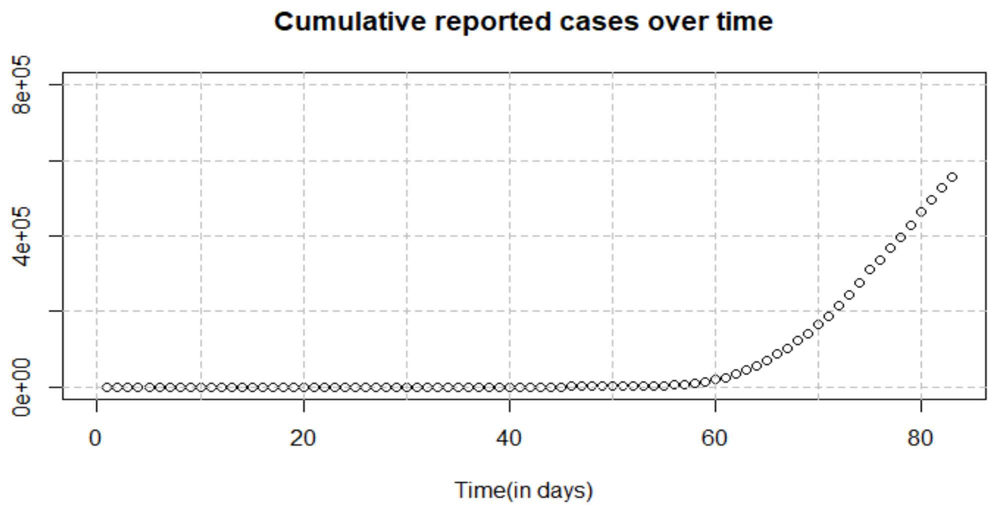
Estimating transmission rate of deterministic mathematical model analysis for COVID-19 in US using least squares method

Giannis Roussos

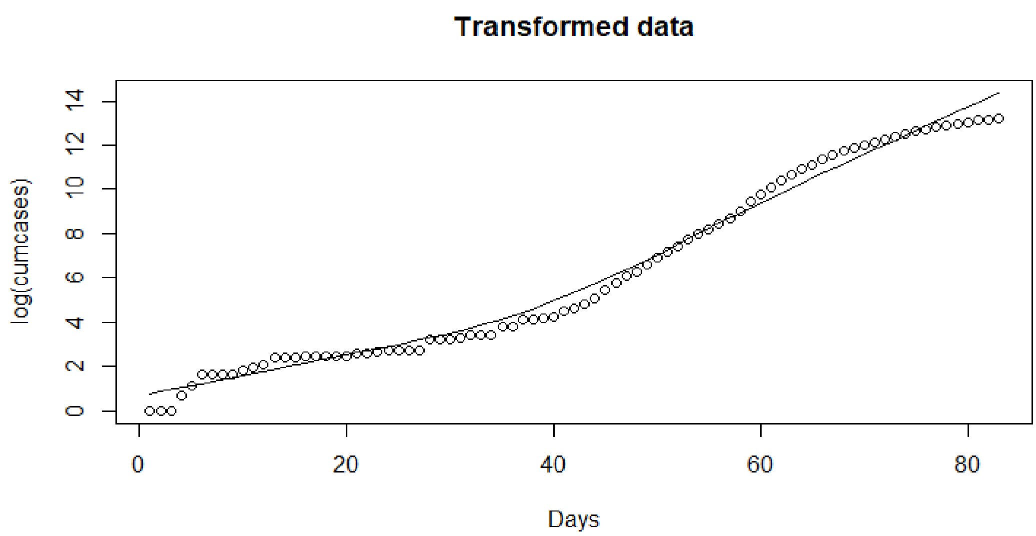
21st April 2020

Data

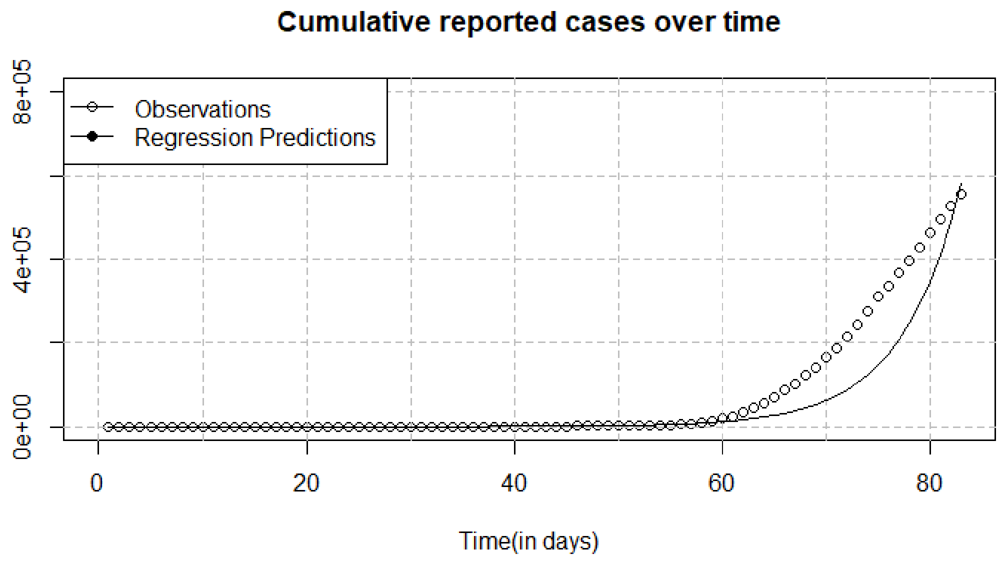
In this project we study the spread of corona virus 19 in US from 21/1/2020 until 83 days after, thus until 12/4/2020.



From the graph above we expect that the spread follows exponential growth. To investigate this more, we transform the cumulative cases to log(cumulative cases) and we plot the data to see if they fit a straight line or not.



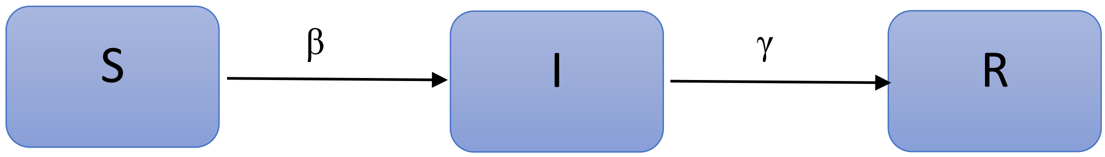
As we see from the plot above the transformed data do not exactly create a straight line, however we can assume that predictions from exponential model are close for describing the pandemic. After fitting an exponential regression model to the data, we predict the cumulative reported cases that we have found from the regression and we plot it together as it is shown to the following plot. As we mentioned before, we do not expect the predictions to fit very good with the data.



SIR Model

In order to study the pandemic we construct an SIR deterministic model. In this model we assume that the general population N remains constant, due to some births and deaths. We discriminate the population to three compartments: "S": susceptibles who cannot transmit the virus, "I": reported infected who transmit the virus, "R": removed after treatment. We do not take into account assymptomatic cases. The susceptibles become infected with a rate β which is the transmission rate that we are going to try to predict. The infected people become removed with rate γ after completing the treatment.

Flow-Diagram



Equations

1) $\frac{dS(t)}{dt} = -\beta \cdot \frac{S \cdot I}{N}$

2) $\frac{dI(t)}{dt} = \beta \cdot \frac{S \cdot I}{N} - \gamma \cdot I$

3) $\frac{dR(t)}{dt} = \gamma \cdot I$

Where $\frac{d(S+I+R)}{dt} = 0 \Rightarrow N = constant$, which is consistent to our hypothesis of the model.

From our model we have that: $S \leq S_0$ and from equation (2) $\Rightarrow$

$\frac{dI(t)}{dt} < I \cdot (\frac{\beta \cdot S_0}{N} - \gamma)$ and in order for the infected people to increase there must be: $\frac{\beta \cdot S_0}{N} - \gamma > 0 \Rightarrow \frac{\beta \cdot S_0}{N \cdot \gamma} > 1 \Rightarrow RE_0 = \frac{\beta \cdot S_0}{N \cdot \gamma} > 1$ (4) , where RE_0 is the basic reproductive number and by estimating β we can also estimate RE_0.

Parameters

Parameters	Symbol	Values	Source
Total population number	N	328.200.000	fixed
Initial susceptibles number	S_0	$N-I_0-R_0$	fixed
Initial infected number	I_0	1	fixed
Initial removed number	R_0	0	fixed
Transmission rate	β	-	fitted
Removal rate	γ	$\frac{1}{14}$	fixed

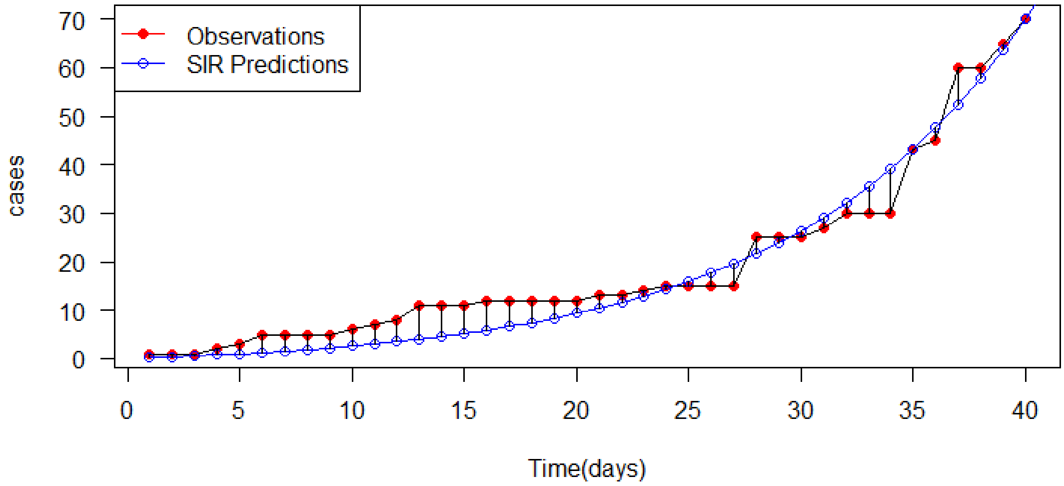
Estimate transmission rate

To estimate β we split our data into 2 parts, the first part is until day 40 and the other part is from day 40 to day 83. We do this because we see from the data that from day 40 until 83 infected people increase much more than from day 1 to day 40 and so we expect the β to change during time. To estimate the transmission rate we use the method of least squares.

And from the first 40 days we estimate that β is:

```
b
## [1] 0.165293
```

And by simulating for this value of the transmission rate we have the following results:



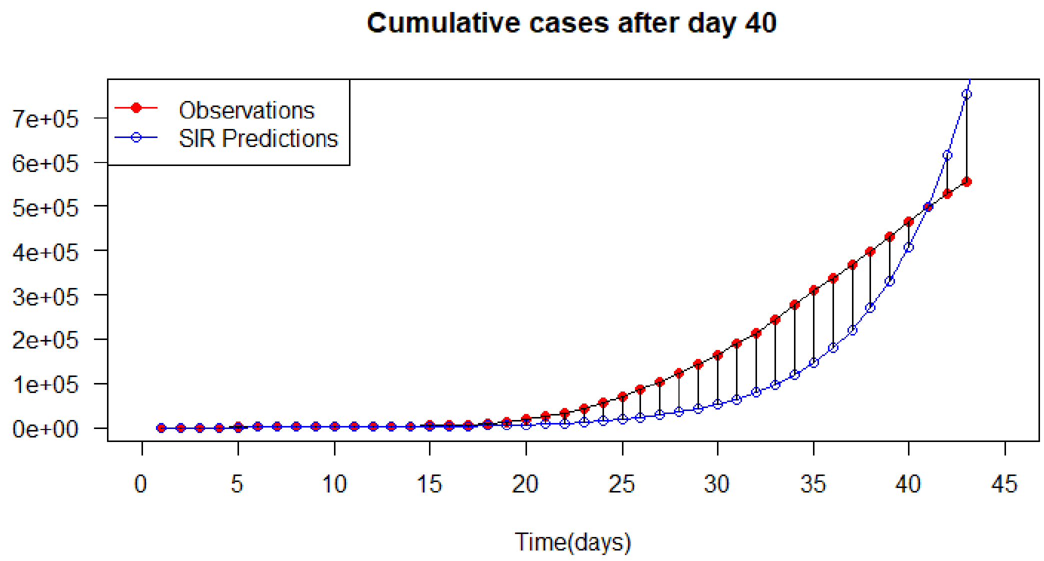
And from equation (4) we estimate the basic reproductive number RE_0:

```
RE0
## [1] 3.886025
```

After day 40 we use again the method of least squares and we find that β is:

```
b
## [1] 0.2775732
```

And by simulating the model for this value of the transmission rate we have the following result:



Thus, totally we have: