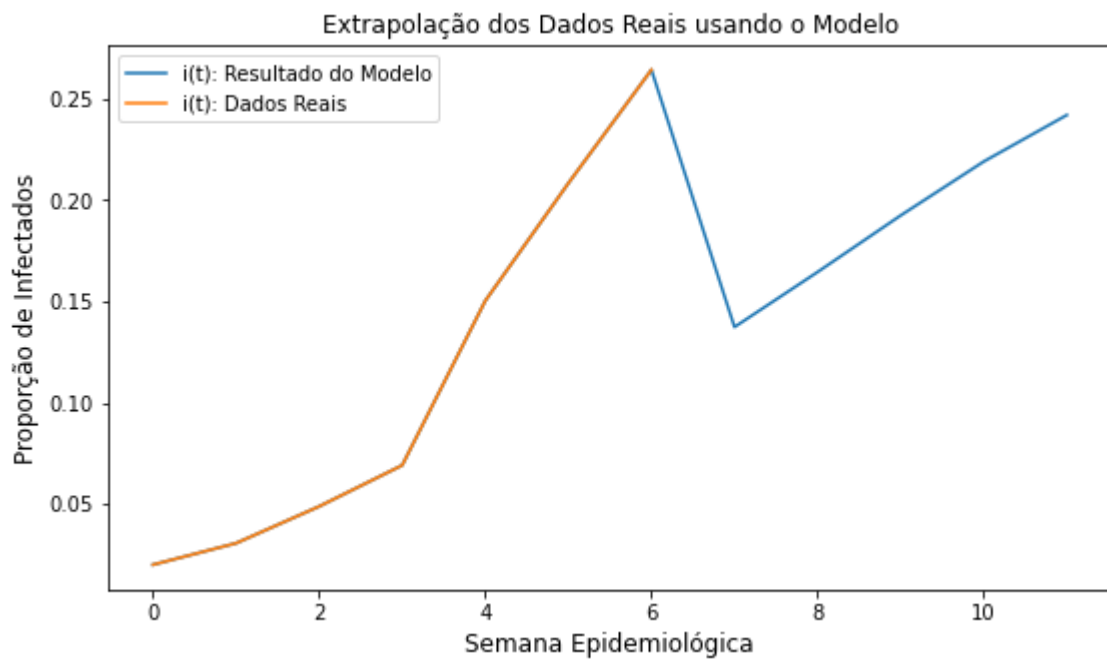


Gráfico 5:

