Lecture 6: SI Models

Radu Balan

Department of Mathematics, NWC
University of Maryland, College Park, MD

Version: February 15, 2023

Epidemiological Models

In this lecture we start discussing epidemiological models. There are two main types of epidemic models:

- deterministic (or, compartmental) model
- stochastic (e.g., agent based) model

We focus on three deterministic models:

1. SI (Susceptible-Infected) Model
2. SIR (Susceptible-Infected-Removed) Model
3. SEIR (Susceptile-Exposed-Infected-Removed) Model

Today we discuss the SI model.

The SI Model: The linear form

Assume a system with two compartments: 'Susceptible' (S) and 'Infected' (I). At time $t_0 = 0$ the system has a total of N individuals (initial total population). Most of them are susceptible $S(0) = S_0$, but some are infected, $I(0) = I_0$. Our intention is to model the time evolution of these populations.

The simplest model assumes a constant rate of exchange:

$$\left\{ \begin{array}{rcl} \frac{dS}{dt} & = & -\beta S \quad , \quad S(0) = S_0 \\ \frac{dI}{dt} & = & \beta S \quad , \quad I(0) = I_0 \end{array} \right.$$

where $\beta \geq 0$ is a parameter.

Analytic solution (exact):

$$S(t) = S(0)e^{-\beta t} \quad , \quad I(t) = N - S(0)e^{-\beta t}$$

with $N = S(0) + I(0)$, the total population. Note $S(t) + I(t) = N$ for all t.

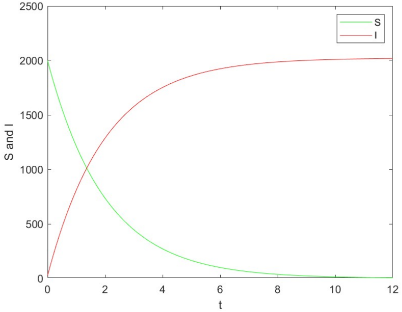
The SI Model: The linear form (2)

Deterministic simulation

Simulation of the linear SI model:

$$\beta = 0.5\ ,\ \ S(0) = 2000\ ,\ \ I(0) = 23$$

Note: at time $\tau = \frac{1}{\beta}$, the susceptible population dropped to $e^{-1}S(0) = 736$ which is about 37% of initial population.



The Euler scheme with step size of about 0.01 produces an error less than 0.2 at time $\tau = 2 = 1/\beta$.

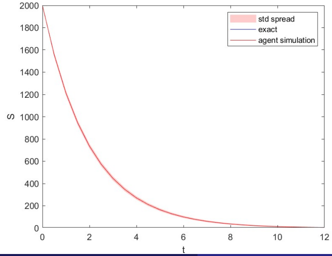
The SI Model: The linear form (3)

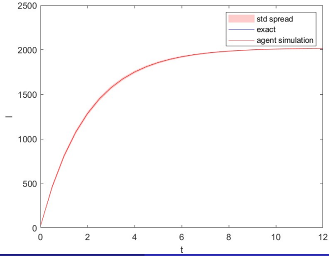
Agent based simulation

As we learned from previous lecture, an agent based simulation implements a Markov chain with the following transition matrix:

$$\Pi = exp(T_0 R) = \begin{bmatrix} e^{-\beta T_0} & 0 \\ 1 - e^{-\beta T_0} & 1 \end{bmatrix}$$

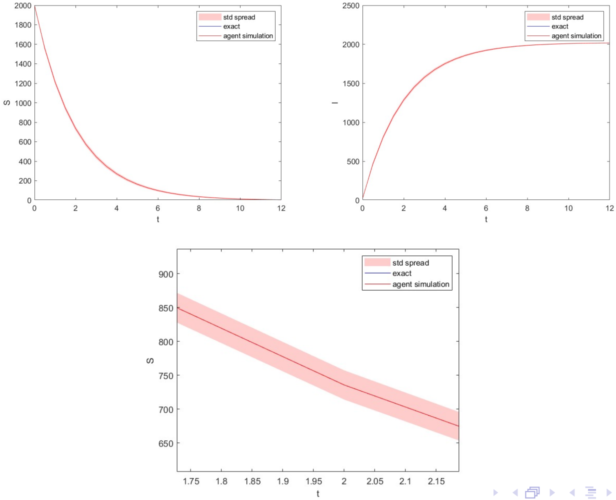
Here are results of a large number of simulations (10^6) for $T_0 = 0.5$. The shadded area is one std of simulations





The SI Model: The linear form (4)

Details of stochastic fluctuations



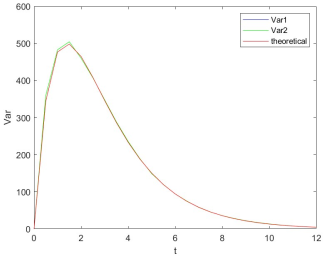
The SI Model: The linear form (5)

Additional stochastic analysis

For agent based simulation oe can compute the first two statistics of simulated S and I:

$$\mathbb{E}[S(t=pT_0)] = e^{-\beta t}S(0) \quad , \quad \mathbb{E}[I(t=pT_0)] = N - e^{-\beta t}S(0)$$

$$Var(S(t=pT_0)) = Var(I(t=pT_0)) = e^{-\beta t}(1-e^{-\beta t})S(0)$$



The SI Model: The Nonlinear form

The nonlinear form

A better modelling of the transition rate should take into account the transmission process of the disease. For infectious diseases (such as Covid-19), transmission occurs only when susceptible people get in contact with infected individuals. Such a process is modeled by a rate proportional to the product SI and not just S. For comparison reasons, we prefer to used a rate of the form $\beta S \frac{I}{N}$ which models the interaction between a susceptible individual with the fraction of infected population. Thus we obtain:

$$\left\{ \begin{array}{rcl} \frac{dS}{dt} & = & -\beta S \frac{I}{N} \quad , \quad S(0) = S_0 \\ \frac{dI}{dt} & = & \beta S \frac{I}{N} \quad , \quad I(0) = I_0 \end{array} \right. \qquad \text{(SI Model)}$$

where $N = S(0) + I(0)$ and $\beta \geq 0$ is transmission parameter.

The SI Model: The Nonlinear form

The normalized form

For reasons of normalizations, we prefer to compute *fractions* of susceptible population and of infected population:

$$s(t) = \frac{S(t)}{N} \quad , \quad i(t) = \frac{I(t)}{N}$$

In which case the model becomes:

$$\left\{ \begin{array}{rcl} \frac{ds}{dt} & = & -\beta si \quad , \quad s(0) = \frac{S_0}{N} \\ \frac{di}{dt} & = & \beta si \quad , \quad i(0) = \frac{I_0}{N} \end{array} \right. \qquad \text{(SI Model)}$$

Note $s(t) + i(t) = 1$ for all t (conservation of total population). The good news: it admits an exact closed form solution: $\frac{ds}{s(1-s)} = -\beta dt$. Hence:

$$s(t) = \frac{s_0}{s_0 + (1 - s_0)e^{\beta t}} \quad , \quad i(t) = \frac{i_0}{i_0 + (1 - i_0)e^{-\beta t}}.$$

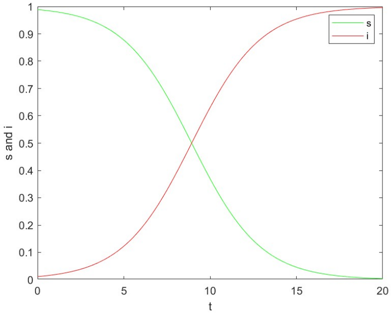
The SI Model: The Nonlinear form

The normalized form: Numerical simulation

Simulation of the non-linear SI model:

$$\beta = 0.5 \ , \ S(0) = 2000 \ , \ I(0) = 23$$

Note: the state converges to the same point $(0,1)$, that is, $\lim_{t\to\infty} s(t) = 0$ and $\lim_{t\to\infty} i(t) = 1$, but at a much slower rate than the linear model. There is also a delay in the onset of infections.



The SI Model

Exact solution vs. Agent Based Modeling

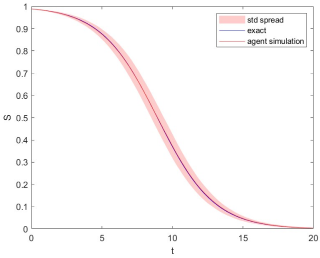
One way of implementing an agent based model is to pretend the nonlinear model is a quasilinear SI model with effective parameter $\tilde{\beta}(t) = \beta i(t)$. Thus we obtain a time-dependent Markov chain. For a discretization step T_0, at time step $p > 0$, the transition matrix to transition from time $(p-1)T_0$ to pT_0 is given by

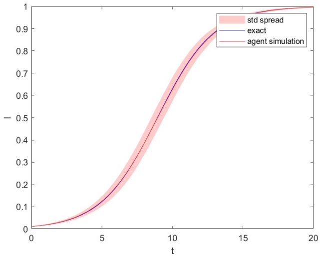
$$\Pi^{(p)} = \begin{bmatrix} e^{-\beta T_0 i((p-1)T_0)} & 0 \\ 1 - e^{-\beta T_0 i((p-1)T_0)} & 1 \end{bmatrix} \quad , p = 1, 2, 3, \ldots$$

The SI Model

Exact solution vs. Agent Based Modeling (2)

Here are results of a large number of simulations (10^3) for $T_0 = 0.01$. The shadded area is one std of simulations





How to Calibrate SI Models

Since $I(t)$, or $i(t)$, is a monotone increasing sequence, the SI model is appropriate only for the time series of the *cumulative* number of infections, not for daily rates of infection.

Give the time series $\{I(0), I(1), \cdots, I(T_{max})\}$ the question is to estimate β and N that best fit the data. The closed form solution for $I(t)$ is:

$$I(t) = \frac{NI(0)}{I(0) + (N - I(0))e^{-\beta t}} \quad , \quad t = 0, 1, 2, \cdots$$

Ideally, we want to minimize:

$$\begin{array}{l} \text{minimize} \;\; J(N,\beta) = \displaystyle\sum_{t=0}^{T_{max}} \left| I(t) - \frac{NI(0)}{I(0) + (N - I(0))e^{-\beta t}} \right|^2 . \\ \quad N \in \mathbb{N} \\ \quad \beta \geq 0 \end{array}$$

How to Calibrate SI Models (2)

Instead of minimizing the residuals, we adopt a different strategy: we transfor algebraically the prediction until we obtain a linear dependency on some parameter. Here is one possible approach:

$$e^{\beta t} = \frac{\frac{I(t)}{N-I(t)}}{\frac{I(0)}{N-I(0)}}$$

$$\beta t = log\left(\frac{I(t)}{N-I(t)}\right) - log\left(\frac{I(0)}{N-I(0)}\right) \quad (*)$$

From where we define the objective function:

$$J(\beta, N) = \sum_{t=0}^{T_{max}} \left|\beta t - log\left(\frac{I(t)}{N-I(t)}\right) + log\left(\frac{I(0)}{N-I(0)}\right)\right|^2 \quad (**)$$

The model (*) is linear in β but nonlinear in N.

How to Calibrate SI Models (3)

The first algorithm assumes we are know (given) the total number N. How to get it? One way is to look for $I(T_{max})$, the total number of infections at time T_{max}. If the end time is large enough, then, according to the SI model, the total number of infections should match the entire population. Once N is estimated ("guessed"), then β is given by:

Algorithm (SI Alg 1 - Known N)

Inputs: The time series of cumulative number of infections $\{I(0), I(1), \cdots, I(T_{max})\}$, and an estimate of the total population N.

Step 1: Compute the least-squares solution of (**) via:

$$\hat{\beta} = \frac{6}{T_{max}(T_{max}+1)(2T_{max}+1)} \sum_{t=1}^{T_{max}} t \cdot log\left(\frac{I(t)}{I(0)} \frac{(N - I(0))}{(N - I(t))}\right).$$

Output: Estimated $\beta = \hat{\beta}$.

How to Calibrate SI Models (4)

If N is to be estimated as well, then compute the objective function $J(\hat{\beta}, N)$ and optimize over N. For instance start with $N = I(T_{max}) + 1$ and increment:

Algorithm (SI Alg 2 - Unknown N)

Inputs: *The time series of cumulative number of infections* $\{I(0), I(1), \cdots, I(T_{max})\}$.

Step 1: Initialize $N = 1 + I(T_{max})$, $J_{old} = \infty$. *Set* $a = \frac{6}{T_{max}(T_{max}+1)(2T_{max}+1)}$.

Step 2: Repeat:

2.1 Compute:

$$J(N) = \sum_{t=1}^{T_{max}} \left| log\left(\frac{I(t)(N - I(0))}{I(0)(N - I(t))}\right)\right|^2 - a\left(\sum_{t=1}^{T_{max}} t \cdot log\left(\frac{I(t)(N - I(0))}{I(0)(N - I(t))}\right)\right)^2.$$

2.2 If $J(N) < J_{old}$ *then: (i) Assign*$J_{old} = J(N)$, *(ii) increment* $N = N + 1$, *and (iii) go to Step 2.1. Else go to Step 3.*

Step 3. For the last value of N, *compute*

$$\hat{\beta} = \frac{6}{T_{max}(T_{max} + 1)(2T_{max} + 1)} \sum_{t=1}^{T_{max}} t \cdot log\left(\frac{I(t)(N - I(0))}{I(0)(N - I(t))}\right).$$

Output: Estimated $\beta = \hat{\beta}$ *and* N.