Disease modeling with social distancing: application to COVID–19

Alen Alexanderian*

Abstract

We present simple illustrations of disease modeling social distancing, as motivated by the present outbreak of COVID–19 and the social distancing measures implemented throughout the world. We focus on a simple SEIR model in the present study to facilitate discussion of incorporating and optimizing social distancing in models of spread of disease.

1 Introduction

We present some simple illustrations of incorporating and optimizing social distancing strategies in models of infectious disease. This is motivated by recent outbreak of COVID–19 and the social distancing measures implemented in various areas. The study here is qualitative in nature; it is not meant to provide exact predictions. We consider a simple SEIR model and focus only on disease transmission. The goal is to illustrate how social distancing measures impact spread of disease as described by standard models, and give insights on optimizing social distancing measures.

2 The model

We consider the following modified SEIR model for spread of COVID–19, as governed by the following system of ordinary differential equations (ODEs):

$$
\begin{aligned}
\frac{dS}{dt} &= -(1-u)\beta\, SI/N, \\
\frac{dE}{dt} &= (1-u)\beta\, SI/N - \sigma\, E, \\
\frac{dI}{dt} &= \sigma\, E - \gamma\, I, \\
\frac{dR}{dt} &= \gamma\, I.
\end{aligned}
\tag{2.1}
$$

Here S, E, I, and R denote the susceptible, exposed, infected, and recovered/removed populations, respectively. The number N is the total population, which is a conserved quantity in the present model. The system (2.1) is a standard SEIR model [1] with the exception of the factor $1-u$ in equations for S and E. The coefficient $u \in [0,1]$ quantifies the level of social distancing: if $u = 0$, there is no social distancing, and if $u = 1$, we have a full lockdown. The way the control parameter u is incorporated in the ODE system is standard; see e.g., [3]. To start, we assume a constant in time u. Later in Section 4, we consider a time-dependent u.

The values of the model parameters are highly uncertain and their estimation is subject to intense research activity. Here we use values adapted from the early estimates reported in [2]: $\beta = 1.94$, $\sigma = 0.25$ and $\gamma = 0.62$. Also, we assume a total population of $N = 10^4$ of which one individual is exposed to the disease at $t = 0$. Here the unit of time is in days. Below, we consider the quantity

$$y(t,u) = I(t,u)/N,$$

as the fraction of the population that is infected.

*North Carolina State University, Raleigh, NC, USA. E-mail: `alexanderian@ncsu.edu`
Last revised: April 9, 2020.

3 Optimizing social distancing

An important question is: how to pick an "optimal" u? A full lockdown is impractical as some people have to be out to maintain essential societal functions; e.g., essential government agencies, hospitals, pharmacies, grocery stores, etc. need to operate. More generally, one can associate a cost with social distancing, in terms of economic/societal impact. Therefore, we seek a u value that strikes a balance between cost of social distancing and the ultimate objective of "flattening the epidemic curve". The reason for the latter is to reduce the number of infections and more critically avoid going beyond the capacity of the healthcare system to treat the infected people. To this end, we formulate a optimization problem as follows:

$$\min_{u \in [0,1]} J(u) := \int_0^T y(t;u)^2 \, dt + cu^2, \tag{3.1}$$

where $y(t;u)$ is the fraction of population that is infected; this is obtained by solving the ODE system (2.1). The first term in the objective function $J(u)$ in (3.1) controls the "flatness" of the curve, $y(t) = y(t;u)$, and the second term penalizes large values for u, in view of the cost associated with social distancing. This is an optimization problem constrained by an ODE system. The specific form of the optimization problem follows standard formulations in optimal control.

Solving the optimization problem (3.1) with $c = 0.06$, we obtain the results in reported in Figure 1 (left). We also plot $J(u)$ as a function of u in Figure 1 (right).

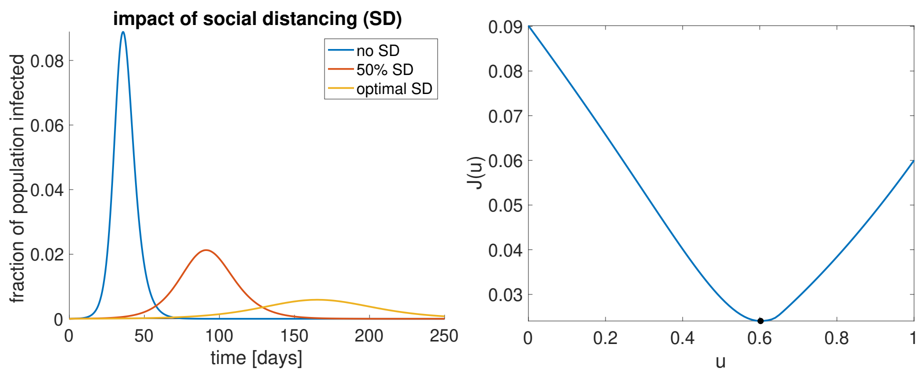

Figure 1: Left: fraction of population that is infected, with $u = 0$ (no social distancing), $u = 0.5$, and $u = 0.6031$ (optimal social distancing); the acronym SD in the legend stands for social distancing. Right: the objective function $J(u)$ with $u \in [0,1]$.

4 Optimizing social distancing with early termination

As mentioned above, social distancing involves economic and societal costs. Also, as one considers the dynamics of infections, a reasonable idea is to ask whether we can stop social distancing at later stages of the epidemic, possibly at the cost of slight increase in number of infections. If number of infections remain low, the healthcare system will be able to handle the infected in the hospitals, as there will be sufficient resources (ventilators, hospital beds, personnel ...). Of course, if social distancing is terminated prematurely, i.e., if we let our guard down too early, the epidemic curve will spike, leading to all kinds of bad outcomes, such as overwhelming of healthcare system and massive loss of life.

With the above in mind, instead of constant u in (2.1) we consider a sigmoid:

$$u(t) = \frac{u_0}{1 + \exp(t - t_s)}, \quad t \in [0, T].$$

In Figure 2, we show a few examples. With this approach, we seek to optimize the choice of u_0 and t_s simultaneously:

$$\min_{(u_0, t_s) \in \mathcal{Q}} J(u_0, t_s) := \int_0^T y(t, u_0, t_s)^2 \, dt + c_1 u_0^2 + c_2 (t_s / t_{\max})^2, \tag{4.1}$$

where c_1 and c_2 penalize large values of $u0$ and t_s due to the cost of maintaining large social distancing levels at long periods of time. Here $\mathcal{Q}$ indicates the set of admissible parameters; we chose $\mathcal{Q} = [0,1] \times$

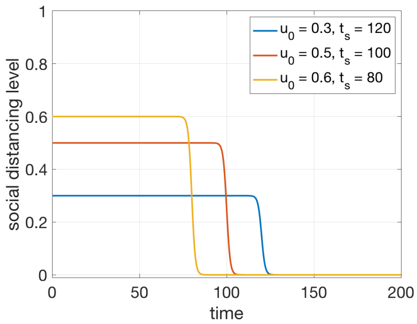

Figure 2: The social distancing function $u(t) = u(t; u_0, t_s)$ with a few choices of parameters. Here u_0 sets the magnitude of social distancing and t_s decides when social distancing is terminated.

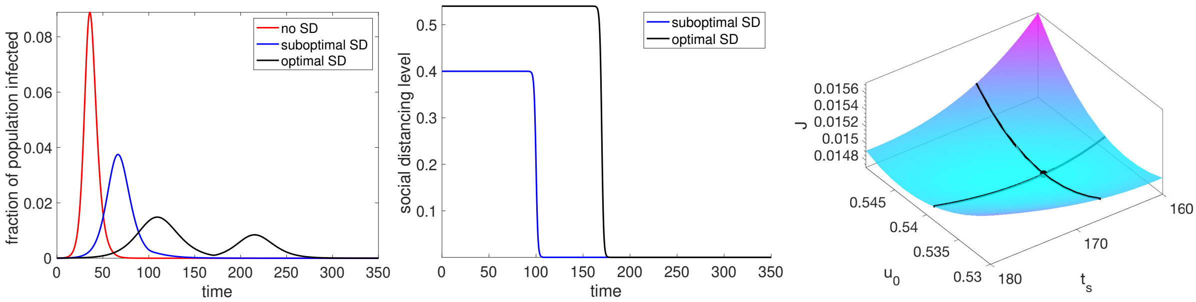

Figure 3: Spread of epidemics with no control, a suboptimal control and optimal control (left), the suboptimal and optimal controls (middle), and the surface plot (right) of the control objective J near the minimizer, which here is $(u_0, t_s) = (0.54, 169.88)$. We use a dot to indicate the location of the minimizer.

$[0, t_{\max}]$ with $t_{\max} = 200$. With $c_1 = 10^{-4}$ and $c_2 = 5 \times 10^{-3}$, we obtain the results in Figure 3. Note that in this case, the optimal social distancing function leads to infections happening in two controlled waves. This is a consequence of early termination of social distancing. Of course this behavior is also a consequence of the form of the model (2.1) and the choices of β, σ, and γ.

5 Sensitivity of the optimal control with respect to model parameters

Here we consider the following question: how sensitive is the computed optimal control (i.e., the social distancing level) u to the parameters β, σ, and γ? Why is this important? The reason is that there are large uncertainties associated with model parameters; these have to be estimated using incomplete observed data and using imperfect models. Therefore, it is important to know how sensitive the solution of the optimization problem, for finding optimal u, is to the (approximate) model parameters. This can guide experimenters in focusing on accurate estimation of the parameters that are most influential.

We focus on the case of constant in time u, for simplicity. Let us denote the parameter vector by $\theta = (\beta, \sigma, \gamma)^\top$. We want to compute

$$D_i = \frac{\partial u^*}{\partial \theta_i}\Big|_{\theta=\bar{\theta}}, \quad i = 1, 2, 3, \tag{5.1}$$

where u^* is the solution of the optimization problem (3.1) and $\bar{\theta}$ is the vector of nominal parameters used in the present study. Computing D_i is facilitated by differentiating through the optimality condition

$$J'(u^*(\theta), \theta) = 0. \tag{5.2}$$

Here J' denotes derivative of J with respect to u. We have used the Implicit Function Theorem (assuming sufficient regularity) to write the solution of the optimization problem as a function of θ in a neighborhood of $(\bar{u}^*, \bar{\theta})$. Here $\bar{u}^*$ is the optimal control corresponding to $\bar{\theta}$.

Differentiating with respect to θ_i, through (5.2), at the point $(u, \theta) = (\bar{u}^*, \bar{\theta})$, we have

$$\Big(\frac{\partial^2 J}{\partial u^2}\Big)\Big(\frac{\partial u^*}{\partial \theta_i}\Big|_{\theta=\theta^*}\Big) + \frac{\partial^2 J}{\partial u \partial \theta_i} = 0, \quad i = 1, 2, 3, \tag{5.3}$$

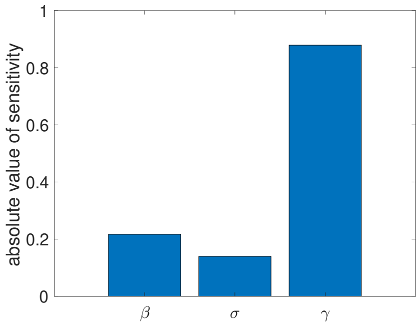

Figure 4: Shown are the absolute values of sensitivities, $|D_i|$ $i = 1, 2, 3$ with D_i as in (5.1).

where $\frac{\partial^2 J}{\partial u^2}$ and $\frac{\partial^2 J}{\partial u \partial \theta_i}$ are evaluated at $(u, \theta) = (\bar{u}^*, \bar{\theta})$; we approximate these using finite-differences:

$$\frac{\partial^2 J}{\partial u^2}(\bar{u}^*, \bar{\theta}) \approx \frac{J(\bar{u}^* - h; \bar{\theta}) - 2J(\bar{u}^*; \bar{\theta}) + J(\bar{u}^* + h; \bar{\theta})}{h^2}, \tag{5.4}$$

$$\frac{\partial^2 J}{\partial u \partial \theta_i}(\bar{u}^*, \bar{\theta}) \approx \frac{J(\bar{u}^* + h; \bar{\theta} + he_i) - J(\bar{u}^* + h; \bar{\theta} - he_i) - J(\bar{u}^* - h; \bar{\theta} + he_i) + J(\bar{u}^* - h; \bar{\theta} - he_i)}{4h^2}. \tag{5.5}$$

Using these, we solve (5.3) for $\frac{\partial u^*}{\partial \theta_i}$. Note that here e_i, $i = 1, 2, 3$, are the standard basis vectors in $\mathbb{R}^3$. We report the magnitude of the computed sensitivities in Figure 4. The results indicate that the computed optimal control u (social distancing level) is most sensitive to γ, but the other parameters also have non-negligible impact on the solution of the optimization problem. It is interesting that the optimal control is least sensitive to σ, which is the inverse of the average incubation period. Note, however, that the present analysis is local in the parameter domain.

6 Outlook

Interesting and important lines of inquiry include:

- **Modeling**: How can one model/estimate the cost of social distancing? How can we improve the model (2.1) for cases like COVID–19? Standard ideas include incorporating disease related mortality rate and impacts of migration in the model; also, the age of the patients being an important factor in COVID–19 dynamics, an important improvement would be to consider an age-structured SEIR model, which would involve a system of hyperbolic partial differential equations. Of course, the predictive capability of the model is highly dependent on accurate estimation of model parameters such as β, σ, and γ. The more complex the model, the more parameters need to be estimated.
- **Optimization**: How can the formulation of the optimal control problem be improved, while remaining practical? This of course is closely tied to the mathematical model of the disease transmission. In the present study, the constant social distancing model is useful as it provides important insight in simple terms. Given a sufficiently realistic model, one can also seek time-dependent controls (social distancing functions). The basic case of a sigmoid function was just an illustrative example, and more complex cases can be considered.

  Alternatively, one can consider computing constant optimal controls on shorter time intervals, $[T_0^i, T_f^i]$, $i = 1, \ldots, K$, iteratively (we assume $T_0^1 = 0$). A brief outline of such a strategy is as follows:

  1. Let the initial state $(S(0), E(0), I(0), R(0))$ be as described in Section 2
  2. For $i = 1, 2, \ldots, K$
  3. Solve the optimization problem

  $$\min_{u \in [0,1]} J(u) := \int_{T_0^i}^{T_f^i} y(t; u)^2 \, dt + cu^2, \quad (y \text{ obtained by solving (2.1) in } [T_0^i, T_f^i])$$

  4. Simulate (2.1) in $[T_0^i, T_f^i]$ with the computed optimal control
  5. If $i < K$, let $(S(T_0^{i+1}), E(T_0^{i+1}), I(T_0^{i+1}), R(T_0^{i+1})) = (S(T_f^i), E(T_f^i), I(T_f^i), R(T_f^i))$
  6. End For

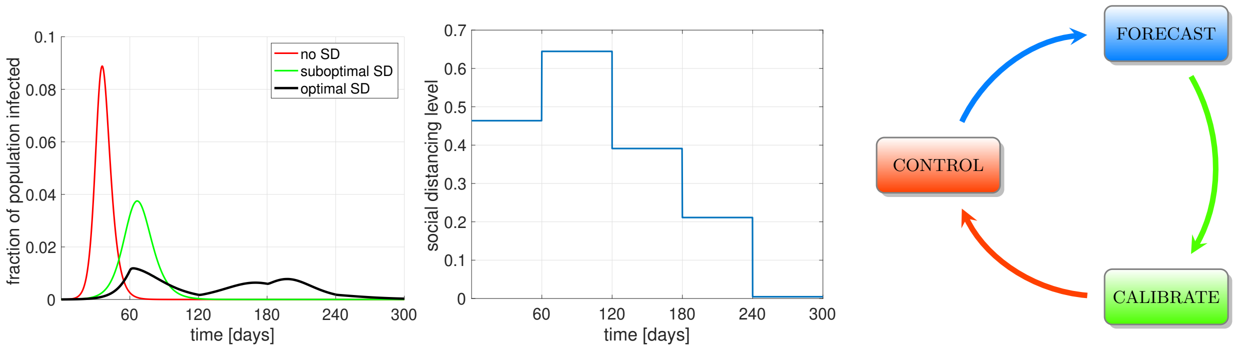

Figure 5: Left: fraction of the population infected with no social distancing (SD), suboptimal SD, and iteratively computed SD levels. Middle: the iteratively computed optimal SD levels; we computed optimal SD every sixty days. Note that this allows for adjusting the SD levels as the epidemic progresses, with more aggressive SD at the peak of the epidemic and with gradual decreases in SD over time. Right: diagram depicting iterative process of model calibration, control, and forecast.

We depict in Figure 5 (left,middle) an illustration of this strategy, with optimal social distancing levels computed every sixty days. In practice this approach allows for making short-term plans, monitoring the situation, and making decisions adaptively as more data becomes available. The results in Figure 5 are computed with a fixed model as described in Section 2. However, one can continuously calibrate the disease model using observed data, thus making successively computed optimal controls more effective. In short, we can follow an iterative process of calibration, control, and forecast untill the outbreak wanes. We depic this process in Figure 5 (right).

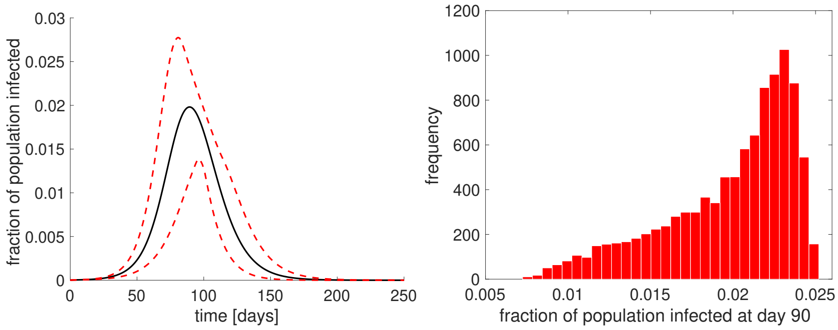

Figure 6: Simulated forecast under uncertainty, assuming the entries of $\theta = (\beta, \sigma, \gamma)^\top$ are distributed uniformly $\theta_i \sim U(0.95\bar{\theta}_i, 1.05\bar{\theta}_i)$. Left: mean of the infected population (solid line) along with 5th and 95th percentiles (dashed lines) with constant social distancing $u = 0.5$. Right: statistical distribution of the fraction of the population infected at $t = 90$ [days].

- **Uncertainty quantification**: parameter estimation provides estimates of model parameters, but also can provide parameter ranges and statistical distributions. Once such information becomes available, incorporating the statistical distribution of model parameters is important in providing forecasts with quantified uncertainties. A simple illustration is provided in Figure 6. In practice parameter distributions based on parameter estimation (using observed data) should be used.

References

[1] Herbert W Hethcote. The basic epidemiology models: models, expressions for r0, parameter estimation, and applications. In *Mathematical understanding of infectious disease dynamics*, pages 1–61. World Scientific, 2009.

[2] 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.

[3] Mami T Wentworth, Ralph C Smith, and Harvey Thomas Banks. Parameter selection and verification techniques based on global sensitivity analysis illustrated for an hiv model. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):266–297, 2016.