Estimation of coronavirus epidemic in "Andhra Pradesh, India" using the SIR model

- Ramdas Omkar Prasad (Roll no.: 1903213)
- Nikhil Yuvraj (Roll no.: 1903211)

Abstract:

Presently, the world is infected by the COVID-19 virus. For controlling the spreading of the virus, we have to do futuristic calculation for infection spreading & accordingly, have to take safety measures. The coronavirus affected the population of the entire world. The nations like USA, Italy, which are considered to be the best in medical services, are also suffering from this pandemic. Hence, from their experience, we, the Indians, have to take more safety measures in order to fight against Corona.

Here, we are using SIR model to study the pattern of spreading of Corona virus in **Andhra Pradesh, India**. From this pattern, we will try to predict the end of this pandemic from the available data, under certain assumptions.

Introduction:

Virus is a nano meter in size which reproduces inside the living cell of an organism. Presently Coronavirus has created a health emergency to the world population and became a pandemic. Initially, this virus is transferred from the bat to the human. Further this virus shows the human to human transmission. Covid-19 virus is spreading to people by the respiratory droplets and contact mode. Previously there are several mathematical models reported. The SIR model is a simple and effective model which can give the prediction of different pandemic situation.

First case of Coronavirus was reported in Wuhan, China in November, 2019. Now, it has spread over 198 countries in the world. However, the first case in India was reported on 30 January, 2020. In Andhra Pradesh, first case was reported on 12 March, 2020. The growth in cases in Andhra Pradesh started on 20 March and now, number of active cases have reached more than 4500. This study explains this whole epidemic growth by using the SIR Model for Andhra Pradesh which can help to control this epidemic.

Methods:

SIR Model:

The **SIR Model** is one of the simplest compartmental models, and many models are derivatives of this basic form. The model consists of three compartments:

- ➢ **S** for the number of Susceptible i.e. individuals having no immunity but are not infected.

➢ **I** for the number of Infected i.e. active cases.
➢ **R** for the number of Recovered i.e. recovered with immunity and deceased individuals.

Also, it assumes that within the outbreak period, no significance population change takes place (e.g., through new births, deaths, migration etc.) and N = S + I + R = Constant.

The SIR model can be expressed by the following set of ordinary differential equations:

Rate of change of Susceptible Population: $\frac{dS}{dt} = -\frac{\beta SI}{N}$

Rate of change of Infected Population: $\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I$

Rate of change of Recovered Population: $\frac{dR}{dt} = \gamma I$

This model was implemented in python using SciPy for solving the differential equations. Key influencers here are the parameters β and γ which are described as:

➢ β: Decides how fast the disease can be transmitted due to exposure. An interesting point here is, this β can be different for the same kind of virus in different societies. A society with less socialization will have a lower value of β.
➢ γ: is a parameter expressing how much part of infected people can be recovered in a specific period

SIR model for Andhra Pradesh, India:

Andhra Pradesh is a state in the South India. It has a population of **49,634,314**.

Variables Used:

N - Total population

S(t) - Number of people susceptible on day t

I(t) - Number of people infected on day t

R(t) - Number of people recovered on day t

β - The expected amount of people an infected person infects per day (normalised by N i.e. β = β/N)

γ: - The proportion of infected recovering per day

As the lockdown was declared in the whole country just few days after rise in the number of infected cases has begun in the Andhra Pradesh, we started our analysis directly assuming that there was lockdown from start!!

Initial values of unknown parameters:

N =49,634,314

I(0) =1

R(0) =0

S(0) =N-R(0)-I(0) =49,634,313

We also used following assumptions for curve fitting along with real data:

Assumptions:

- The Total Population, N is fixed. In other words, we will assume that time scale is short enough that births and natural deaths (other than deaths from this disease) can be neglected.
- The only way a person can leave the susceptible group is to become Infected (no quarantine).
- Recovered individuals can't be infected again because either they have developed immunity or died.
- Age, Sex, social status, and race do not affect the probability of being Infected.
- There is no inherited immunity.
- The member of the population mixes homogenously (have the same interactions with one another to the same degree).
- A person become Infectious as soon as he/she gets the disease.
- No vaccine is available.

With These assumptions we can Fit our model to following equations obtained after Integrating the set of first-order ODE's defined above with some approximations:

$$
\begin{aligned}
&\text{i.} \quad S(t) = s0 * e^{-\frac{\beta R(t)}{\gamma}} \approx s0 \left(1 - \frac{\beta R(t)}{\gamma} + \frac{\beta^2 R(t)^2}{\gamma^2} * \frac{1}{2}\right) \\[2mm]
&\text{ii.} \quad R(t) = \rho^2 * \frac{\frac{s0}{\rho} - 1}{s0} + \frac{\rho^2}{s0} * \alpha * \tanh\left(\frac{\gamma \alpha t}{2} - \emptyset\right) \\[2mm]
&\text{iii.} \quad \frac{dR}{dt} = \frac{\gamma \rho^2 \alpha^2}{2s0} * \text{sech}^2\left(\frac{\gamma \alpha t}{2} - \emptyset\right) = \gamma I(t) \\[2mm]
&\text{iv.} \quad I(t) = \frac{\rho^2 \alpha^2}{2s0} * \text{sech}^2\left(\frac{\gamma \alpha t}{2} - \emptyset\right)
\end{aligned}
$$

Here S(t), R(t), I(t) denotes the Number of Susceptible, Recovered, Infected People on day t.

And constants $\gamma, \beta, \rho, \alpha, \emptyset$ are related by the following expressions:

- $\rho = \frac{\gamma}{\beta}$

- $\alpha = \sqrt{\left(\frac{s0}{\rho} - 1\right)^2 + 2s0 * \frac{(N - s0)}{\rho^2}}$

- $\emptyset = \tanh^{-1}\left[\frac{\left(\frac{s0}{\rho} - 1\right)}{\alpha}\right]$

Curve fitting procedure:

- For Curve fitting and error minimization, we used curve_fit() function from scipy. optimize library of python.

- After searching for data of Andhra Pradesh regarding model on various websites (mentioned in citation), we get data for Infected and Recovered compartments.
- However, there were many unexpected variations in Infected data, so we use the Recovered data (which is actual data) to fit it to the curve R(t).

Coding:

I used the following function to fit the curve:

```python
import numpy as np
def recover(x,c0,c1,c2):
    return c0*np.tanh(c2)+c0*np.tanh(c1*x-c2)
```

**Here constants c0,c1,c2 are defined as:**

$C0 = \rho^2 * \alpha / S_0$

$C1 = \gamma * \alpha / 2$

$C2 = \emptyset$

One point to note is that in actual equation, there were only two unknowns α and γ. But, now we have three. The reason is that the actual expression was difficult to fit to the curve, so we convert it to the above exressions with three independent variables. But, this will lead to **three different values** of α and γ (from each of the two). So, I calculate all of them and check the error by taking **average value**. Error was less than **2%.**

R(t) curve after fitting (data upto 19 May):

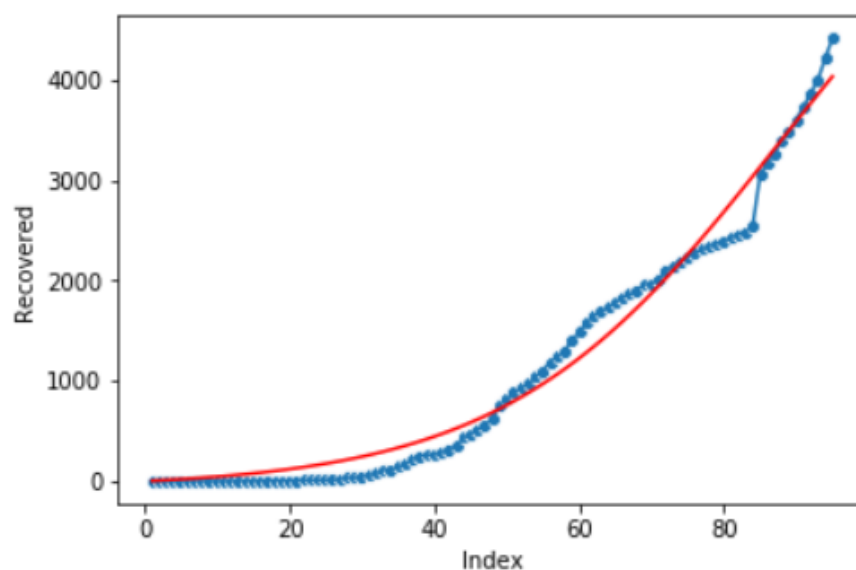Out[3]: [<matplotlib.lines.Line2D at 0xbe0bcc0>]
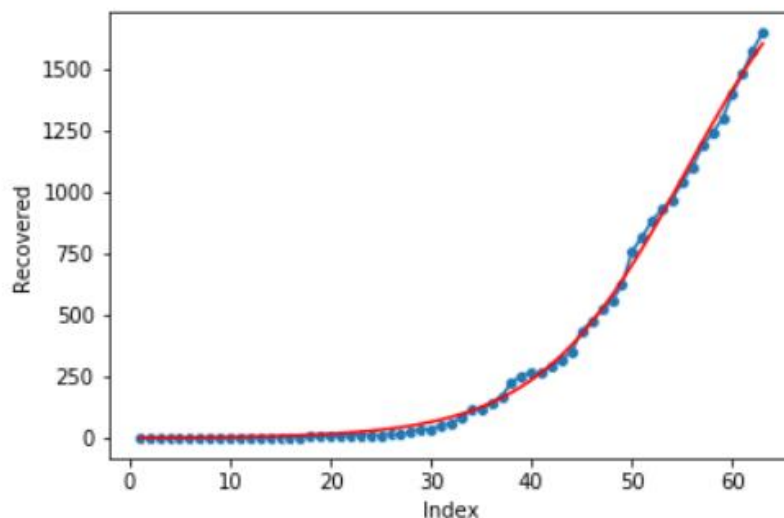
But, when we tried to predict the end of epidemic and maximum recovered, we get very different results far from current scenario!!

We found that the reasons behind this were

- *Migration of labourers from other states*
- *Relaxation of lockdown in AP(from 21 May)*

So, we collected more data (up to 21 June) and apply the same procedure as before & get this curve R(t):

```
Out[5]: [<matplotlib.lines.Line2D at 0xbe0fd68>]
```


This gives us more accurate results.

Our SIR model ends here.

Results:

The values of constants:

If the lockdown has been followed strictly:
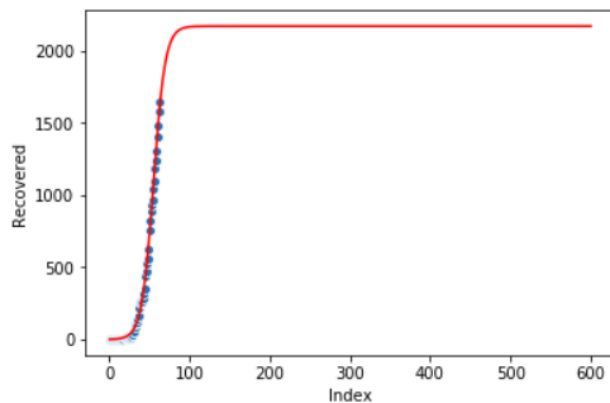
β=0.622884439* 10^-7

γ=2.95502463

Reproduction constant=1.0462

Graph and predictions:

```
maximum recovered : 2169
Day at which maximum recovered people observed: 119
(2169, 2169, 2169, 2169)
Pandemic ends after 120 days from start of pandemic if lockdown has been followed strictly.
```
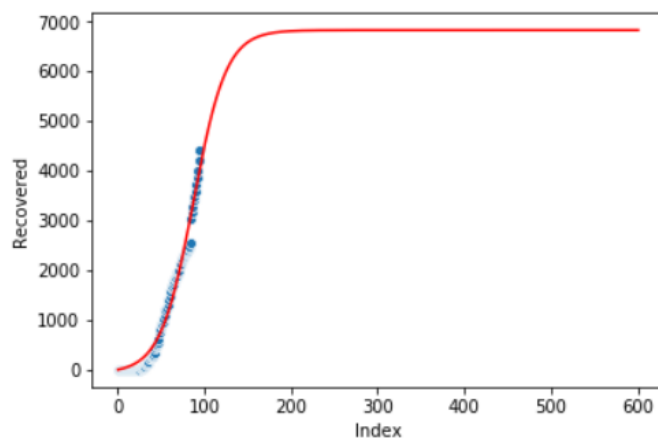

As the effect of migration and relaxation:

$\beta = 1.03311041 * 10^{-6}$

$\gamma = 48.2258649$

Reproduction constant=1.063

Graph and predictions:

```
maximum recovered : 6825
Day at which maximum recovered people observed: 263
(6825, 6825, 6825, 6825)
Due to relaxation of lockdown, Pandemic will end after 264 days from start of pandemic
```


Discussion:

- If lockdown has been followed strictly, pandemic would have been in the end phase. But, now it will take more days to end.

- The reason behind the unexpected variations in infected compartment is also the same.
- In spite of these, lockdown has reduced reproduction constant considerably(from 2.0 to 1.046) (expected value of R is from Nature article)
- Also, as a result of migration and relaxation of lockdown, more data should be available for perfect results which unfortunately was not.
- Another point is that this is very simplistic model and does not contain compartments like quarantine, exposed, etc. which are very important compartments in cases like COVID. Thus, SEIQR is the better option than this.
- Another limitation is that we have neglected migration in this model which affects highly if the data values are low. (Eg. Our case)

Citations:

- http://hmfw.ap.gov.in/covid_19_dailybulletins.aspx
- https://en.wikipedia.org/wiki/Template:COVID-19_pandemic_data/India/Andhra_Pradesh_medical_cases_chart
- https://www.covid19india.org/state/AP
- https://en.wikipedia.org/wiki/Andhra_Pradesh
- https://www.nature.com/articles/d41586-020-00758-2