A case study of the 2019-2020 Lombardy COVID-19 outbreak using a SIR Model.

Umberto Zerbinati

March 11, 2020

Abstract

1 SIR Model

The SIR model is a system of ODEs design to model the dynamics of an epidemics not affected by birth and death. This model was originally developed by Kermack-McKendrick in 1927, and is expressed by the following ordinary differential equations:

$$
\begin{aligned}
S'(t) &= -\beta S(t)\frac{I(t)}{N} \qquad N = S + I + R, \\
I'(t) &= \beta S(t)\frac{I(t)}{N} - \gamma I(t) \qquad S(0) = S_0, \\
R'(t) &= \gamma I(t) \qquad I(0) = I_0,\ R(0) = 0.
\end{aligned}
\tag{1}
$$

In particular β is defined as the product between the average number of contacts per person times the probability the disease transmit at each contact, and $\frac{1}{\gamma}$ is the duration of the infection in days. An other important quantity is the average number of secondary infection, i.e. R_0, in particular in SIR model it is defined as $R_0 = \frac{\beta}{\gamma}$. This is known to be a some sort of magic number when it comes to epidemic spread. If we define the two quantity:

$$
x(t) = \frac{S(t)}{N} \quad y(t) = \frac{I(t)}{N}
\tag{2}
$$

we can rewrite (1) as:

$$
x'(t) = \beta x(t)y(t),
\tag{3}
$$

$$
y'(t) = (\beta x(t) - \gamma)y(t),
\tag{4}
$$

there the initial growth in infected is given by the ODE:

$$
y'(t) = (\beta x(0) - \gamma)y(t),
\tag{5}
$$

this tell us that if $R_0 x(0)$ is greater then one we will have an exponential growth in the infected. More detail regarding the significance of $R0$ in particular related to the COVID-19 outbreak can be found in [1]. The R_0 number for the COVID-19 has been estimated around 3.28, more information can be found in [2]. More sophisticated model that take into have been introduced to model pandemics. Birth and death can be included in the SIR model, but given the fast dynamics of the COVID-19 outbreak this was not necessary. Model where the immunity is not achieved after the infection, called SIS model, have been developed, but given the limited period and the localization of the case study those model where not taken into consideration. Last but not least model that take in consideration age structure can be used, and given the way COVID-19 affect differently elderly population, this would have been appropriate. But due to the complexity in building the contact matrix was decided not to use those type of models. More complex deterministic model have been applied to the Chinese COVID-19 outbreak in [3] and [4]. All the model here discussed are deterministic model, stochastic model are an other valid approach. In particular we can see a branching model applied to the COVID-19 outbreak in [5]. More information about the epidemic modeling can be found in [6].

2 Data

Our aim was to fit the SIR model to the COVID-19 outbreak in Lombardy, to find the right configuration of two parameters i.e. γ, β. We estimated those parameters confronting model simulation with data collected from the first 12 days of the outbreak, starting from the 21st of February to the 3rd of May. We limit the data taken in consideration to the first 12 days because due to change of the testing policy in Lombardy the number of new infection could not be considered completely reliable. In particular we decided to discharge the data from the day the mortality rate surpassed the world average, which is 3.4% according to [7], by more then the variance in the death rate, i.e. 0.52.

3 Fitting The Model

As previously mentioned is important to chose the most accurate β and γ parameters for the model to correctly represent the COVID-19 outbreak. The γ parameter was fixed such that $\gamma^{-1} = 9$ this assuming the COVID-19 has an active infective period equal to the period from ofset to hospitalization. In particular this number comes from [8] and [9] and the assumption that the γ parameters is the same between the Lombardy and the Whuan outbreak. We then estimated the β value that minimize the following error,

$$\mathcal{E}(I, R, \{i_n\}, \{r_n\}) = \sum_{k=1}^{d} \frac{|I(k) - i_k|}{i_k} + \frac{|R(k) - r_k|}{r_k}, \tag{6}$$

where $\{i_n\}$ and $\{r_k\}$ are the recorded data from the Lombardy outbreak, as described in section 2, and d is equal to 12. To minimize $\mathcal{E}$ we used a gradient descent method with starting point β_0 ranging in $[0.1, 2.0]$ and γ_0 ranging in $[\frac{1}{23}, 1]$. In particular automatic differentiation based on centered finite difference approximation is used to compute the gradient. The parameters for which the model best fitted the data, according to the error function $\mathcal{E}$, were found to be $\beta = 0.76$ and $\gamma = \frac{1}{6.8}$. This give us an initial $R0$ number equal to 5.2, which placed the $R0$ number for the Lombardy outbreak just below the worst estimate for the Hubei province outbreak, [2].
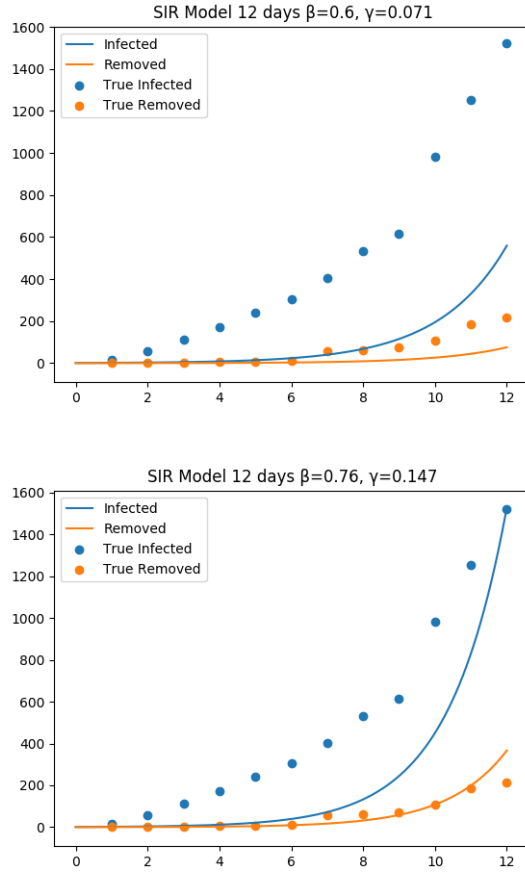

Figure 1: The figures above show the how the SIR model changes after the optimal β is computed.

3

4   Results

Once we run the simulation for the SIR modeling over a full period of 120 days we find out that the peak in infection will be reached between the 27th and the 31st day of the infection. Furthermore the critical proportion of the people to be vaccinated, to gain hard immunity, in case a vaccine is developed need to be around the 80% of the population. This model did not take into consideration the measure that were adopted by the Government after the 1st of March to control the spread of COVID-19.
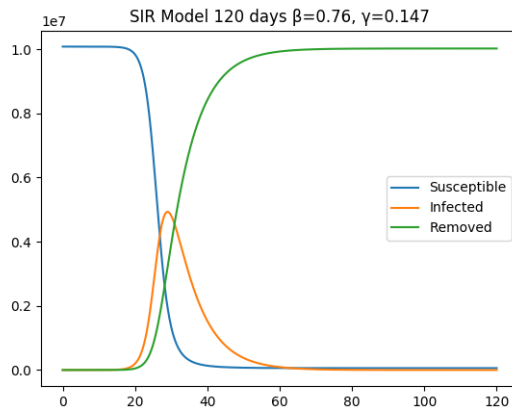

Figure 2: The figures above show the simulation for the above described SIR model over the duration of 120 days.

References

[1] Giulio Viceconte and Nicola Petrosillo. Covid-19 r0: Magic number or conundrum? *Infectious Disease Reports*, 12(1), 2020.

[2] Ying Liu, Albert A Gayle, Annelies Wilder-Smith, and Joacim Rocklöv. The reproductive number of covid-19 is higher compared to sars coronavirus. *Journal of Travel Medicine*, 2020.

[3] Jonathan M Read, Jessica RE Bridgen, Derek AT Cummings, Antonia Ho, and Chris P Jewell. Novel coronavirus 2019-ncov: early estimation of epidemiological parameters and epidemic predictions. *medRxiv*, 2020.

[4] Biao Tang, Xia Wang, Qian Li, Nicola Luigi Bragazzi, Sanyi Tang, Yanni Xiao, and Jianhong Wu. Estimation of the transmission risk of the 2019-ncov and its implication for public health interventions. *Journal of Clinical Medicine*, 9(2):462, 2020.

[5] Joel Hellewell, Sam Abbott, Amy Gimma, Nikos I Bosse, Christopher I Jarvis, Timothy W Russell, James D Munday, Adam J Kucharski, W John Edmunds, Fiona Sun, et al. Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*, 2020.

[6] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals.* Princeton University Press, 2011.

[7] T. Adhanom. "who director-general's opening remarks at the media briefing on covid-19 - 3 march 2020", 2020.

[8] Jantien A Backer, Don Klinkenberg, and Jacco Wallinga. Incubation period of 2019 novel coronavirus (2019-ncov) infections among travellers from wuhan, china, 20–28 january 2020. *Eurosurveillance*, 25(5), 2020.

[9] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy SM Leung, Eric HY Lau, Jessica Y Wong, et al. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*, 2020.