Modelling spread of COVID-19

Khariton Gorbunov

University of Oxford

Abstract

The rapid spread of the virus COVID-19, also known as the Coronavirus, has resulted in loss of many lives, global economic recession and other issues such as increased racism and xenophobia. The virus is now present in many of the biggest countries in the world, such as China, US, Russia and most of Europe. It is therefore of utter importance to investigate how the virus will spread, the characteristic trends, as well as figuring out how the variables can be tuned to reduce spread, while maintaining social and economic activities.

Keywords: COVID-19, Coronavirus, pandemic, mathematical modelling, simulation

Framework

The framework adapted in this investigation is that of having different states, such as infected portion of the population, recovered portion etc, defining transitional terms between the states. I.e., transition from state A to B is denoted by T_{AB}. Note that the transition terms are not symmetrical, so $T_{AB} \neq T_{BA}$. Transition terms are defined for some period Δt, such that the net flux of these transitions corresponds to the total change of that state per unit time:

$$\frac{A(t+\Delta t) - A(t)}{\Delta t} = \sum_{k \in K} T_{kA} - \sum_{m \in M} T_{Am}$$

Where K denotes the set of all states coming in to A, and M is the set of states which A flows out to.

After figuring out all the particular transitions, these equations can be converted to sets of ODEs and solved to generate the characteristic curves.

This is the framework adapted by the SIR model[1], which we will modify and look at further in this work.

1. SI model

Firsty, we will considere the simplest model which includes two groups and only, the susceptible and infected, having a single transition T_{SI}. Define the variables as following:

- N - number of people in the system
- S - number of people susceptible to infection
- I - number of people infected

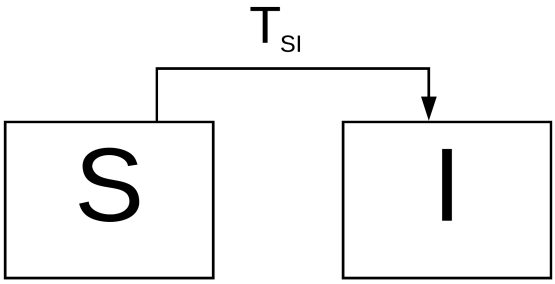

Figure 1: SI diagram

This means people can only get sick and never recover, and although highly unrealistic, it will build a foundation for future models. Consider an influx of people over timeframe Δt into the infected group:

$$I(t + \Delta t) - I(t) = T_{SI}\Delta t$$

In this case, T_{SI} will be the expected number of people being infected per unit time over Δt. If we call the random variable that denotes number of people people being infected over Δt as i, then

$$T_{SI}\Delta t = \mathbb{E}[i]$$

Consider how people are infected: infections are transmitted through interactions, and for every interaction, a person is either infected or not. Therefore,

a single interaction is a Bernoulli trial, hence for n interactions taking place in Δt will be a binomial distribution dictating number of infected people i. The aim is therefore to get the expected value of i. We know that

$$i \sim B(n, p_{single})$$

so

$$\mathbb{E}[i] = np_{single}$$

p_{single} is the probability of infection in an interaction. For that to happen, two things must occur: the interaction has to be between an infected and a non infected person, as well as the virus must be passed on

$$p_{single} = P(\text{transmission} \cap \text{infectious interaction})$$

Simplify to:

$$p_{single} = P(\text{transmission}|\text{infectious interaction})P(\text{infectious interaction})$$

$P(\text{transmission}|\text{infectious interaction})$ is the probability of virus transmission given a susceptible and infected person interact, and will depend on factors such as immunity, mode of transmission (cough, handshake) etc. It is therefore instructive to keep as a hyperparameter, denoting it with a lowercase p. For the latter term, we need to pick two people, where one of them is infected. Considering picking people to be independent events with uniform probability, we get the following:

$$P((I \cap S) \cup (S \cap I)) = 2P(I)P(S)$$
$$2P(I)P(S) = 2\frac{I}{N}\frac{S}{N} = \frac{2IS}{N^2}$$

Combining everything gives:

$$\mathbb{E}[i] = \frac{2pISn}{N^2}$$

This is for an arbitrary n interactions over time Δt, so we can instead express this as an interaction rate $r = \frac{n}{\Delta t}$, such that

$$I(t + \Delta t) - I(t) = \frac{2pISr\Delta t}{N^2}$$

Dividing by Δt and letting it go to zero gives us the following ODE:

$$\frac{dI}{dt} = \frac{2pISr}{N^2}$$

Since we only have two states, whatever goes into I comes out of S, so

$$\frac{dS}{dt} = -\frac{dI}{dt}$$

and can be formally shown noting that $S + I = N$.

We can solve the equations by simply substituting $S = N - I$:

$$\frac{dI}{dt} = \frac{2prI(N-I)}{N^2}$$

and given initial conditions $I(t_0) = I_0$, has the expected logistic solution:

$$I(t) = \frac{N}{1 + \frac{N-I_0}{I_0}\exp(\frac{-2pr(t-t_0)}{N})}$$

and the characteristic curves displayed on figure 2.

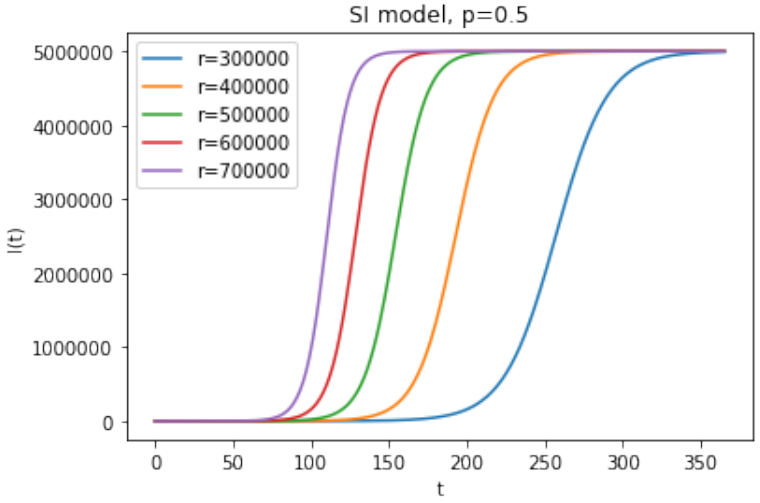

Figure 2: SI model

2. SIR - modified

The SIR model [1] deals with 3 states:

- S - number of susceptible people
- I - number of infected people
- R - number of recovered people

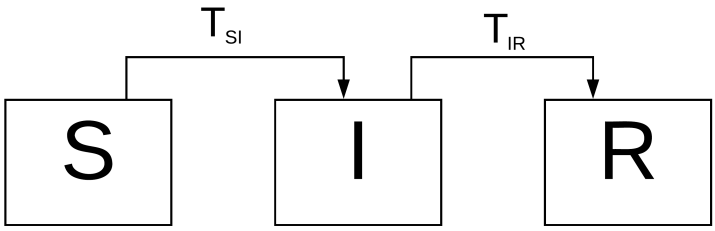

Figure 3: SIR diagram

It assumes that every transition is proportional to the real time quantities of the states, and gives rise to the following equations:

$$\begin{aligned}\frac{dS}{dt} &= -T_{SI} \\ \frac{dI}{dt} &= T_{SI} - T_{IR} \\ \frac{dR}{dt} &= T_{IR}\end{aligned}$$

The transition T_{SI} is of similar form to what we developed in the previous chapter, however, I believe that the rate of recovery should not necessarily be proportional to the number of infected people.
The proposal is to consider a person to be recovered after a fixed number of days being infected, let that period be τ. That means that the transition T_{IR} is equal to the transition T_{SI} just delayed by τ. What that means in practice, is that everyone who got sick over a period Δt at time t will recover

over the same time period Δt at time $t + \tau$
So we amend the term $T_{IR}(t) = \gamma I$ to become

$$T_{IR}(t) = T_{SI}(t - \tau)$$

Note that T_{SI} is simply the flux leaving S, so $S(t + \Delta t) - S(t) = -T_{SI}$.
T_{SI} derived in previous section remains the same, as the only thing changed is N, which now encapsulates R as well, such that $N = S + I + R$. We can therefore rewrite our governing equation for the net flux into I:

$$I(t + \Delta t) - I(t) = \frac{2pISr\Delta t}{N^2} + S(t + \Delta t - \tau) - S(t - \tau)$$

dividing through by Δt and letting it go to zero gives:

$$\frac{dI(t)}{dt} = \frac{2pISr}{N^2} + \frac{dS(t - \tau)}{dt}$$

Depending on the method of solution, it may be more convenient to rewrite the equation as:

$$\frac{dI(t)}{dt} = \frac{2pr[I(t)S(t) - I(t - \tau)S(t - \tau)]}{N^2}$$

The total system can be then summarized as following:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{2pISr}{N^2} \\
\frac{dI(t)}{dt} &= \frac{2pISr}{N^2} + \frac{dS(t - \tau)}{dt} \\
\frac{dR(t)}{dt} &= -\frac{dS(t - \tau)}{dt}
\end{aligned}$$

Curves for different interaction rates displaying number of infected people as a function of time can be found in figure 4.

Also, we can plot different curves for the various recovery periods τ, displayed on figure 5.

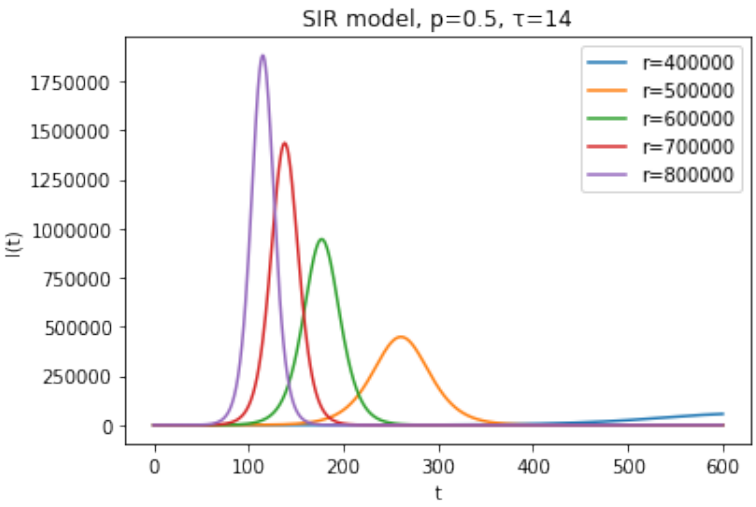

Figure 4: SIR model for various rates of interaction

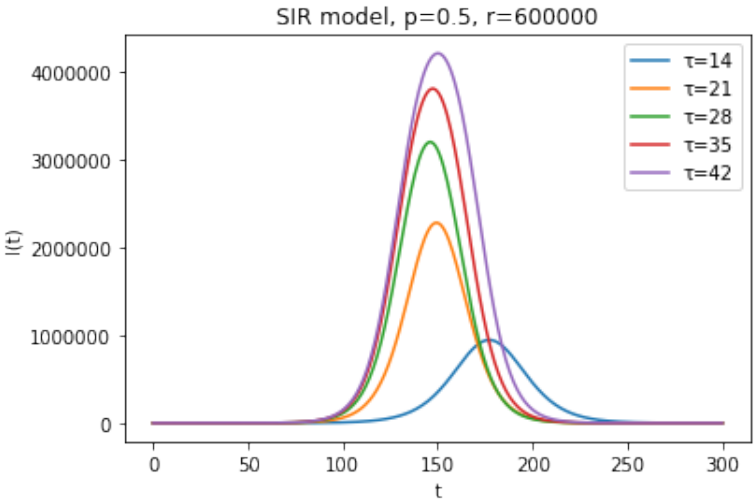

Figure 5: SIR model for various recovery times

3. Modelling incubation and perfect quarantine

Consider an SIQR model:

- S - number of susceptible people

- I - number of infected people in the incubation period. Not showing symptoms but are infectious
- Q - number of infected people in quarantine with perfect isolation.

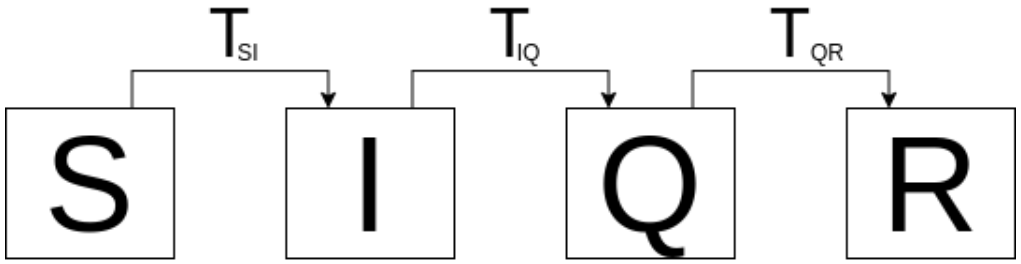

Figure 6: SQIR diagram

Suppose that once a person starts showing symptoms, they are placed in a perfect quarantine. That implies that a person from Q cannot interact with anyone from S, and transition T_{SI} therefore becomes

$$T_{SI} = \frac{2pISr}{(S+I+R)^2}$$

. For incubation of period τ_i, T_{IQ} will be a delayed version of T_{SI}, so

$$T_{IQ}(t) = T_{SI}(t - \tau_i)$$

. As in the modified SIR model, we will say that there is a constant time for people to recover, which dictates quarantine time τ_q. That means sickness period from being infected is given by $\tau_i + \tau_q$, and

$$T_{QR}(t) = T_{IQ}(t - \tau_q) = T_{SI}(t - (\tau_i + \tau_q))$$

Another thing to consider is the fact the interaction rate $r = \frac{n}{\Delta t}$ will decrease when people are quarantined, as they no longer interact with the rest of the society. We can compensate for that by assuming that there is an average interaction rate per person $\bar{n}$, which means that total number of interactions over a Δt is $n(t) = \bar{n}(S + I + R)$. We know that $n(0) = \bar{n}N$ and $n(0) = r(0)\Delta t$, so $r(t) = \frac{r(0)}{N}(S + I + R)$.

Hence, we can write down our set of equations:

$$\begin{aligned}
\frac{dS(t)}{dt} &= -\frac{2pISr}{(S+I+R)^2} \\
\frac{dI(t)}{dt} &= \frac{2pISr}{(S+I+R)^2} + \frac{dS(t-\tau_i)}{dt} \\
\frac{dQ(t)}{dt} &= -\frac{dS(t-\tau_i)}{dt} + \frac{dS(t-(\tau_i+\tau_r))}{dt} \\
\frac{dR(t)}{dt} &= -\frac{dS(t-(\tau_i+\tau_r))}{dt} \\
r(t) &= \frac{r(0)}{N}(S+I+R)
\end{aligned}$$

Note that at ay given time, the total number of infected people is $I(t) + Q(t)$

4. Parallels to statistical mechanics and spatial simulations

Simple physical models of gasses often neglect inter-molecular forces, and in the simplest case, monatomic gasses can be modelled as Newtonian spheres which elastically collide in a confined space. From these assumptions, we can derive the collision frequency of these atoms [2]. Consider N spherical atoms with diameter d existing in a volume V that move at an average speed v. Over an interval Δt, one of these atoms will sweep a volume of $Cfv\Delta t$, where C is the collision cross section area and f is the relative speed factor to take into account different speed directions, and both depend on the dimension of the problem (2D or 3D). In that volume, we expect there to be an average of $\frac{N}{V}Cfv\Delta t$ collisions, so the rate of collision for that particular sweep is $\frac{NCfv}{V}$. To count up the total number of collisions for all the actors in the system, we multiply that rate by N, and divide by two to avoid double counting. Therefore, the collision rate of the system is:

$$r = \frac{CfvN^2}{2V}$$

Suppose that the atoms actually represent two different species A and B. For every collision, there is a finite probability that species A will become

species B when the two collide. To find the rate of collision, we can do a similar analysis, and we arrive at

$$r_{AB} = \frac{N_a N_b C f v}{V}$$

If species A has a probability of morphing into species B on collision with a probability p (in statistical mechanics that would actually be dictated by the kinetic energy distribution and would be an intrinsic property of the system), we can say that the change in species B over time Δt is given by:

$$N_b(t + \Delta t) - N_b(t) = \frac{N_a N_b C f v p}{V} \Delta t$$

. We can let the interval go to zero, and we obtain a differential equation:

$$\frac{dN_b}{dt} = \frac{N_a N_b C f v p}{V}$$

This already resembles our system by form, but there are physical parameters that do not make sense in a human interaction simulation such as cross section and relative velocity. Therefore, we can eliminate the factor $\frac{Cfv}{V}$ by substituting in rate of collision, and our new equation becomes

$$\frac{dN_b}{dt} = \frac{2rpN_a N_b}{N^2}$$

.

Which is of exactly the same form as the term T_{si} in the modified SIR model. What this means is that the differential equations can be simulated using a spatial simulation of colliding objects, and a rate of interaction can be fixed by adjusting the intrinsic parameters such as energy of the system and cross sectional area of the colliders.

There are however limitations on actual simulations, and that is that the density is assumed to be uniform, but in reality, there will be density gradients due to the mechanics of the system, which will arguably simulate real human interaction better due presence of clustering in social systems.

Bibliography

References

[1] Wikipedia contributors, Compartmental models in epidemiology — Wikipedia, the free encyclopedia, 2020. [Online; accessed 16-March-2020].

[2] P. Atkins, J. de Paula, Physical Chemistry for the Life Sciences, W.H. Freeman and Company, New York, NY, 2006.