

## Project Proposal

I plan to make a mathematical model of COVID-19. I will use the basic SIR model to do this. My model makes the following assumptions:

1. The population remains constant
2. The rate of infectives is directly proportional to the number of contacts.
3. Infected people recover/die at contact rate

Variables and constants:

S=Susceptible cases ||  $S_0$  =initial number of susceptible cases

I=Infected cases ||  $I_0$  =initial number of infected cases

R=Recovered cases ||  $R_0$  =0

Differential Equations:

Rate of change of total population with time,  $\frac{d(I+R+S)}{dt} = 0$  (Using assumption 1) -(1)

Rate of change of susceptible cases with time,  $\frac{dS}{dt} = -\gamma IS$  (Using assumption 2, where  $r$  is a constant) -(2)

Rate of change of recovered cases with time,  $\frac{dR}{dt} = \alpha I$  (Using assumption 3) -(3)

Rate of change of infected cases with time,  $\frac{dI}{dt} = \gamma IS - \alpha I$  (Using assumption 2 & 3) -(4)

## Theoretical Analysis

With this theoretical analysis of the SIR model used for epidemic spread, I plan to answer 3 questions and draw a conclusion for the answers to these three questions.

Question 1. What determines if the epidemic will spread?

From equation (2), we know that the rate of change of susceptible cases is negative. Thus, with time the number of susceptible cases goes down. Therefore,  $S \leq S_0$ . On plugging this inequality into eq (4), we get the following inequality:

$$\frac{dI}{dt} \leq I(rS_0 - \alpha) \quad - (5)$$

According to inequality (5), the epidemic will spread if  $(rS_0 - \alpha) > 0$ , or  $S_0 > \frac{\alpha}{r}$ .

$\frac{r}{\alpha} = q$ , **contact ratio**. Contact ratio is the fraction of the population that comes into contact with an infected individual during the period they are infectious.

*Conclusion* – Contact ratio  $q$  determines if the epidemic will spread or not. The virus will spread if  $S_0 > \frac{1}{q}$ , as it was in COVID-19's case.

Question 2. What is the maximum number of people that will have the disease at one time,  $I_{max}$ ?

On dividing equation 4 with equation 2, we get the following equation:

$$\frac{dI}{dS} = \frac{\gamma IS - \alpha I}{-\gamma IS} = -1 + \frac{1}{qS}$$

Integrating both sides,

$$I + S - \frac{1}{q} \ln(S) = I_0 + S_0 - \frac{1}{q} \ln(S_0) \quad - (6)$$

The number of infected people at any time is when  $\frac{dI}{dt} = 0$ . Using equation 4, this happens when  $S = \frac{1}{q}$ .

Therefore,

$$I_{max} = I_0 + S_0 - \frac{1}{q}(1 + \ln(qS_0)) \quad - (7)$$

*Conclusion* –  $I_{max}$  is clearly a function of  $q$  as the rest of equation (7) consists of constants. A large value of  $q$  would lead to a large value of  $I_{max}$ . COVID-19 has a large value of  $q$ .

Question 3. How many people will catch the disease in total?

The epidemic ends when the number of infectious cases becomes equal to 0. At the end of the epidemic:

$$I = 0; S = S_{end}; R = R_{end} \quad - (8)$$

Plugging (8) end values into equation (6), we get:

$$S_{end} - \frac{1}{q} \ln(S_{end}) = I_0 + S_0 - \frac{1}{q} \ln(S_0) \quad - (9)$$

When we analyze equation (9), we realize that the value of  $S_{end}$  will be very low if the value of  $q$  is large.  $q$  is large for COVID-19.

We know that at any point in time,

$$R + I + S = I_0 + S_0 \quad - (10)$$

Plugging (8) end values to (10), we get,

$$R_{end} + S_{end} = I_0 + S_0 \quad - (11)$$

*Conclusion* – If  $S_{end}$  is small in value, that it is because of the large value of  $q$  for COVID-19, the value of  $R_{end}$  will be large. Thus, it is likely that a great proportion of the population will catch the disease. This result was also determined by the contact ratio,  $q$ .

After answering all three questions, we notice that **the contact ratio,  $q$ , plays a key role in the spread of the epidemic.**

The basic SIR model I analyzed does not have the power to make a strong prediction of COVID-19's outcomes because determining its spread and susceptibility involves a lot more complex variables which are not accounted for in our simple SIR model. Furthermore, it is still not known if a person can get infected twice.

## Numerical Analysis

Code for graphing the system of differential equations<sup>1</sup>:

```
#required libraries

import numpy as np
from matplotlib import pyplot as plt

class ODESolver:

    def __init__(self, f):
        self.f = f

    def set_initial_conditions(self, U0):
        if isinstance(U0, (int, float)):
            # Scalar ODE
            self.number_of_eqns = 1
            U0 = float(U0)
        else:
            # System of eqns
            U0 = np.asarray(U0)
            self.number_of_eqns = U0.size
        self.U0 = U0

    def solve(self, time_points):

        self.t = np.asarray(time_points)
        n = self.t.size

        self.u = np.zeros((n, self.number_of_eqns))

        self.u[0, :] = self.U0

        # Integrate
        for i in range(n - 1):
            self.i = i
            self.u[i + 1] = self.advance()

        return self.u[:i+2], self.t[:i+2]

class ForwardEuler(ODESolver):
    def advance(self):
        u, f, i, t = self.u, self.f, self.i, self.t
        dt = t[i + 1] - t[i]
        return u[i, :] + dt * f(u[i, :], t[i])
```

---

<sup>1</sup> Winther, Greg, "<https://www.youtube.com/watch?v=mwJXjxMTwAw>"

```

class SIR:
    def __init__(self, alpha, lambda, S0, I0, R0):
        """
        alpha, lambda: parameters in the ODE system
        S0, I0, R0: initial values
        """

    def __call__(self, u, t):

        S, I, R = u

        return np.asarray([
            -self.lambda*S*I, # Susceptibles
            self.lambda*S*I - self.alpha(t)*I, # Infected
            self.alpha*I # Recovered
        ])

if __name__ == "__main__":

    sir = SIR(0.1, 0.0005, 1500, 1, 0)

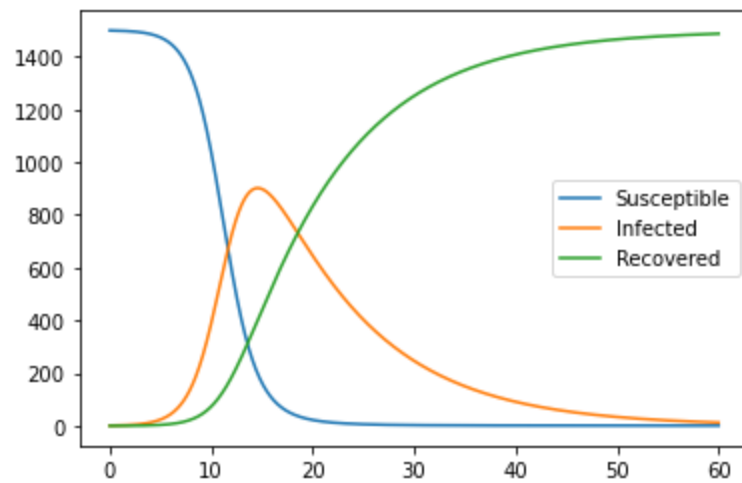
    #last 3 parameters can be changed
    #to get the plot for different initial values
    #Here we have set S0=1500, I0=1, and R0=0.

    solver = ForwardEuler(sir)
    solver.set_initial_conditions(sir.initial_conditions)

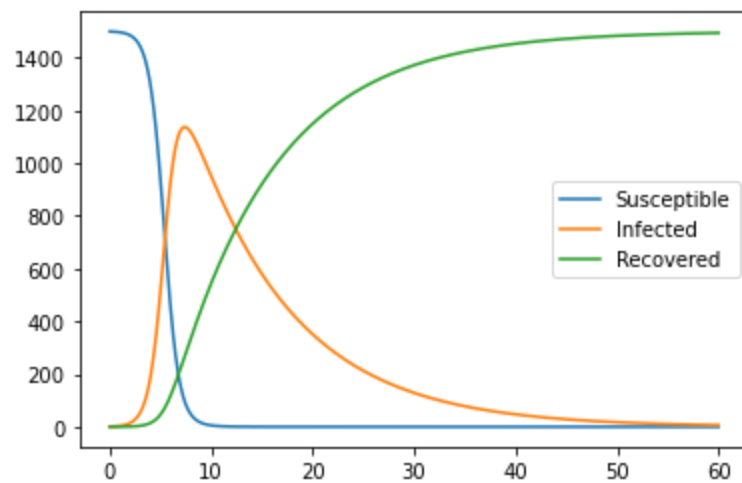
    time_steps = np.linspace(0, 60, 1001)
    u, t = solver.solve(time_steps)

    plt.plot(t, u[:, 0], label="Susceptible")
    plt.plot(t, u[:, 1], label="Infected")
    plt.plot(t, u[:, 2], label="Recovered")
    plt.legend()
    plt.show()

```



Discussion: The graph we obtain provides a very basic idea of how an epidemic spreads. In the theoretical analysis, one of the key takeaways was the key role of contact ratio,  $q$ , in the spread of the disease. If we double the value of  $q$ , we get the following graph:



The key change we notice is the increased value of the maximum number of infected people. This was discussed in question 2 of theoretical analysis. Corona virus has a very large contact ratio. Thus, a large population will be infected by the disease.

We also notice that increasing the value of contact ratio leads to more total people to be infected and recover from the disease by the end of the epidemic. Corona virus has a large contact ratio. This implies that a larger population of people will catch the disease. This was discussed in question 2 of theoretical analysis. The numerical analysis adds weight to our claims in theoretical analysis.

## Final Write-Up

For the final write up I have decided to go a bit deeper on the theoretical analysis by updating the model to be more realistic.<sup>2</sup>

The new model allows people to move around in space. The pandemic can be viewed as spreading through the population in the form of a wave. The initial model only had time dependence. This model will also have space dependence. Adding space dependence,  $x$ , turns our initial system of equations to a system of partial differential equations.

We follow three assumptions for the movement of the population:

1. Susceptible cases don't move  $\frac{\partial S}{\partial t} = -\gamma IS$  ----(1)
2. Infected cases move at a constant rate randomly  $\frac{\partial I}{\partial t} = \gamma IS - \alpha I + D \frac{\partial^2 I}{\partial x^2}$  ----(2)
3. Recovered/Removed cases do not move  $\frac{\partial R}{\partial t} = \alpha I$  ----(3)

To recover a characteristic element in this set of equations to better understand this model, we non-dimensionalize the system of equations. When we non-dimensionalize the system of partial differential equations,  $S$ ,  $I$ , and  $R$  are converted from complete population numbers to population ratios of a given population. The key element we get from non-dimensionalization is reproductive factor,  $R_0$ .  $R_0 = S_0 q$ , where  $S_0$  = initial number of susceptible cases and  $q$  is the contact ratio.

After non-dimensionalization, we also use the technique of change of variables to further simplify the equations. We want a new variable,  $y$ , which gives us a new system of differential equations of  $S$ ,  $I$ , and  $R$  only depending on this new variable  $y$ , eliminating  $t$ , and  $x$  dependence.

$$y = x - ct$$

Plugging  $y$  into equations (1) & (2), we get new derivative equations for susceptible cases (4) and recovered cases (5) with respect to the variable  $y$ . We ignore the new derivative equation for recovered cases with respect to  $y$  for now and only focus on the other two equations.

$$4. \quad 0 = c \frac{dS}{dy} - IS \quad ----(4)$$

$$5. \quad 0 = \frac{d^2 I}{dy^2} + c \frac{dI}{dy} + I(S - \frac{1}{R_0}) \quad ----(5)$$

To solve this set of differential equations we need boundary conditions.

---

<sup>2</sup> <https://www.youtube.com/watch?v=uSLFudKBnBI>



As  $t \rightarrow -\infty$  (Past)

$$y \rightarrow \infty$$

$$I \rightarrow 0$$

$$S \rightarrow S_0 = 1 \text{ (} S_0 \text{ becomes 1 after non-dimensionalizing)}$$

As  $t \rightarrow \infty$  (Future)

$$y \rightarrow 0$$

$$I \rightarrow 0$$

The  $S$  in future or  $S_{end}$  will help us find  $R_{end}$  which will help us find how many people got infected by the disease. I was interested in finding out how the answer to this question would change from the first model which used the most basic SIR model having only time dependency.

Using eq (4), we get,  $IS = c \frac{dS}{dy}$

Plugging this into equation (5), we get

$$\frac{dI}{dy} + cI + c(S - \frac{1}{R_0} \ln(S)) = k, \text{ where } k \text{ is a random constant.}$$

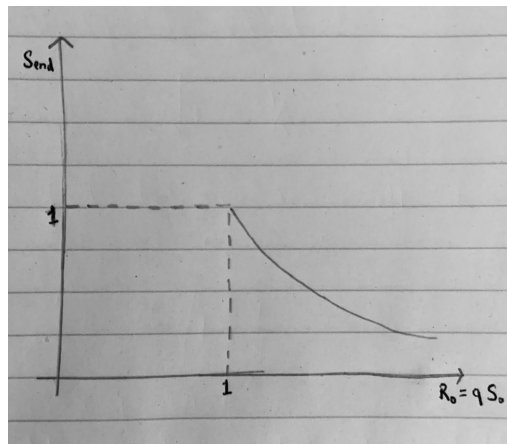
Using initial value condition for  $t \rightarrow -\infty$ , we find that  $k=c$ .

$$\text{Therefore, } \frac{dI}{dy} + cI + c(S - \frac{1}{R_0} \ln(S)) = c.$$

To find  $S_{end}$  we use  $t \rightarrow \infty$  and  $I \rightarrow 0$  which gives us the following equation:

$$S_{end} - \frac{1}{R_0} \ln(S_{end}) = 1$$

We get the following graph which describes the relationship between  $S_{end}$  and  $R_0$ .



Now,  $R_{end} + S_{end} = I_0 + S_0$  and total population,  $R + I + S = I_0 + S_0$ .

Therefore, to have least number of people affected by the virus,  $R_{end}$  we must maximize  $S_{end}$ .

$S_{end}$  can have the maximum value 1 (because of non-dimensionalization) and  $R_0$  can have a minimum value of 1. For  $R_0 = q S_0$ ,  $S_0$  value is fixed but  $q$ , contact ratio, value can be reduced by following

preventive measures against the spread of pandemic. Even though the new model is more realistic compared to our initial SIR model, **contact ratio is still the key factor** deciding the severity of the pandemic. To reduce contact ratio, **we must wash our hands, practice social distancing, and wear masks in public.**

