Covid-19 via the SIR model

August 22, 2020

The year 2020 will be remembered, amongst other things, because of the Covid-19 pandemic that hit virtually every country in the world. Currently with a death toll of over 400000, it became a real threat, making our governments take measures unseen for most of us. In this project you will use the SIR model to study the pandemic and to determine how these decisions impacted the spread of the virus. It can be written individually or in pairs. Each group will be assigned a unique country/state/region in the world. Please select the country in the spreadsheet `https://docs.google.com/spreadsheets/d/1Oipj8tAGOABxYMyS8Cp6m-5tnhmXTDcJYlmn_vNks5o/edit#gid=0`.

- Model: Basic SIR model with immunity. Recovered people R can't catch the disease. For the purposes of this model, dead people can be counted together with the recovered.
- Data: Several websites have public data of all countries. Local governments or agencies may also have publicly available more detailed data on specific countries. You will need to download and store data of daily new cases, deaths and recovered. The data has to include between 10-20 days before and after the lockdown. You will also need the country/region's population, and the date when strict lockdown measures were adopted. Recall that S, I and R are fractions of population, so make sure you have the data in the right units.

  Some useful websites:

  - `https://www.worldometers.info/coronavirus/`
  - `https://ourworldindata.org/coronavirus-data#comparing-data-sources`
  - `https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdowns`
- Estimating parameters: To estimate the parameters you may use regression tools from any statistics classes you may have taken. If this isn't the case, you may alternatively:
  - Plot the graphs $I(S)$, compare to various values of the parameters and choose the best fit.
  - Discretize the equations and average over days. For example, the value of β can be determined with the data of recovered R and infected I people. Approximating $dR/dt = \beta I(t)$ by $R(t+1) - R(t) = \beta I(t)$, then plugging in values and solving for β. You will get a different value every day, from which you can take the average.

- Submission format: The answer should be typed and submitted on a .pdf file.
- NOTE: Not all countries provide the same data. Please contact us if you don't find the needed info. In particular, if data on recovered cases is not available, you may take $\beta = 1/10$.

1. (10 points) In lecture we showed that solution curves of the SIR model with $S(0) \simeq 1$ and $I(0) \simeq 0$ are graphs of the function

$$I(S) = -S + \frac{\beta}{\alpha}\ln(S) + 1.$$

   Let $\rho = \beta/\alpha$.

   (a) Determine the maximum number of infected people in terms of ρ.

   (b) The quantity $R_0 = 1/\rho$ is called "basic reproduction number." Justify why if $R_0 < 1$, the epidemic cannot get started.

   (c) Assume $\alpha = 0.2$ and $\beta = 0.1$. What is the maximum number of infected people at a time? How many people get infected after the epidemic is over?

2. (10 points) Consider the data of the spread **before** official lockdown dates. Use technology to answer the following questions:

   (a) Determine the value of β. As mentioned above, you may approximate $dR/dt = \beta I(t)$ by $R(t+1) - R(t) = \beta I(t)$ and average over all days.

   (b) Determine the value of α that best fits the data.

   (c) Plot the graph I versus S with the parameters determined above and compare it with the graph of the real data.

3. (10 points) Repeat the previous question with the data **after** the lockdown.

4. (10 points) Compare the values of α and R_0 before and after. Explain why you wouldn't expect β to vary. Comment on the validity of the SIR model in this case.

5. (BONUS points) Estimate the maximum number of infected people and total number of infections in the long term for both situations.