EE 691

RESEARCH AND DEVELOPMENTAL PROJECT

ESTIMATION OF ASYMPTOMATIC POPULATION BASED ON STOCHASTIC EPIDEMIOLOGICAL MODELLING

Syomantak Chaudhuri
Roll Number 170070004

Guide : Prof. Vivek S. Borkar

Electrical Engineering Department
Indian Institute of Technology, Bombay

ABSTRACT

Based on the widely known 3 compartment SIR technique, we have presented a stochastic modelling technique with 4 compartments. In this work, we focus on demonstrating the effectiveness of the approach that we propose. In general, stochastic epidemiological modeling is much tougher to analyze and study than a deterministic differential equation-based technique. In our proposed parametric method, two-fold time-varying discretization of population volume and time is used to simplify the problem. We then use a variant of the Stochastic Approximation Expectation-Maximization algorithm to obtain the underlying model parameters.

1 INTRODUCTION

One of the first significant work in epidemiological modelling was by Kermack and McKendrick [1] in 1927. Conventionally, the model (SIR model) has 3 compartments - S compartment denoting the susceptible population, I compartment denoting the infected population, and R compartment containing the recovered population. Population in I compartment can infect susceptible population in S. Infected people also recover at some rate. The models can be both stochastic or deterministic. The deterministic variants often use differential equations to describe the dynamics and is often easier to handle. Let $S(t), I(t), R(t)$ denote the compartment population. The differential equation governing the SIR is as shown.

$$\begin{aligned}\dot{S}(t) &= -\beta S(t) I(t)/N \\ \dot{I}(t) &= \beta S(t) I(t)/N - \gamma I(t) \\ \dot{R}(t) &= \gamma I(t)\end{aligned}$$

β is known as the infection rate, γ is known as the recovery rate, and $N = S(t) + I(t) + R(t)$ is a the total population which is constant. Epidemiological modelling can help health agencies understand the dynamics of the epidemic which, in turn, would help them access the impact of various interventions and choose the optimal intervention for the region.

While their SIR technique ignores the variation in the demography, the varying intervention efforts with time such as complete lockdown and partial lockdown, and population changes, it has been extensively studied since the work by Kermack and McKendrick. In this work as well, we make the same assumptions and use a stochastic four compartment model similar to [2]. We assume that the population size is constant, ie, the number of deaths is not significant. We have also assumed that a recovered person would not get reinfected. We have used a stochastic variant of the Expectation Maximization algorithm [3] based on [4],[5] to estimate the parameters involved in our model.

In section 2, we describe out model and in section 3 we discuss the method we used to estimate the parameters. In section 4, the estimated parameters are used to predict the dynamics of the epidemic and the results are shown. We have listed some conclusions and limitations of our work in section 5.

2 THE MODEL

Our primary aim is to show the validity of our method of solving for the parameters and thus, estimating the asymptomatic population; we are not focused on the underlying compartmental model and its validity. To this end, we divide the population into 4 compartments:

- Compartment S (susceptible) denotes the population segment who are at risk of contracting the virus.
- Compartment I (symptomatic) denotes the population segment who have contracted the virus, currently symptomatic and recorded by official agencies.
- Compartment A (asymptomatic) denotes the population segment who have contracted the virus, currently asymptomatic along with those who are symptomatic but not recorded by the agencies.
- Compartment R (recovered) denotes the population segment who had contracted the virus but are presently not at risk of spreading.

R compartment contains dead people as well as the recovered people. Although slightly counter-intuitive, clubbing unaccounted symptomatic COVID-19 patients in A compartment makes sense as strictly, our goal is to estimate the unaccounted COVID-19 patients. Henceforth, 'infectious' term would refer to any COVID-19 patient (irrespective of compartment I or A).
We aim to predict the asymptomatic population dynamics given the data of number of symptomatic (I compartment) COVID-19 patients. This data is readily available on most COVID-19 tracking websites. Further, to make the job easier, we can also use the data of number of new symptomatic cases (per day) and number of recovered symptomatic cases (per day). Note that it is not possible to get any data on asymptomatic patients in most regions where the testing is not widely available and only reserved for the symptomatic patients.

Based on the proposed compartments, a graph depicting the connectivity of the compartments are as shown in figure 1.

In the modelling, we assume constant population, i.e., no birth or death (COVID-19 deaths do not matter). Let total population be N. Denote the population in compartment S at time t as $S(t)$ and other similarly. We model the differential equations governing the dynamics as described

$$\dot{S}(t) = -\lambda S(t)(\alpha I(t) + A(t))/N, \tag{1}$$

$$\dot{I}(t) = \lambda\mu S(t)(\alpha I(t) + A(t))/N - \gamma I(t), \tag{2}$$

$$\dot{A}(t) = \lambda(1-\mu) S(t)(\alpha I(t) + A(t))/N - \beta A(t), \tag{3}$$

$$\dot{R}(t) = \gamma I(t) + \beta A(t) \tag{4}$$

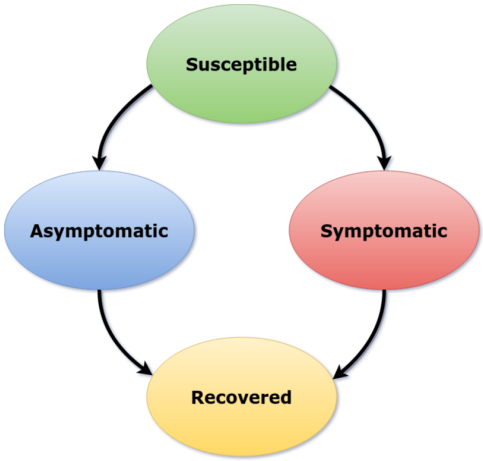

Figure 1: Graph depicting flow between the compartments

$\lambda, \alpha, \mu, \gamma, \beta$ are parameters that we wish to find. Based on the differential equations, one can infer the following about the parameters

- λ can be seen as the average number of contacts per person in a unit time multiplied by the probability of disease transmission when a susceptible person comes in contact with an infectious person.
- $\lambda S(t)(I(t)+A(t))/N$ would be the rate of decrease of $S(t)$. However, we add a parameter $\alpha \in [0,1]$ in this equation to reflect the fact that those who are symptomatic COVID-19 patients are much less likely to come in contact with a susceptible person (than random collision probability of $I(t)S(t)/N^2$) due to quarantine protocols implemented in various places.
- μ reflects the probability of an infected person to be symptomatic or asymptomatic.
- β and γ are the inverse of recovery time for asymptomatic and symptomatic COVID-19 patients respectively.

Although it has been speculated that a recovered person might get infected again as COVID-19 antibodies do not remain in the body for long, a recent study [6] suggests the opposite. If such a phenomenon was to be included to our model, then a simple transition between R and S compartment could be added.

3 METHODOLOGY

Since we are only focused on empirically showing the validity of our methods, we generated synthetic data based on some values of the parameters which we wish to solve. The synthetic data was generated based on the set of differential equations (1) to (4) and an appropriate initial condition.

We assume that $\beta = \gamma$ - this would also make the task of learning the parameters easier. We can view the differential equations (1) to (4) as an approximation of an underlying continuous time Poisson process, with the possible transitions shown in figure 1.
We discretize the system two fold - we discretize the possible transitions from each compartment and we discretize the time as well. Each day is divided into T intervals. Further, for any compartment $X \in \{S, I, A, R\}$, on the n-th day, for each interval of size T^{-1} day, there can only be one transition out of the state, and this transition is of size X_n/D, where X_n is the value of compartment X at the start of the day and D is a constant.
Since in this approximation we are not updating the transition size of X_n/D for all the T intervals in a given day, we have to add an additional constraint that $T = D$ to the possibility of the entire compartment population becoming 0 after a day and to restrict the possibility of compartment population becoming negative.
Let $\Delta_S^n = S_n/D$. Similarly define Δ_I^n, Δ_A^n. For a time interval $1/T$ on the n-th day, the transition probabilities P_S (of moving out) of the states can be found as based on the differential equations 1-4. The decrease in population of S compartment is in time $1/T$ is given as $dS = -\lambda S(\alpha I + A)/(NT)$. The S, I, A values in the expression is the value at the start of the time interval. Now, for our Poisson process, the expected decrement in S population would be $P_S^n \Delta_S^n$ (more on [7]). Therefore (shown for compartment S)-

$$\begin{aligned} P_S^n \Delta_S^n &:= \lambda S_n(\alpha I_n + A_n)/NT \\ P_S^n &= \lambda(\alpha I_n + A_n)/N \qquad \text{(Using D=T)} \end{aligned} \tag{5}$$

Similarly,

$$P_A^n := \gamma \tag{6}$$

$$P_I^n := \gamma \tag{7}$$

Note that there is an approximation here when we use the same value of S_n, I_n, A_n for all the T time intervals in each day. Ideally, we should be updating the values as soon as there is any transition but that approach does not seem to be feasible for simulation based algorithm that we propose.

At the start of the next day, we update the Δ^n values based on the new compartment populations due to influx (from other compartments) and outflux of people. People moving out of compartment S will move into I or A with probability μ or $1-\mu$ respectively. For a particular day n, let D_S^n be a binomial random variable $\sim Bin(T, P_S^n)$. Similarly, $D_{IR}^n \sim Bin(T, P_I)$ and $D_{AR}^n \sim Bin(T, P_A)$. Let $D_{SI}^n \sim Bin(D_S^n, \mu)$ (and $D_{SA}^n = D_S^n - D_{SI}^n$). The outflux from compartment X ($X \in \{S, I, A\}$) will be $\Delta_X^n D_X^n$ and we need to account the influxes accordingly. Concretely, the update of variables after a day would

be

$$S_{n+1} = S_n - \Delta_S D_S^n \tag{8}$$

$$I_{n+1} = I_n - \Delta_I D_{IR}^n + \Delta_S D_{SI}^n \tag{9}$$

$$A_{n+1} = A_n - \Delta_A D_{AR}^n + \Delta_S D_{SA}^n \tag{10}$$

$$R_{n+1} = R_n + \Delta_I D_{IR}^n + \Delta_A D_{AR}^n \tag{11}$$

3.1 Stochastic Approximation of Expectation Maximization Algorithm

To estimate the parameters given the data of S-I and I-R daily transitions, we use the stochastic approximation version of Expectation Maximization algorithm (SAEM) [4] [5]. This is done because the expectation computation is quite complex for such a long time series.

Let the given data be for N days and we are given the initial values of each compartment. Let us represent the given data as $X^N = (X_0, X_1, .., X_{N-1})$ and each X_i has 2 components - SI transitions ($X_{i,0}$) and IR transitions ($X_{i,1}$) between day i to $i+1$. Note that the initial condition is also given and is implicit. Thus, we would represent the first k transition data as X^k while we would use subscript to refer to k-th transition data X_k.

On day i, suppose we have some estimates of $\hat{S}_i$ and $\hat{A}_i$, then we can use these estimates (true value of I_i can be obtained using X^{i-1} and initial condition) to get corresponding chunk sizes Δ_S^i, Δ_I^i, and Δ_A^i. We describe the estimated S-A and A-R transitions on day i using D_{SA}^i and D_{AR}^i respectively in units of the chunk sizes. For example, for a particular day, using (6), we can estimate that the transition from A to R is the random variable $D_{AR}^i \Delta_A^i$ (using the corresponding values of $Delta_A^i$ for that particular day). Note that describing the exact dynamics of the compartment for each day is equivalent to giving the $D_{SI}^i, D_{SA}^i, D_{IR}^i, D_{AR}^i$ values and the initial condition. For SAEM, we want to use the population (time series) in compartments S and A as the hidden variable, therefore, we can instead represent this times series using the D values.

Let Z^N be the trajectory of the hidden variables, ie, $Z_n = \{D_{SA}^n, D_{AR}^n\}$ for $n = 0, .., N-1$. Let $\vec{\theta}$ be the vector of parameters $\lambda, \alpha, \mu, \gamma$. We can now describe the SAEM algorithm in terms of our variables -

Algorithm 1: SAEM

random initialize $\vec{\theta}^{(0)}$;
zero initialize $Q^{(0)}(\vec{\theta})$;
while *Not Converged* **do**
  $z^N \sim P(Z^N | X^N, \vec{\theta}^{(t)})$;
  $Q^{(t)}(\vec{\theta}) \leftarrow (1-\gamma_t) Q^{(t-1)}(\vec{\theta}) + \gamma_t \log(P(z^N, X^N | \vec{\theta}))$;
  $\vec{\theta}^{(t+1)} \leftarrow \vec{\theta}^{(t)} + \eta_t \nabla_{\vec{\theta}} Q^{(t)}(\vec{\theta})$;
end

Note that the distributions $P(Z^N | X^N, \vec{\theta})$ is quite complex due to the complicated dependence of D

random variables across the days. For example, suppose we are sampling z^N sequentially (day-wise) according to the distribution $\prod_{i=0}^{N-1} P(Z_i|Z^{i-1}, X^N, \vec{\theta}^{(t)})$ and at the n-th day we have some estimate of S_n based on the D values of the previous days then, $d_{SI}^n = TX_{n,0}/S_n$ and if $Z_{n,0} = D_{SA}^n = T - d_{SI}^n$ then it means that $S_{n+1} = S_n - \Delta_S^n(d_{SI}^n + D_{SA}^n) = 0$, and hence, any $X_{k,0},\ k > n$ can not be greater than 0 which might contradict the data available. Thus, we can't arbitrarily and independently sample the D values without taking into account the later X^N data.

Further, the SAEM algorithm described requires too much memory after a few iteration. We solved this by taking $\gamma_t = 1 \forall\, t$ and essentially only considering the log-likelihood of the current sample for the gradient ascent.

3.2 Distribution Approximation

In order to solve the problem, we make an approximation and sample independently the D_{SA} and D_{AR} independently of the future data. Note that while D_{SA} and D_{AR} are sampled independently of each day and each other, the evolution of S, A compartments are cumulative effects and are dependent on all the D values till the time in consideration. Based on this scheme, we describe the sampling distribution.

$$z^N \sim \tilde{P}(Z^N|X^N, \vec{\theta}^{(t)})$$

$$\begin{aligned} \tilde{P}(Z^N|X^N, \vec{\theta}^{(t)}) &= \prod_{i=0}^{N-1} P(Z_i|Z^{i-1}, X^N, \vec{\theta}^{(t)}) \\ &\approx \prod_{i=0}^{N-1} P(Z_i|Z^{i-1}, X^i, \vec{\theta}^{(t)}) \end{aligned} \tag{12}$$

$$\tag{13}$$

Inspecting each $P(Z_i|Z^{i-1}, X^i, \vec{\theta}^{(t)})$ term, since Z^{i-1}, X^i is given, we can calculate S_i, I_i, A_i based on it as shown in equations (8)-(11). We convert the $S-I$ transition data to $d_{SI}^i = TX_{i,0}/S_i$. We can also get the value of P_S^i similarly. Define

$$P_c^i := \frac{P_S^i(1-\mu^{(t)})}{1-\mu^{(t)}P_S^i} \tag{14}$$

Then for random variables $Z_{i,0} = D_{SA}^i$ and $Z_{i,1} = D_{AR}^i$, we can write the probabilities as (dropping the

sub/super-script i for the RHS terms in 15-16)

$$P(Z_i|Z^{i-1},X^i,\vec{\theta}^{(t)}) = P_{i1}(d_{SA})P_{i4}(d_{AR})$$

$$P_{i1}(d_{SA}) := \binom{T-d_{SI}}{d_{SA}} (P_c)^{d_{SA}} (1-P_c)^{T-d_{SI}-d_{SA}} \tag{15}$$

$$P_{i4}(d_{AR}) := \binom{T}{d_{AR}} (P_A)^{d_{AR}} (1-P_A)^{T-d_{AR}} \tag{16}$$

Similarly, we derive the log-likelihood expression (note that there is no approximation in this part).

$$\log P(Z^N,X^N|\vec{\theta}) = \sum_{i=0}^{N-1} \log(P(Z_i,X_i|Z^{i-1},X^{i-1},\vec{\theta})) \tag{17}$$

dropping sub/super-script in RHS of 18-21,

$$P(Z_i,X_i|Z^{i-1},X^{i-1},\vec{\theta}) = P_{i1}(d_{SA},d_{SI})P_{i2}(d_{SA},d_{SI})P_{i3}(d_{IR})P_{i4}(d_{AR})$$

$$P_{i1} := \binom{T}{d_{SA}+d_{SI}} (P_S)^{d_{SA}+d_{SI}} (1-P_S)^{T-d_{SI}-d_{SA}} \tag{18}$$

$$P_{i2} := \binom{d_{SI}+d_{SA}}{d_{SI}} (\mu)^{d_{SI}} (1-\mu)^{d_{SA}} \tag{19}$$

$$P_{i3} := \binom{T}{d_{IR}} P_I^{d_{IR}} (1-P_I)^{T-d_{IR}} \tag{20}$$

$$P_{i4} := \binom{T}{d_{AR}} (P_A)^{d_{AR}} (1-P_A)^{T-d_{AR}} \tag{21}$$

As T is large, we can approximate these by

$$\log(P_{i1}) \approx -T\,\mathrm{D}\left(\frac{d_{SA}+d_{SI}}{T}\middle\|P_S\right) \tag{22}$$

$$\log(P_{i2}) \approx -(d_{SA}+d_{SI})\,\mathrm{D}\left(\frac{d_{SI}}{d_{SA}+d_{SI}}\middle\|\mu\right) \tag{23}$$

$$\log(P_{i3}) \approx -T\,\mathrm{D}\left(\frac{d_{IR}}{T}\middle\|P_I\right) \tag{24}$$

$$\log(P_{i4}) \approx -T\,\mathrm{D}\left(\frac{d_{AR}}{T}\middle\|P_A\right) \tag{25}$$

where $\mathrm{D}(a\|b)$ represents the KL-divergence between the 2 Bernoulli distributions $Bern(a)$ and $Bern(b)$.

4 NUMERICAL EXPERIMENTS

For implementing the SAEM algorithm, we chose $\gamma_t = 1 \forall t$. We did not use the simple stochastic gradient ascent algorithm and instead opted to use Adam [8] present in PyTorch [9]. and $\eta_t = \frac{1}{1+t/P}$ where we set $P = 40$ heuristically (same for all the parameters). The learning rate was tuned for each parameter heuristically to $5e-3$. We used $T = D = 10,000$ for our simulation. Further, we clip the parameter values to $[0.01, 0.99]$ during the gradient ascent as a sanity check.

In the sampling step in SAEM, since our distribution is approximate and ignores the future X^N data, it is quite possible that the sampling process would end up with an error during the trajectory as explained in section 3.1. This is mainly observed when the parameter estimates are way off and cause too much outflux from S compartment. However, if one starts with parameters which cause relatively low outflux from S compartment, then the sampling step does not tend to cause any error throughout the process even as parameter values get updated. Another way to avoid such errors would be to set $\mu = 0.99$ initially. This would mean that most of the outflux from S compartment are to I compartment, and that is already given as X^N (which is obviously consistent) and hence, there is very low probability of getting an error with such initialization.

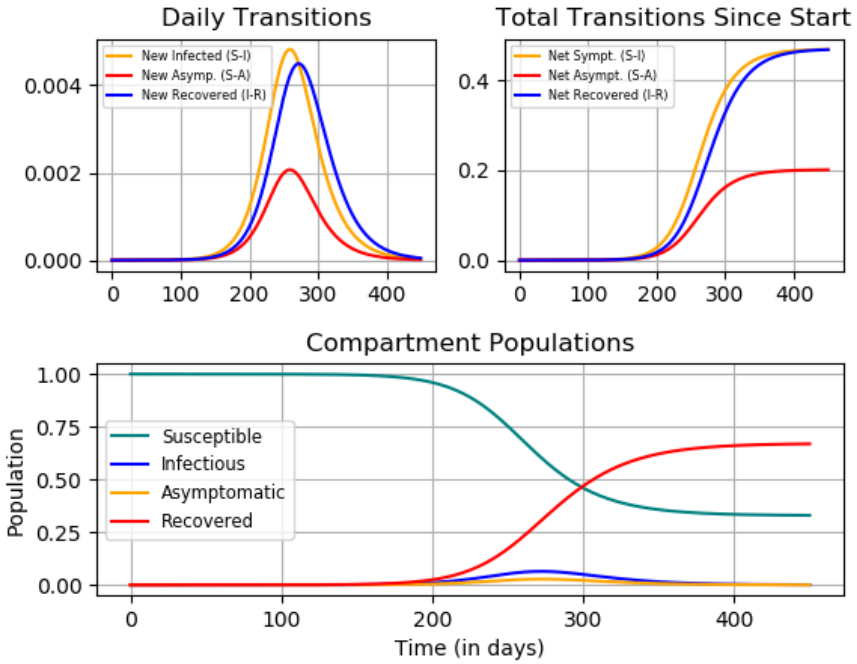

Figure 2: In clockwise order from top left - (1) Transitions between compartments per day during the epidemic. (2) Total transitions along the 'edge' since $t = 0$. (3) Compartment population on each day.

For data generation, we used the set of differential equations 1-4 to synthesize our own dataset. We used some suitable initial condition and some parameter values closely related to the COVID-19 epidemic [10]. One such dataset is shown in figure 2. Our method, however, could not be directly used

on this dataset. This can be understood as a trade-off between precision and computation complexity. Our choice of T, D is not large enough for the initial and final tail of the S curve as the value of P_S in the initial and final days are extremely small due to the small value of S. Thus, during sampling of D_{SA}, all the T samples are 0 as they are nearly $Bern(0)$. This could perhaps be solved by taking a large value of D so that at least a few samples are 1 but this is computationally infeasible. Instead, we chose to set the $t = 0$ for our algorithm at about day 230 and considered the data only till day 320. Using this partial data the result of the SAEM algorithm is shown in figure 3 and 4. In this experiment, we started with $\mu = 0.5$, ie, equal probability of a person being asymptomatic or symptomatic. We took $\lambda = 0.01$ to avoid the error explained at the beginning of the section. γ and α were initialised randomly in $(0, 1)$. As can be seen from figure 4, γ and μ converge to their true values but the other 2 parameters seem to converge to some other value.

Using the final parameters obtained from our modeling, we plotted the predicted SIAR dynamics (for time horizon of the experiment) and they are shown in figure 5. Note that we have used the differential equations 1-4 with the estimated parameters to generate the estimated trajectory. In terms of our stochastic model, this would be equivalent to the expected trajectory. We have reported the Root Mean Squared fractional Error (RMSfE) as a performance metric. If the true value of a compartment is X_t and the predicted value is $\hat{X}_t$, then the RMSfE is given by

$$RMSfE(\hat{X}^H) = \sqrt{\frac{1}{H}\sum_{t=1}^{H}\left(\frac{X_t - \hat{X}_t}{X_t}\right)^2}$$

Note that one could also replace the $\hat{I}^H$ by the given data at the cost of inconsistency in conservation of total population. This would yield RMSfE for I compartment as 0. Another performance metric we have reported is the normalized cross-correlation coefficient (NCC) between the estimated and the actual trajectory of each compartment in figure 6.

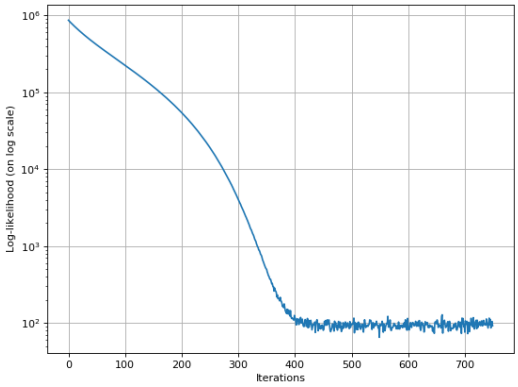

Figure 3: Negative of log-likelihood during the gradient ascent. The Y-axis is in log scale.

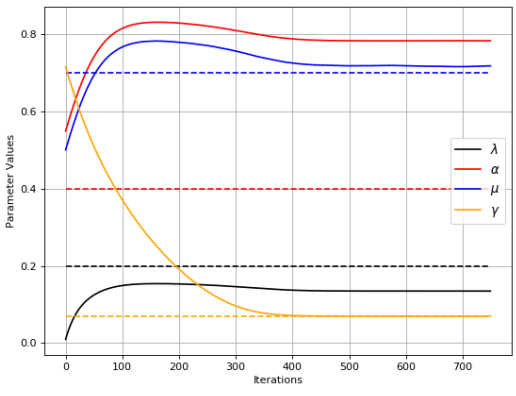

Figure 4: The change in parameter values during gradient ascent is shown. The dashed lines indicate the true parameter values used to generate the dataset.

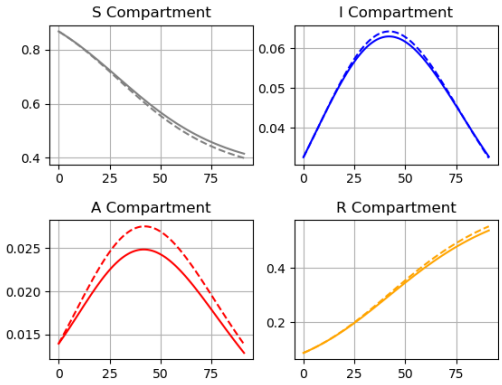

Figure 5: Trajectory of the epidemic estimated using our model.
The dashed lines represent the true trajectory.

	RMSfE (in %)	NCC (in %)
S	2.4	99.9
I	1.2	99.9
A	8.2	99.9
R	2.1	99.9

Figure 6: Results of the experiment

We conducted the same experiment with differential random initial conditions for γ and α. While γ was converging to the correct value in all the experiments, the results were not so consistent for the other parameters. Results with different initial parameter conditions are shown in figure 8 and 7. Thus, the initial condition affects the parameters obtained significantly. To investigate further, we initialised the parameters to their true values and used the SAEM algorithm to get the graph in figure 9. While λ, μ, γ change slightly which might be expected due to some discretization artefacts, α seems to be rapidly changing after a point. To inspect this further, we froze the other 3 parameters to their true value and again performed the optimization process and the result is shown in 10. As can be seen from the graph, α does not seem to be changing significantly. A possible explanation is discussed in section 5.

5 Further Discussions

Perhaps one of the most striking limitations of our modeling technique is the initial condition requirement. While this seems acceptable for early days in the epidemic as we can assign some estimates of the compartments, this seems quite outlandish when we restrict our dataset to regions where S and I compartments have significant population - when the epidemic has significant population in each compartment, it is quite unreasonable to require the exact data in each compartment as our model's initial condition. One of the possible solutions to this problem would be to have a variable D_n, T_n for the n-th day. When the value of I_n (given) is low, we can increase the value of D_n, T_n to ensure that at least some samples are non-zero, ie, $D_{SA}^n > 0$. However, this would be at the cost of increased computational complexity.

Further, our modeling has the same limitations as that of EM algorithm and it can not converge to the global minimum. The optimization process would get stuck in local optima. This is reflected by the rapid variation of the estimated parameters with the initialisation. This is observed in figures 4, 8,

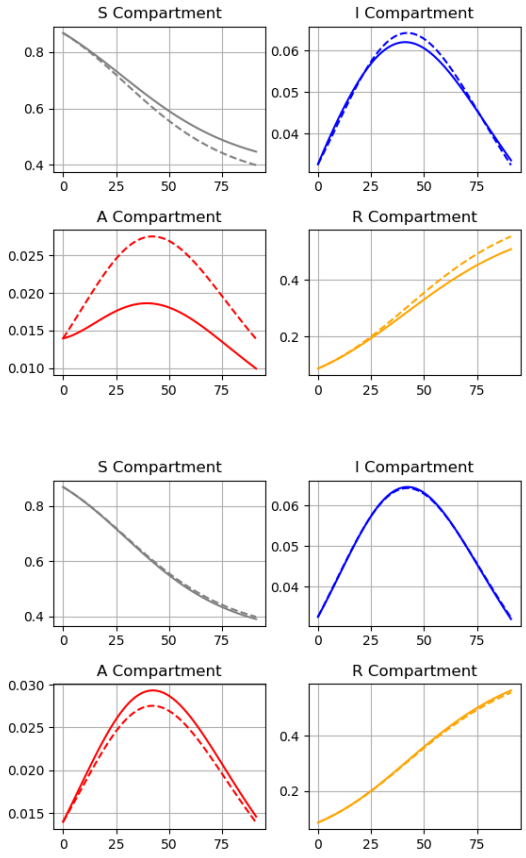

Figure 7: Comparison of the estimated and true trajectories.

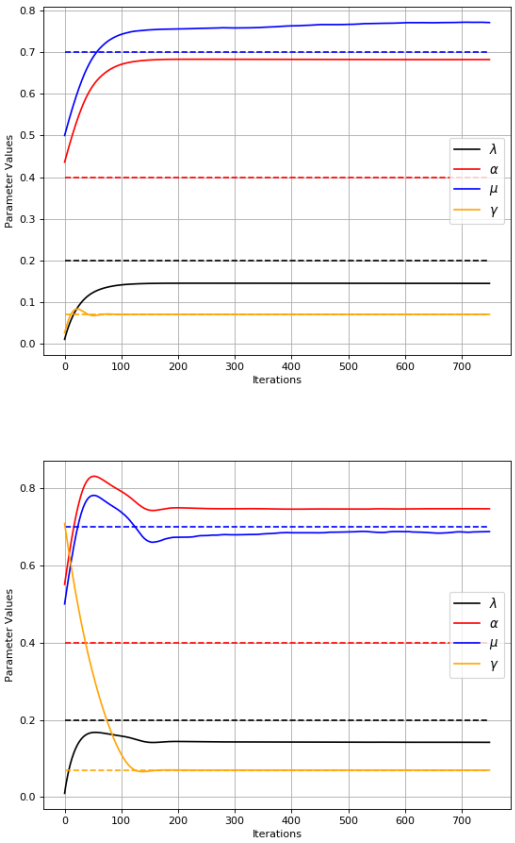

Figure 8: The change in parameter values during gradient ascent.

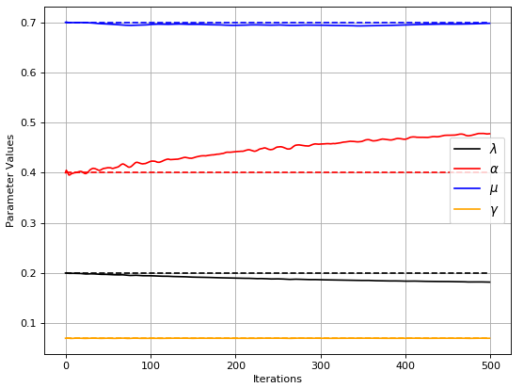

Figure 9: Initialised with true values, we plot the change in the parameters during gradient ascent.

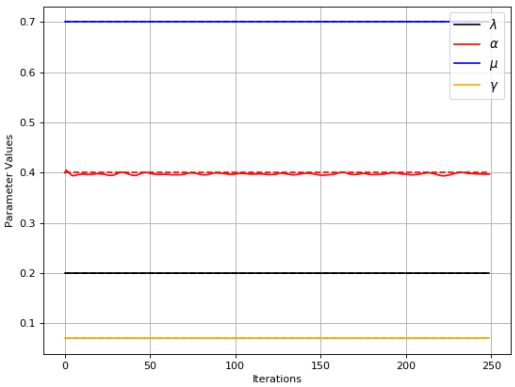

Figure 10: Initialised with true values and keeping λ, μ, γ fixed, α was optimized.

and 9.

In the experiment in figure 9, one would expect that parameters to remain roughly constant but a sharp change in α can be observed. However, when we fix the other parameters, then no such change in α can be observed in figure 10. One possible explanation to this can be that α is extremely sensitive to the values of the other parameters can perhaps, this effect is amplified due to discretization artefacts. It is clear from the results that γ is quite easy to learn and it can be understood in terms of the data given to us already. The $I-R$ transitions and the I population is enough to estimate γ and hence, it comes as no surprise that our model can easily learn the true value of γ.

It is quite interesting to note that in SAEM, we sampled from one distribution but calculated the log-likelihood for a different distribution! Yet, our estimated trajectory does not seem to be too far off the true trajectory. Perhaps, a better approximation can increase the performance of our method. Another possible improvement can be to use several datasets while estimating the parameters, assuming the datasets have the same underlying true parameters. We would independently sample trajectories corresponding to each dataset, calculate the log-likelihood for each trajectory and then perform gradient ascent with sum of the log-likelihoods as the cost function.

REFERENCES

[1] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.

[2] M. Agrawal, M. Kanitkar, and M. Vidyasagar, "Modelling the spread of sars-cov-2 pandemic-impact of lockdowns & interventions," *The Indian journal of medical research*.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[4] O. Cappé and E. Moulines, "On-line expectation–maximization algorithm for latent data models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 3, pp. 593–613, 2009.

[5] W. Jank, "Implementing and diagnosing the stochastic approximation em algorithm," *Journal of Computational and Graphical Statistics*, vol. 15, no. 4, pp. 803–829, 2006.

[6] J. M. Dan, J. Mateus, Y. Kato, K. M. Hastie, C. E. Faliti, S. I. Ramirez, A. Frazier, E. D. Yu, A. Grifoni, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette, and S. Crotty, "Immunological memory to sars-cov-2 assessed for greater than six months after infection," *bioRxiv*, 2020.

[7] P. E. Greenwood and L. F. Gordillo, *Stochastic Epidemic Modeling*, pp. 31–52. Dordrecht: Springer Netherlands, 2009.

[8] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations*, 12 2014.

[9] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 8024–8035, Curran Associates, Inc., 2019.

[10] D. K. Bagal, A. Rath, A. Barua, and D. Patnaik, "Estimating the parameters of susceptible-infected-recovered model of covid-19 cases in india during lockdown periods," *Chaos, Solitons Fractals*, vol. 140, p. 110154, 2020.