SIR model analysis for the spread of COVID-19 in Texas

MATH 6345: Mathematical Methods in Medicine and Biology
The University of Texas at Dallas
Dec 07 , 2021

Indrajith Wasala Mudiyanselage

SIR model analysis for the spread of COVID-19 in Texas

Indrajith Wasala Mudiyanselage

1 Introduction

Corona virus outbreak was first reported in late 2019 in Wuhan, China. It spread across the world in no time as it is a highly contagious virus. A widespread occurrence of an infectious disease like COVID 19 (The disease caused by the Corona virus) is called an epidemic. This disease primarily attacks the human respiratory system, and it can cause fatal conditions. The first confirmed case found in the USA was on January 20th, 2020. The state of Texas confirmed its first case on February 13th, 2020. Since then, Texas has been recorded as one of the states with the highest COVID-19 cases. In this project, we have used Texas COVID-19 confirmed cases data from March 5th , 2020, to May 31st, 2021 (A portion of the data set) to model the underlying spread mechanisms.

Various mathematical models are available to study infectious diseases to identify the underlying spread mechanisms that influence the transmission and control of the diseases. Among these models, Susceptible-Infectious-Removed/Recovered (SIR) model, a model that contains three differential equations is widely used in practice. SIR describes how susceptible individuals become infected when encountering an infected person and how infected individuals recover (or die) from the disease and thereafter become immune to getting the disease again. Therefore, the SIR model is one of the simplest and effective models among the mathematical tool that we have to analyze this epidemic.

1.1 SIR Model

Various mathematical models are available to study infectious diseases in order to identify the underlying spread mechanisms that influence the transmission and control of these diseases [1] . Among them Susceptible-Infectious-Removed/Recovered (”SIR”) model, model that contains three differential equations is the foundation of mathematical models of infectious disease, and widely used in practice [2]. SIR describes how susceptible individuals become infected when coming into contact with an infected person and how infected individuals recover from the disease and thereafter become immune to getting the disease again. The SIR model can describe many contagious diseases such as H1N1, SARS (Severe Acute Respiratory Syndrome), Ebola and Zika [3].

When a disease spreads in a population The SIR model split population into three states:

Susceptible(S): The class of individuals who are healthy but can contract the disease. These are called susceptible individuals or susceptibles. The size of this class is usually denoted by S.

Infected(I): The class of individuals who have contracted the disease and are now sick with it, called infected individuals.

Recovered/Removed(R): The class of individuals who have recovered and cannot contract the disease again are called removed/recovered individuals. The class of recovered individuals is usually denoted by R.

The number of individuals in each of these classes changes with time, that is, $S(t)$, $I(t)$, and $R(t)$ are functions of time t.

The total population size N is the sum of the sizes of these three classes $N = S(t) + I(t) + R(t)$.

Basic SIR models based on several assumptions,

- The total population size remains constant
- Individuals are born into the susceptible class
- Susceptible individuals have never come into contact with the disease and are able to catch the disease, after which they move into the infected class.
- Infected individuals spread the disease to susceptibles, and remain in the infected class (the infected period) before moving into the recovered class.
- The infected people could die from the disease but, immune people and dead people cannot spread the disease to others anymore, they can be removed from the future consideration. Since

immune and dead people have the same influence on the model dynamics we can consider them as Recovered (or Removed) class.

A key parameter for all these models was the basic infectious contact number (reproductive number) $\Re_0$ (the product of rate of transmission from an infectious individual and the infectious period) [4]. $\Re_0$ is thus the ratio between infection rate and the recovery rate. If infection rate is denoted by α and the recovery rate is denoted by β

The equation is:

$$\Re_0 = \frac{\alpha}{\beta}$$

If $\Re_0 < 1$, this case shows that the disease will not persist for a long time in the area and will be abolished very soon (disease will not spread). $\Re_0 = 1$ then the disease will remain in the area and will present in the area in any condition with a constant proportion The higher values $\Re_0$, the more infectious the disease will be.

Figure 1: Flowchart of SIR model

1.2 SIRS Model

SIRS model is also mathematical model that compute the number of people infected with a contagious illness in a population over time. The difference between SIR and SIRS model is that SIR model assumes Those infected that recover gain permanent immunity whereas, SIRS model assumes Those infected that recover gain temporary immunity In this case, the SIRS model is used allow recovered individuals return to a susceptible state. The SIR model can describe The SIRS model can describe another class of airborne diseases, for example seasonal influenza.

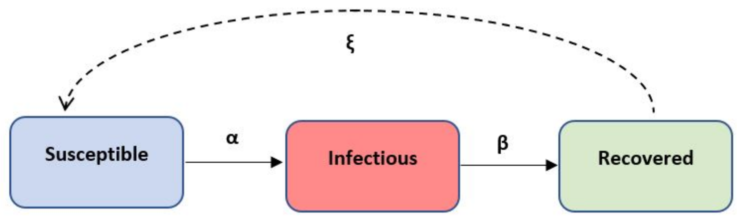


Figure 2: Flowchart of SIRS model

2 Initial value problem

In our model we study the dynamics of the following variables

$S(t)$ = number of susceptible individuals
$I(t)$ = number of infected individuals
$R(t)$ = number of recovered individuals
N = total population size(Does not change over time)

note that,

$$S(t) + I(t) + R(t) = N$$

2.1 System of equations for SIR model

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\alpha}{N}S(t)I(t)\\ \frac{dI}{dt} &= \frac{\alpha}{N}S(t)I(t) - \beta I(t)\\ \frac{dR}{dt} &= \beta I(t)\end{aligned} \tag{1}$$

Where; α = transmission rate parameter (each infected person has a fixed number α of contacts per day that are sufficient to spread the disease)and β = Recovery rate parameter(if duration of infection is 5 days then $\beta = 1/5$)

2.2 System of equations for SIRS model

$$
\begin{aligned}
\frac{dS}{dt} &= -\frac{\alpha}{N} S(t) I(t) + \xi R(t) \\
\frac{dI}{dt} &= \frac{\alpha}{N} S(t) I(t) - \beta I(t) \\
\frac{dR}{dt} &= \beta I(t) - \xi R(t)
\end{aligned} \tag{2}
$$

where ξ is the rate which recovered individuals return to the susceptible statue due to loss of immunity.

3 Parameter Estimation

In different literature, there are many approaches to estimate model parameters.

3.1 Estimate recovery parameter β

Suppose number of infected individuals in a given time is a constant $I(t) = I_0$. Then we get,

$$\frac{dR}{dt} = \beta I_0 \tag{3}$$

This implies that,

$$R(t) = \beta I_0 t \tag{4}$$

If time of recovery is $t = T$ days, and $R(T) = I_0$, or $\beta T = 1$. Therefore we get,

$$\beta \approx \frac{1}{T} \tag{5}$$

3.2 Estimate transition parameter α

In the initial stage of the disease $S(t) \approx N$. Then we can obtain,

$$
\begin{aligned}
\frac{dI}{dt} &= \frac{\alpha}{N} N I(t) - \beta I(t) \\
\frac{dI}{dt} &= (\alpha - \beta) I(t)
\end{aligned} \tag{6}
$$

If the initial number of infections is I_0, then by integration we get,

$$I(t) = I_0 e^{(\alpha-\beta)t}$$
$$\Rightarrow \ln I(t) = (\alpha-\beta)t + \ln I_0 \tag{7}$$
$$\Rightarrow \ln I(t) = mt + \ln I_0$$

We can estimate m by fitting a least square line (Best fitted line) for the initial stage of the disease. Then

$$\alpha = m + \beta \tag{8}$$

4 Numerical result

For the numerical analysis, we have used JHU CSSE COVID-19 dataset ("time series covid confirmed US"). Using this dataset, we filtered Texas confirmed cases from March 5th, 2020 to May 31st, 2021 (Fifteen months period) to fit the SIR model.

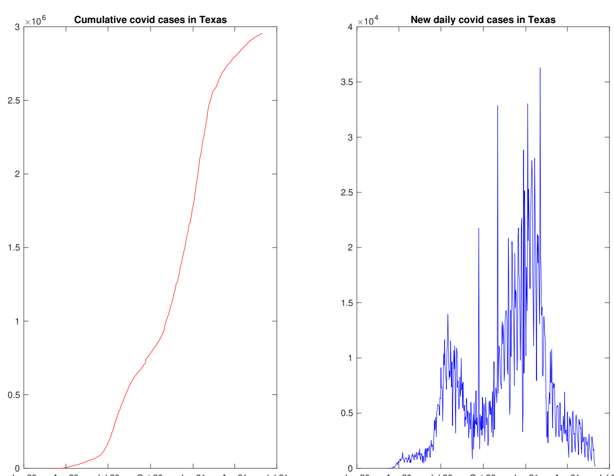


Figure 3: Initial data distribution for the selected time period

4.1 Parameter estimates

4.1.1 Estimation of β

Early research suggested that it could take 2 to 6 weeks to for a COVID-19 patient to recover from the disease. Longer time will take to recover if the case is sever but, most of the cases are mild. Therefore, assume that it will take a patient 4 weeks (28 days) to fully recover form COVID-19.

$$\begin{aligned} \beta &\approx \frac{1}{T} \\ &\approx \frac{1}{28} \\ &\approx 0.035714 \end{aligned} \tag{9}$$

4.1.2 Estimation of α

According to the equation number (7) we need to fit a least square line considering initial stage of the disease. Then by the slope of the line (m) we can estimate the α value. Here, as the initial stage we will consider the first 244 days period (t=0 to t=243) as in this period the number of infections $I(t)$ is less 10000 (Then $S(t) \approx N$) . Note that we assume the total population of Texas N, (29,145,505 by April 2020) is fixed for the selected time period. The estimated $m = 0.071526$. The the transmission rate,

$$\alpha = m + \beta = 0.107240 \tag{10}$$

Then the infectious contract number $\Re_0 = \frac{\alpha}{\beta} = 3.0027 > 1$. Therefore average of 3 people infected by a single infected individual. Therefore epidemic occurs.

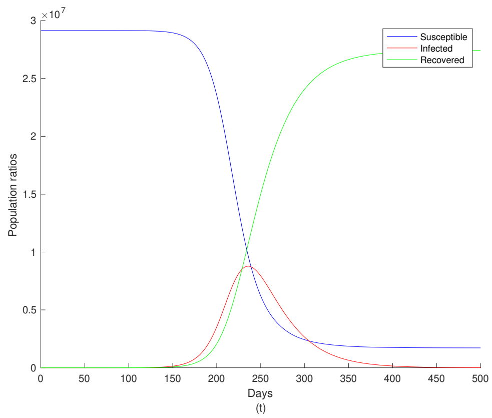
4.1.3 Estimation of ξ (For SIRS model)

If we assume a person's antibodies last after COVID-19 recovery for 8 months. Then a recovered person will fall into susceptible group after 8 moths. Assume,

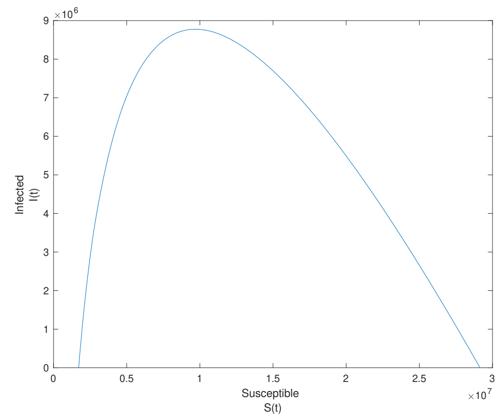
$$\xi = \frac{1}{8(30)} \approx 0.00417 \tag{11}$$

5 Fitted SIR model

Figure 4 gives SIR model results from ODE45 (Assume this as exact solution) for 500 days. Approximately after 500 days, the model reveals that the number of infections will be close to zero and the number of recovered patients will be around 2.7 million. This means 93% of the total population got infected and recovered with in approximately 16 months. Moreover there are 7% of susceptible people in the population will not get infected.



(a) SIR model with constant population



(b) Infected Vs Susceptible plot

Figure 4: Plots for the estimated model

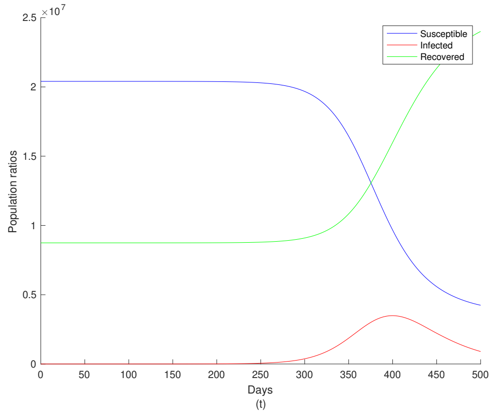
5.1 Intervention Strategies

Consider the following intervention strategies and study the effectiveness of each strategy.

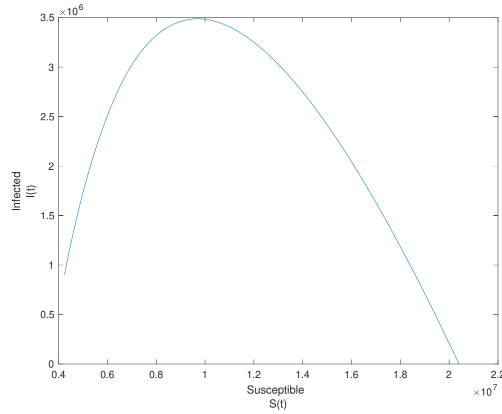
5.1.1 Constant vaccination strategy.

Give yearly vaccinations to proportion p of the population (with $0 < p < 1$), keeping them immune to COVID-19 at all times. This intervention would reduce the total population participating in disease. In other words this will reduce the number of susceptible individuals and we can consider them as recovered from the disease (immune to the disease).

Assume that 30% (p=0.3) of the population had the vacation before the infection. Then in the initial recovered population $R(0) = N \cdot p \approx 8743652$.



(a) SIR model with constant population

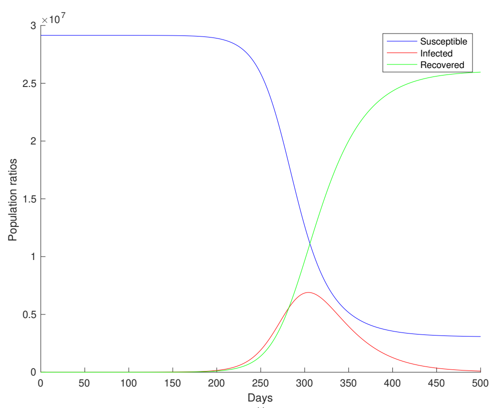


(b) Infected Vs Susceptible plot

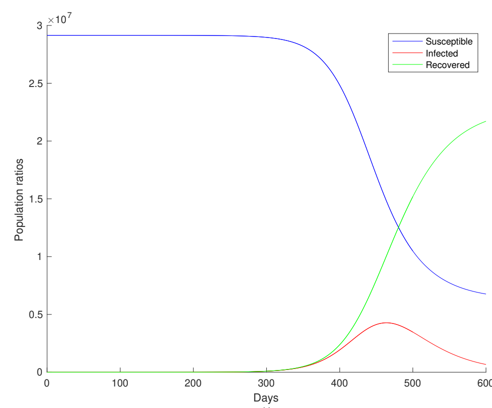
Figure 5: Plots for the estimated model

5.1.2 Intervention by good hygiene

We could reduce the rate of infection by educating the population of good hygiene (such as wearing face masks, washing hands and covering coughs and sneezes) and encourage sick people to avoid contact with others.



(a) SIR model with $\alpha = 0.09$

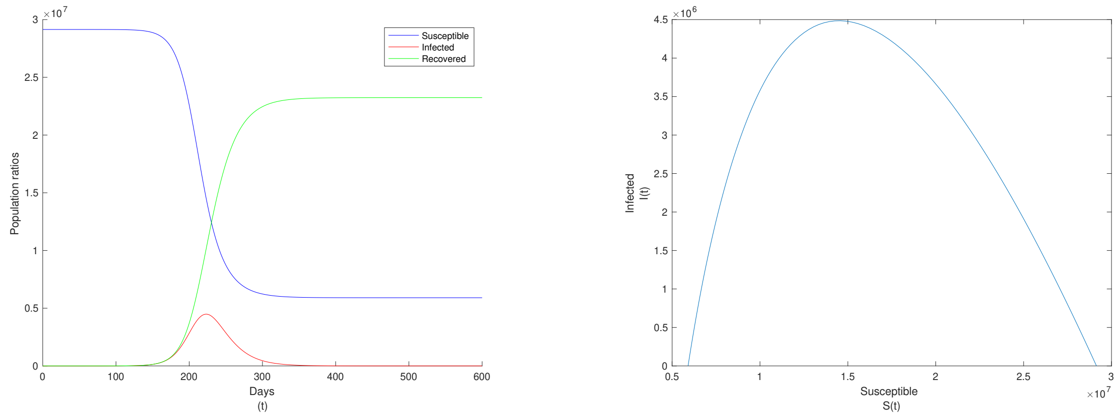


(b) SIR model with $\alpha = 0.07$

Figure 6: Plots for the estimated model

5.1.3 Intervention by drugs

We could administer drugs to help the infectives fight off the infection (increase the rate of recovery). Assume that now it will take 2 weeks to recover then $\beta = 1/14 \approx 0.071428$. Thus $\alpha = 0.142954$



(a) SIR model with $\beta = 1/14$

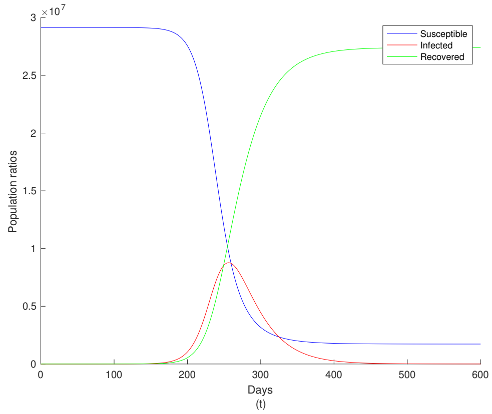
(b) Infected Vs Susceptible plot

Figure 7: Plots for the estimated model

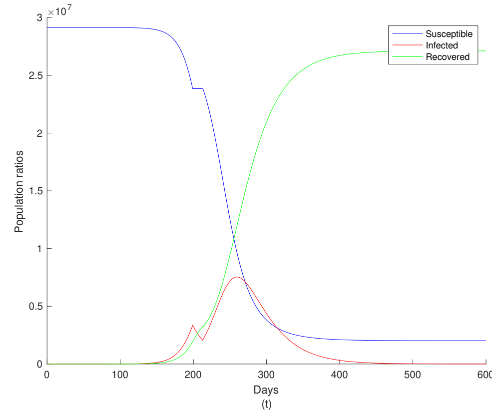
The model reveals that if we can increase the recovery rate (if possible, reduce the time for recovery by 50%) by introducing new treatments methods and drugs, it will significantly decrease the infection rate and quickly control the spread of the disease.

5.1.4 Intervention by school closure

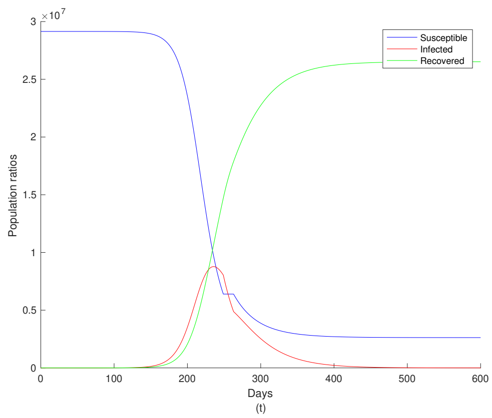
We assume that during the period of school closure, the rate of infection α will be zero. Then we simulate what will happen if you close the school for 14 days during the epidemic. Here we consider 4 possible cases; which are school closure happens after 100 days, after 200 days, after 250 days and finally after 300 days.



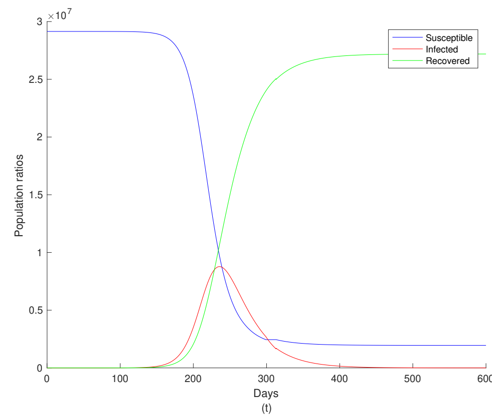
(a) When school closure happens after 100 days



(b) When school closure happens after 200 days



(c) When school closure happens after 250 days



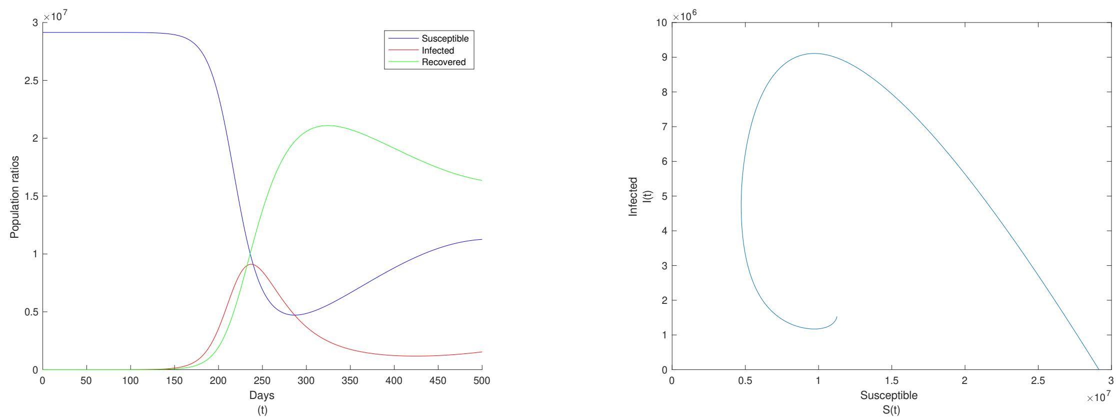
(d) When school closure happens after 300 days

Figure 8: Four possible cases for school closure, after identifying first infections

If the school closure happens after 100 days, there is not much effect on the spread of the disease. But, if the school closure happens approximately around 200 days to 250 days, there is a clear impact on new infections. Therefore, this will reduce the total number of infected people in the population.

6 Fitted SIRS model

Figure 9 gives SIRS model results from ODE45 (Assume this as exact solution) for 500 days. Approximately after 240 days, the model reveals that the number of infections will approach to to zero and again will increase. Also the number of recovered patients will go down while susceptible individuals will increase. It seem like this model is more accurate with the real data.



(a) SIRS model with constant population

(b) Infected Vs Susceptible plot

Figure 9: Plots for the estimated model for $\xi = 1/8(30)$

7 Discussion

Those four intervention strategies (vaccines, good hygiene, drugs, and school closure) could change the pandemic trend with the plot figure by influencing different parameters. Vaccines could reduce the susceptible population while good hygiene decreases the transmission rates, and the drugs make a shorter recovery period. When it comes to school closure, it is better to close schools around 200 to 250 days to control the spread of the disease. It makes sense that in the real world, those first three strategies could be used at the same time so that the pandemic will end shorter than one strategy only if strategies go well.

Based on our result and the rest data do not choose in our project, the SIRS model shows more precise in the COVID-19 pandemic. In the real data after the period we have chosen, the daily confirmed data goes increase again. Those data fit the SIRS model, that recovery individuals start to lose their antibodies and fall back to the susceptible group. It also shows that in an optimal situation, if the COVID-19 could end within around 15 months, the COVID-19 pandemic will terminate and shows a SIR model. Considering the daily confirmed data, it is not not typical to get a single peak figure,. The paper, "Estimation of SIR model's parameters of COVID-19 in Algeria" [7] gives an idea that it may be caused by the difference daily maximum detected or reported ability. Therefore, if we analysis COVID-19 data in different regions or countries, those conditions should be considered to build a SIR model.

Virus mutation is also an important phenomenon in COVID-19. A new mutation may show a different transmission rate (α), recovery rate (β), and immunity loss rate (ξ). If we take mutations into account, the model will be much complicated. So, when we analyze mutations, it may be better to classify data into mutation groups then use different mutations data to build different SIR/SIRS models for the analysis.

References

[1] Siettos, C. I., and Russo, L. (2013). Mathematical modeling of infectious disease dynamics. Virulence, 4(4), 295–306. http://doi.org/10.4161/viru.24041

[2] Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc Royal Society A: Math Phy Eng Sci. 1927;115:700.

[3] Z W, Q W, DJ K. Simulation Study on Effects of Order and Step Size of Runge-Kutta Methods that Solve Contagious Disease and Tumor Models. Journal of computer science and systems biology. 2016;9(5):163-172. doi:10.4172/jcsb.1000234.

[4] Maia Martcheva. book An Introduction to Mathematical Epidemiology - 2015 Series Texts in Applied Mathematics Doi 10.1007/978-1-4899-7612-3 Affiliation University of Florida, Gainesville, United States. Chapter 2 Introduction to Epidemic Modeling. pages9-31

[5] Cooper I, Mondal A, Antonopoulos CG. A SIR model assumption for the spread of COVID-19 in different communities. Chaos Solitons Fractals. 2020;139:110057. doi:10.1016/j.chaos.2020.110057

[6] Adekola HA, Adekunle IA, Egberongbe HO, Onitilo SA, Abdullahi IN. Mathematical modeling for infectious viral disease: The COVID-19 perspective [published online ahead of print, 2020 Aug 17]. J Public Aff. 2020;e2306. doi:10.1002/pa.2306

[7] Lounis M, Bagal DK. Estimation of SIR model's parameters of COVID-19 in Algeria. Bull Natl Res Cent. 2020;44(1):180. doi:10.1186/s42269-020-00434-5

8 Appendix

8.1 MATLAB codes

```
% Data filtering and parameter estimation

clear all;
Texas = readtable('TexasT.csv');

Day=Texas{:,1};
t=Texas{:,2};
C_Tot=Texas{:,3};
D_Tot=Texas{:,4};

figure
subplot(1,2,1)
plot(Day,C_Tot,'r')
title('Cumulative covid cases in Texas')
subplot(1,2,2)
plot(Day,D_Tot,'b')
title('New daily covid cases in Texas')

% Calculate m
regfit=fitlm(t(1:244),log(C_Tot(1:244)/3),'Intercept',false);
regfit;
plot(regfit)

%m=0.071526

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
clear all;

% SIR and SIRS model fitting

%beta=1/28=0.035714   recovery rete
%alpha= m + beta = 0.107240   transmission rate
%xsi=0.0; recovered changing rate to susceptible(This is zero in SIR modle)


a=0.142954;   % transmission rate
b=0.071428;   % recovery rete
c=0; % recovered changing rate to susceptible (This is zero in SIR modle)
N=29145505; %Total population in Texas
p=0; % The population had the vacation before the infection
y0=[29145502-0*N;3;p*N]; % s(0)=N-3=29,145,505-3 , i(0)=3 , r(0)=N*p
h=1; % time interval is one day

f=@(t,yy)[-(a/N)*yy(1)*yy(2)+c*yy(3);(a/N)*yy(1)*yy(2)-b*yy(2);
    b*yy(2)-c*yy(3)];
[T,yy]=ode45(f,0:h:600,y0); % Run 500 days

hold on
r1=plot(T,yy(:,1),'b');
r2=plot(T,yy(:,2),'r');
r3=plot(T,yy(:,3),'g');
hold off

xlabel({'Days','{(t)}'})
ylabel('Population ratios')
```

```
legend('Susceptible','Infected','Recovered')

%For ODE45 Infected Vs Susceptible plot
plot(yy(:,1),yy(:,2))
xlabel({'Susceptible ','{S(t)}'})
ylabel({'Infected','{I(t)}'})

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all;

% Modeling school closure

a=0.107240;  % transmission rate
b=1/28; % recovery rete
c=0; % recovered changing rate to susceptible (This is zero in SIR modle)
N=29145505; %Total population in Texas
y0=[29145502-0*N;3;0*N]; % s(0)=N-3=29,145,505-3 , i(0)=3 , r(0)=N*p
h=1; % time interval is one day

f=@(t,yy)[-(a/N)*yy(1)*yy(2)+c*yy(3);(a/N)*yy(1)*yy(2)-b*yy(2);
     b*yy(2)-c*yy(3)];
f1=@(t,yy1)[0*yy1(1)*yy1(2)+c*yy1(3);0*yy1(1)*yy1(2)-b*yy1(2);
     b*yy1(2)-c*yy1(3)];
f2=@(t,yy2)[-(a/N)*yy2(1)*yy2(2)+c*yy2(3);(a/N)*yy2(1)*yy2(2)-b*yy2(2);
     b*yy2(2)-c*yy2(3)];

[T,yy]=ode45(f,0:h:249,y0); % Run 500 days
[T1,yy1]=ode45(f1,249:h:263,yy(250,:)); % Run 500 days
[T2,yy2]=ode45(f2,263:h:600,yy1(15,:));
```

```
yynew=[yy;yy1;yy2];
Tnew=[T;T1;T2];

hold on
r1=plot(Tnew,yynew(:,1),'b');
r2=plot(Tnew,yynew(:,2),'r');
r3=plot(Tnew,yynew(:,3),'g');
hold off

xlabel({'Days','{(t)}'})
ylabel('Population ratios')
legend('Susceptible','Infected','Recovered')
```