Estimating time-varying transmission rates of the SIR model

Sang Woo Park

A thesis presented for Math 4P06

Department of Mathematics & Statistics
McMaster University
March 22, 2019

Contents

1 Introduction

Over the last few decades, a plethora of computational tools has been developed to study epidemic time series. Central to these studies are mechanistic models. Mechanistic modeling of a disease allow us to study the qualitative dynamics of a system and infer underlying mechanisms that drove the observed patterns in the data. For example, we now know that the complex dynamics of measles outbreaks are driven by seasonal transmission patterns (Earn et al., 2000; Dalziel et al., 2016).

Major challenges in mechanistic modeling of disease dynamics come from developing the actual model and parameterizing the model. One must make different sets of assumptions (e.g., discretization of a continuous-time system) depending on the system as well as their statistical model. Fortunately, the Susceptible-Infected-Recovered (SIR) model, one of the most basic epidemic models, is able to describe the measles dynamics sufficiently well. Focusing on a simple, well-understood model allows us to compare statistical methods in their ability to infer underlying parameters without making significant changes to the model. Here, we focus on the time-varying transmission rate of measles.

One way to estimate the transmission rate $\beta(t)$ is by matching the deterministic solution of the system (1) with the observed data by minimizing their "distance", often measured by sum of squares difference or negative log-likelihood (Riley et al., 2003; Chowell et al., 2004). This method of fitting a deterministic model, often referred to as trajectory matching, assumes that *all* variation in data can be explained by observation error alone (Bolker, 2008). Similarly, it is possible to match gradients of the system (1), where the gradients of the observed time series can be estimated by approximating the time series with a smooth curve, such as piecewise polynomials, and taking its derivative (Ellner et al., 2002). The main advantage of these methods is that they are computationally efficient and easy to implement.

On the other hand, Monte Carlo methods allow us to model *all* sources of variation (e.g., observation error, process error, and environmental stochasticity) explicitly. In particular, Sequential Monte Carlo methods (i.e., particle filtering), popularized by King et al. (2015), provides a powerful technique for testing hypotheses and inferring underlying parameters due to its ability to approximate the likelihood through a series of simulations. Approximate likelihood obtained by particle filtering can be maximized by iterated filtering (Ionides et al., 2011, 2015) or can be incorporated into a Bayesian scheme using a particle Monte Carlo Markov Chain (MCMC). [**SWP:** *Something about Bayesian Monte Carlo methods*]

This is not at all an exhaustive list of all existing methods. Generalized profiling balances bewteen trajectory matching and gradient matching by fitting a deterministic model (Hooker et al., 2010).

Laplace approximation based methods are also available.

Of all these methods, the time series SIR (TSIR) method serves a special place.

has been particularly successful in studying dynamics of ??? diseases, espe-

cially measles.

There are many ways to estimate underlying parameters of the SIR model (1); yet their differences are unclear because they all rely on different set of assumptions. Which one should one use? Here, we compare TSIR each method against simulated data and study how differences in the assumptions affect our conclusions. Then, we apply each method to measles time series from Boston and discuss implications.

2 Methods

2.1 SIR model

The Susceptible-Infected-Recovered (SIR) model is one of the most basic models in mathematical epidemiology. It describe how disease spreads in a homogeneously mixing population. Here, we present the model as a set of ordinary differential equations, but gillespie algorithm can be used to introduce demographic stochasticity (Gillespie, 1976):

$$
\begin{aligned}
\frac{dS}{dt} &= b(t) - \left(\beta(t)\frac{I}{N} + \mu \right) S \\
\frac{dI}{dt} &= \beta(t)S\frac{I}{N} - (\gamma + \mu)I \\
\frac{dR}{dt} &= \gamma I - \mu R
\end{aligned}
\tag{1}
$$

where S is the number of susceptible individuals, I is the number of infected individuals, and R is the number of recovered (or removed) individuals. We assume that individuals enter the susceptible pool at rate $b(t)$. Susceptible individuals become infected at a rate $\beta(t)$ by contacting infected individuals. Infected individuals recover at a rate γ; in other words, infectious period is exponentially distributed with mean $1/\gamma$ time units. Finally, all individuals experience natural mortality at a rate μ. A stochastic simulation of the SIR model is performed using the gillespie algorithm (Gillespie, 1976).

2.2 Epidemic time series

In order to estimate the parameters of the SIR model, it is important to know what kinds of data are available. Epidemic time series can be classified into three categories: incidence, mortality, and prevalence. Incidence is defined as the number of newly infected individuals generated over a reporting period. Then, true incidence of the SIR model (1) between time t and $t - t_{\mathrm{rep}}$, where t_{rep} is the length of reporting time step, can be obtained by integrating total infection rate over the reporting period:

$$
i_t = \int_{t-t_{\mathrm{rep}}}^{t} \beta(s)S\frac{I}{N}ds.
\tag{2}
$$

3

Similarly, true mortality case, which is defined as the number of individuals that died during a reporting period, can be obtained by integrating total death rate over the reporting period:

$$m_t = \int_{t-t_{\text{rep}}}^{t} \gamma I ds. \tag{3}$$

Finally, prevalence is defined as the number of infected individuals that are present in the population and corresponds to the state variable I; to distinguish the continuous dynamics from the discrete reports, we write p_t to denote true prevalence at time t. Incidence, mortality, and prevalence for a stochastic simulation is defined in the same manner as described above.

Reported incidence, mortality, and prevalence cases are drawn from a beta-binomial distribution with reporting rate ρ and dispersion parameter θ_{rep} (Morris, 1997). We assume that cases are reported instantaneously but a more realistic model can incorporate a delay between infection time and reporting time (e.g., infected individuals are likely to report after symptom onset but before recovery). We do not explore these ideas here.

Instead, we focus on how different assumptions about an epidemic time series (whether it represents incidence, mortality, or prevalence) affect our inference. While this question is initially motivated by the TSIR framework, where incidence is assumed to be equal to prevalence when reporting time is equal to mean generation time, the actual problem of identifying the correct representation of an epidemic time series is far more common. For example, incidence has been assumed to be measured either upon transition from latent to infectious stage (Hooker et al., 2010; Althaus, 2014) or upon recovery (He et al., 2009), where the latter assumption is equivalent to our definition of mortality reports. Understanding these differences is crucial in order to correctly estimate underlying parameters as well as their associated uncertainties. For all other purposes, we use incidence time series throughout this study.

2.3   Fitting methods

Time-series SIR model

The time-series SIR (TSIR) method was developed by Bjørnstad et al. (2002) to estimate seasonally varying transmission rates of measles. The TSIR method begins by discretizing the infection process of the SIR model, assuming that disease generation time is equal to reporting interval:

$$\begin{aligned}
S_{t+1} &= B_t + S_t - i_{t+1} \\
i_{t+1} &= \beta_t S_t \frac{i_t^{\alpha}}{N_t}
\end{aligned} \tag{4}$$

where α is often referred to as a conversion factor from a continuous-time model to a discrete-time model or a heterogeneity parameter (Glass et al., 2003). Tak-

ing log on both sides, estimation of the transmission rates β_t becomes a regression problem:

$$\log i_{t+1} = \alpha \log \beta_t + \log S_t + \alpha \log i_t - \log N_t + \epsilon_t. \tag{5}$$

A key assumption behind the TSIR model is that all, if not most, individuals become infected eventually. In other words, the total number of births should be equal to the total number of cases over a long period of time. Rewriting true incidence i_t as a product of reported cases $i_{t,\mathrm{rep}}$ and reciprocal of reporting rate $1/\rho_t$ and iteratively solving (4), we get

$$\sum_{t=1}^{N} B_t = \sum_{t=1}^{N} \frac{i_{t,\mathrm{rep}}}{\rho_{t+1}} + Z_{t+1} - Z_1, \tag{6}$$

where $Z_t = S_t - \bar{S}$. Under this assumption, fitting a regression model to cumulative birth as a function of cumulative cases allows us to estimate time-varying reporting rate ρ_t from the recirpocal of the slope and susceptible dynamics Z_t from the residuals (Finkenstadt and Grenfell, 2000). Conditional on the known susceptible dynamics, we can apply the TSIR regression to epidemic time series.

Trajectory matching

Trajectory matching is a way of estimating underlying parameters of an ODE model by matching the observed time series and a deterministic solution of the model. Given observed incidence $i_{t,\mathrm{rep}}$ (or equivalently, observed prevalence or mortality cases), we can define its mean dynamics using the solution of the SIR model:

$$\mu_t = \frac{i_t}{\rho} = \frac{1}{\rho} \int_{t-t_{\mathrm{rep}}}^{t} \beta(s) S \frac{I}{N} ds. \tag{7}$$

Since it is difficult to integrate of the state-space, we include another state variable C that represents cumulative incidence where $dC/dt = \beta SI/N$ and numerically solve for C by integrating alongside the SIR model; taking the difference in C between two consecutive reporting periods allows us to obtain true i_t. Finally, the likelihood of observing the data can be defined as a deviation from this mean trajectory:

$$i_{t,\mathrm{rep}} \sim \mathrm{NegBin}(\mu_t, \phi_{\mathrm{obs}}), \tag{8}$$

where ϕ_{obs} is the negative binomial dispersion parameter.

Gradient matching

Gradient matching is a way of estimating underlying parameters of an ODE model by matching the gradients of a system, rather than the deterministic solution of the system (Ellner et al., 2002). Gradient matching provides a flexible way of modeling a system as it does not require the parameteric form of the rate equations to be pre-specified (e.g., we do not need to that per-capita rate

of transmission is a linear function of I). Instead, parameters can be estimated by fitting a non-parametric regression model to the "observed gradients" as a function of state variables that are relevant. Therefore, all state variables must be observed in order to apply gradient matching.

In practice, it is impossible to observe all state variables of an epidemic model, even for simple models such as the SIR model. When either incidence or mortality cases are reported, *no* state variables are observed. Nonetheless, it is still possible to estimate the transmission rates of the SIR model using gradient matching by making similar assumptions as the TSIR model. Here, we briefly outline the estimation process; note that the purpose of section is to introduce the method, rather than to provide a rigorous justification of the statistical model.

First, we begin by assuming that the mean generation time is equal to the length of reporting period and that incidence is proportional to prevalence, i.e., $i_t = aI_t$ for some $a > 0$. Note that this is a weaker assumption than the condition for the TSIR method. We then fit a smooth curve $f(t)$ to reported incidence $i_{t,\text{rep}}$ using a negative binomial family with a log link:

$$i_{t,\text{rep}} \sim \text{NegBin}\left(\exp(f(t)), \phi_{\text{smooth}}\right), \tag{9}$$

where natural cubic splines are used to model the function f [**SWP:** *CITE spline package*]. Basis knots are placed every 6 generations.

The smooth curve $f(t)$ represents the dynamics of the log of prevalence ($\log I_t$) up to a constant. Taking the derivative of the curve $f(t_i)$, we have an estimate of the derivative of the log of prevalence:

$$E\left[\frac{df(t)}{dt}\right] = \beta S - (\gamma + \mu). \tag{10}$$

There are several ways we can estimate the transmission rate using this equation. First, we can model transmission rate as a function of susceptibles S by fitting a non-parametric regression by treating known values of γ and μ as offset terms:

$$\frac{df(t)}{dt} \sim \text{Normal}(g(S_t) - (\gamma + \mu), \sigma^2), \tag{11}$$

where g is a non-parametric function that represents the transmission rate and the unobserved susceptible population S_t can be estimated from susceptible reconstruction. When transmission varies over time, we can either (1) model transmission rate as an interaction term between a factor variable that represents time of the year and a smoothing term (i.e., fit a separate nonparametric curve g for each factor level) or try to separate transmission rate from the susceptible dynamics using a log link function:

$$\frac{df(t)}{dt} + (\gamma + \mu) \sim \text{Normal}\left(\exp\left(\log \beta_t + \log S_t\right), \sigma^2\right), \tag{12}$$

where log number of susceptible individuals $\log S_t$ is treated as an offset term and the log transmission rate can be estimated is estimated as a function of time of year using a nonparametric regression.

Sequential Monte Carlo

Calculating the exact likelihood for a nonlinear stochastic system is practically impossible. Instead, Sequential Monte Carlo (particle filtering) allows the likelihood to be approximated via simulations. Let $X_t = \{\mathbf{x}_{t,i} \,|\, i = 1, 2, \ldots, n\}$ be the set of states (particles) at time t and h be the process model. For the SIR model, each particle $\mathbf{x}_{t,i} \in X$ consists of number of susceptible individuals S_t and number of infected individuals I_t. Then, we can predict the states at next time step $t + 1$ for each particle using the process model: $h(\mathbf{x}_{t,i})$. For each prediction $h(\mathbf{x}_{t,i})$, we can calcuate the likelihood $w_{t,i}$ of the observation at time $t+1$ and define marginal likelihood by taking the average: $\bar{w}_t$. Finally, we draw n weighted samples from the set $\{h(\mathbf{x}_{t,i}) \,|\, i = 1 = 2, \ldots, n\}$, where the weight of each $h(\mathbf{x}_{t,i})$ is proportional to w_i, to obtain the next set of particles X_{t+1}. This algorithm is repeated until the final time step. Then, complete likelihood of the data is given by the product of the marginal likelihood at each time step: $\prod \bar{w}_t$.

While Sequential Monte Carlo provides a powerful technique for approximating the likelihood, it is still impractical to fit a continuous-time stochastic model. For example, it can take more than [**SWP:** *10*] hours to simulate an 20-year-long measles time series with gillespie algorithm. Computing a likelihood with Sequential Monte Carlo would take more than a few days, let alone maximizing the likelihood.

Instead of fitting a continuous-time model, we can assume that an epidemic process is a partially observed discrete-time Markov process and discretize the system by using a hazard approximation of the transition probabilities between states:

$$
\begin{aligned}
B(t) &\sim \text{Poisson}(b(t)\Delta t) \\
F_S(t) &\sim \text{Binomial}(S_t, 1 - \exp(-(r_1 + r_2)\Delta t)) \\
F_I(t) &\sim \text{Binomial}(I_t, 1 - \exp(-(r_3 + r_4)\Delta t)) \\
F_{S,I}(t) &\sim \text{Binomial}\left(F_S(t), \frac{r_1}{r_1 + r_2}\right) \\
S(t + \Delta t) &= B(t) + S(t) - F_S(t) \\
I(t + \Delta t) &= FS, I(t) + I(t) - F_I(t) \\
C(t + \Delta t) &= C(t) + FS, I(t)
\end{aligned}
\tag{13}
$$

where $r_1 = \beta I/N$, $r_3 = \gamma$, and $r_2 = r_4 = \mu$. Random variables $F_S(t)$ and $F_I(t)$ represent the total number of individuals that move out from susceptible S and infected I states. Likewise, $F_{S,I}(t)$ represent the total number of individual move from susceptible state S to infected state t. Finally, the state variable $C(t)$ keeps track of cumulative incidence. Again, observation likelihood is defined using a negative-binomial distribution:

$$
i_{t,\text{rep}} \sim \text{NegBin}(\rho(C(t) - C(t - t_{\text{rep}})), \phi_{\text{obs}}).
\tag{14}
$$

Renewal equation

Previously discussed models assume that the shape of generation-interval distribution is exactly known. The continuous time SIR model assumes that generation intervals are exponentially distributed. The discrete-time SIR model represented as a partially observed Markov process assumes that generation intervals are geometrically distributed. Finally, the TSIR assumes that generation intervals are fixed. All these assumptions are unrealistic.

The renewal equation provides an explicit way of modeling incidence $i(t)$ as a function of previous incidence and generation-interval distributions $g(t)$:

$$i(t) = \mathcal{R}_0 S(t) \int_0^t g(t-s)i(s)ds. \tag{15}$$

is a way of modeling incidence as a function of previous incidence and generation-interval distributions explicitly.

The renewal equation allows us to estimate the shape of the generation-interval distributions as well as the transmission rates by modeling modeling incidence as a function of previous incidence and the generation interval distribution:

$$s \tag{16}$$

(Li et al., 2018)

3 Results

3.1 Types of time series

When reporting period is equal to mean generation time, incidence and prevalence are often assumed to be equivalent (Fig. 1). Indeed, for the simple SIR model (1), it can be shown that incidence, prevalence, and mortality are similar to each other under Euler approximation when reporting period is equal to mean generation time (see Appendix). Their deterministic dynamics are sufficiently similar that we expect these three reports to be nearly indistinguishable from each other when demographic stochasticity and observation error is introduced. However, this is not necessarily true for more complicated models. For a more realistic model with exposed period (SEIR model), incidence is similar to prevalence when reporting period is equal to mean *generation time* whereas mortality is similar to prevalence when reporting period is equal to mean *infectious period*. [**SWP:** *TODO: confirm*]

Even if incidence, mortality, and prevalence are expected to be exhibit similar dynamics other when reporting period is equal to mean generation time, it is important to distinguish them when we are estimating underlying parameters of the SIR model, especially when mean generation time $1/\gamma$ is not exactly known. To demonstrate this idea, we first simulate an epidemic time series representing prevalence over time by solving the SIR model numerically and calculating prevalence, $I(t)$, for 20 reporting periods, where reporting period
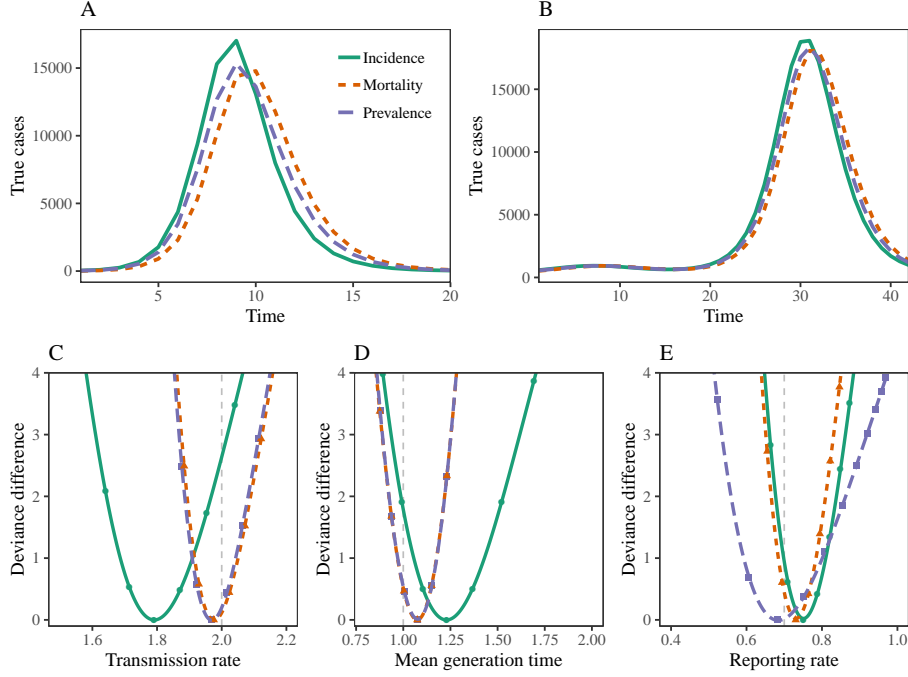
Figure 1: **Comparison of incidence, prevalence, and mortality from dynamical and estimation perspective.** A: dynamics of incidence, mortality and prevalence of the SIR model for a single outbreak (no natural birth/death and fixed transmission rate) when mean generation time ie equal to reporting period. The parameters and initial conditions are $\beta = 2$, $\gamma = 1$, $N = 1 \times 10^5$, $I(0) = 10$, and $S(0) = N - 10$. B: dynamics of incidence, mortality and prevalence of the SIR model with sinusoidal transmission rate ($\beta(t) = b_0(1 + b_1 \cos(2\pi t/26))$) when mean generation time ie equal to reporting period. The parameters and initial conditions are $b_0 = 500/26$, $b_1 = 0.15$, $\gamma = 1$, $\mu = 1/(50 \times 26)$, $N = 5 \times 10^6$, $I(0) = 0.0001N$, $S(0) = 0.05N$. C, D, E: deviance difference (profile likelihood minus the maximum likelihood) of each parameters when incidence, mortality, and prevalence are fit to same time series using trajectory matching. Grey dashed line represents the true value.

is equal to the mean generation time. Then, we draw a beta-binomial random variable at each time step with a reporting rate of 70% and an overdispersion parameter of 10. Using this time series, we try to estimate four parameters – transmission rate β, recovery rate γ, reporting rate ρ, and an initial number of infected individuals $I(0)$ – by treating the same time series as if it were incidence, mortality, and prevalence report (Fig. 1).

We find that fitting prevalence and mortality curves to the same time series yields almost identical estimates of the transmission rate β and the recovery

rate γ (as well as related uncertainty in the estimates) whereas fitting incidence and mortality curves to the same time series yields consistent estimates of the reporting rate ρ (as well as related uncertainty in the estimates). As mortality is approximately a lagged reflection of prevalence regardless of reporting period (reported mortality case at time t is approximately equal to $\rho\gamma I(t - t_{\mathrm{rep}})t_{\mathrm{rep}}$), it is possible to generate a mortality curve that matches its corresponding prevalence curve by adjusting the reporting rate ρ. This explains why estimates of dynamical parameters, β and γ, are similar whether we treat the same time series as prevalence or mortality. On the other ahnd, incidence report no longer matches prevalence report when reporting period is not equal to the mean generation time (i.e., when γ is being estimated); therefore, fitting incidence curve yields a different estimate of the dynamical parameters, β and γ, from fitting mortality or prevalence curves. Fitting incidence and mortality curve yields similar estimates of the reporting rate ρ because they both contain equivalent information about the final size of an epidemic: total *true* incidence and total *true* mortality is equal to the final size of an epidemic.

Even when mean generation time is exactly known, we find that estimates of transmission rate β and reporting rate ρ depends on our assumptions about the data. We find that fitting mortality and prevalence yields consistent estimates of both β and ρ in this case, whereas fitting incidence yields similar but still different estimates. Treating incidence report as prevalene report (and vice versa) should be avoided.

[**SWP:** *Say something more*]

4 Discussion

5 Data

6 TSIR model

6.1 Probability of infection

The TSIR model assumes that probability of infection is a nonlinear function of prevalence: $\beta I^{\alpha}/N$. This contrasts with the linear assumption of the Euler approximation ($\beta I/N$) or the exponenital assumption of the hazard-based approximation $(1 - \exp(-\beta I/N))$. Inclusion of the parameter α has been previously attributed to heterogeneity in the population or changes from continuous to discrete time model. However, there is still a room for clearer understanding of the actual role of the parameter α: for example, a few studies recognize that the estimated α does not always correspond to the best predictive α and resorted to assuming $\alpha = 0.97$.

Glass et al derived an analytical expression for optimal values of α by comparing the changes in the infected population at equilibrium using a second order Taylor expansion. Based on their approximation, optimal value of α equals 1

when there is no birth term. Instead, when we rewrite the TSIR model as

$$\log\left(\frac{I_{t+1}}{S_t}\right) = \log\left(\frac{\beta I_t^\alpha}{N_t}\right) + \epsilon_t. \tag{17}$$

it becomes much clearer that the TSIR model is trying to match the log probability of infection (when an epidemic is ongoing, changes in the susceptible population S_t through natural birth or death is negeligible over a generation, and I_{t+1}/S_t can be interpreted as the probability of infection over a generation). Then, we almost always expect α to be less than one because we expect the probability of infection (i.e., probability of finding a new host) to experience density dependence.

Though the difference in the interpretation of α is subtle, this perspective allows us to derive a useful connection between the discrete-time model and the continuous-time model. A naive estimate of basic reproductive number from the TSIR model is β. However, as discussed earlier, we often prefer to use hazard-based approach to approximate the continuous-time model. Then, we can find $\hat{\beta}$ such that $1 - \exp(-\hat{\beta}I/N)$ matches $\beta I^\alpha/N$ and use $\hat{\beta}$ as our estimate of the basic reproductive number instead. A simple way to match two functions is to match their values at the mean observed incidence. This corrections reduces bias in the estimate of the reproductive number by a large amount.

6.2 Process model

The TSIR model is often associated with three different process models: deterministic ($I_{t+1} = \beta_t S_t I_t^\alpha/N$), poisson ($I_{t+1} \sim \text{Poisson}(\beta_t S_t I_t^\alpha/N)$), and negative binomial ($I_{t+1} \sim \text{NB}(\beta_t S_t I_t^\alpha/N, I_t)$).

6.3 Susceptible reconstruction

One of the main assumptions behind the TSIR model is that the dynamics susceptible population S_t must be known. Assuming that all newborn infants become infected eventually, it is possible to recover the reporting rate, ρ_t, as well as the dynamics of the susceptible population as a deviation from its mean, $Z_t = S_t - \bar{S}$, by fitting a regression model to between cumulative births and cumulative cases:

$$\sum_{t=1}^{N} B_t = \sum_{t=1}^{N} \frac{C_{t+1}}{\rho_{t+1}} + Z_{t+1} - Z_1, \tag{18}$$

where C_t is the number of observed cases. Then, the mean susceptible population $\bar{S}$ can be estimated using profile likelihood with the TSIR regression.

While this method of reconstructing the susceptible dynamics has been successfully used with the TSIR method, several questions remain to be answered. First, how sensitive is our inference to regression methods? Second, does Z_t represent deviation from the overall population mean $\bar{S}$ or from some sort of moving average (assuming that population changes over time)? Likewise, how does changes in population size or reporting rate over time affect our estimate
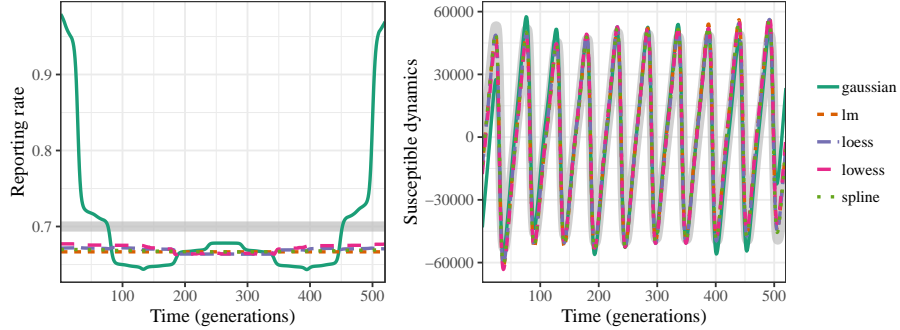
Figure 2: [**SWP:** *TODO*]

of ρ_{t+1} or Z_t? Finally, how do assumptions about the population size affect estimate of the transmission rate?

We first simulate an epidemic using the continuous deterministic SIR model for 10 years and apply 5 different regression models available from the `tsiR` package to test their ability to infer reporting rate and susceptible dynamics (Fig. **??**). We find that estimates of transmission rates are highly sensitive to regression methods. The gaussian regression is particularly inaccurate in boundaries. Other methods provide similar estimates. In general, estimates of susceptible dynamics appear to be robust regardless of the regression method.

Then, we compare estimates of susceptible dynamics and reporting rate under different scenarios (Fig. 3). We allow for population size to stay constant or increase over time and also allow for reporting rate to stay constant or increase over time; these combinations yields 4 different scenarios. We find that estimating time-varying reporting rate is difficult. We also find that Z_t estimated form regression better matches $S_t - \bar{S}$ rather than $S_t - \sigma N$, i.e., any long term dynamics in the susceptible population is captured in Z_t.
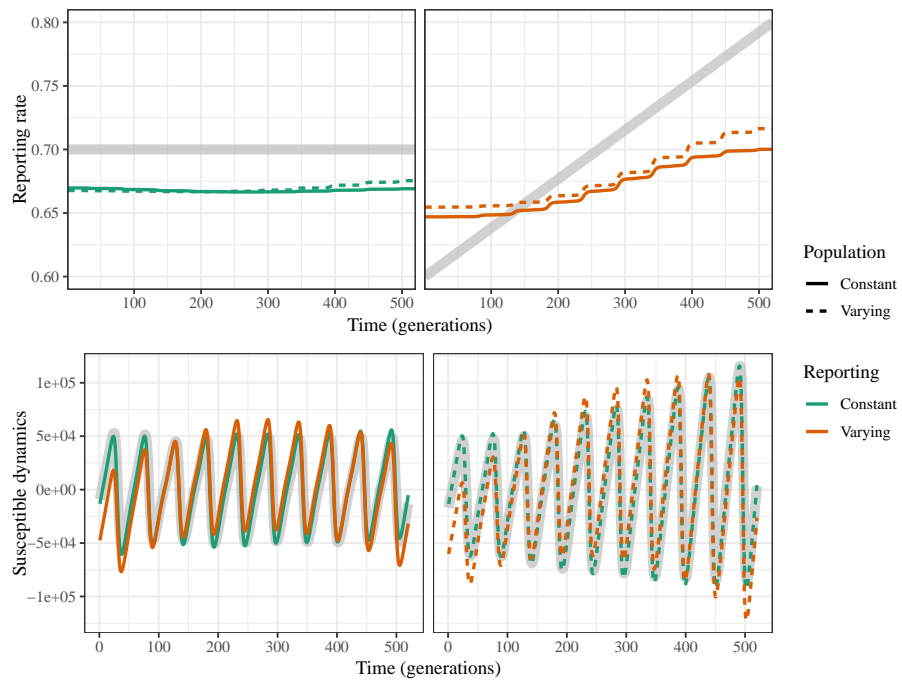
6.4   Splines

Figure 3: **[SWP:** *TODO*]

13

7 Other fitting methods

7.1 Trajectory matching

7.2 Gradient matching

7.3 Particle filtering

7.4 Renewal equation

8 Numerical comparison of different methods

8.1 Single outbreak

8.2 Time-varying transmission rate

9 Case study

10 Discussion

11   Appendix

11.1   Equivalence of incidence, mortality, and prevalence

First, recall that mortality case can be written as

$$\int_{t-t_{\text{rep}}}^{t} \gamma I ds. \tag{19}$$

When reporting period is equal to mean generation time, i.e., $t_{\text{rep}} = 1/\gamma$, it follows that:

$$\int_{t-t_{\text{rep}}}^{t} \gamma I ds \approx \gamma I(t - t_{\text{rep}})t_{\text{rep}} = I(t - t_{\text{rep}}) \tag{20}$$

Likewise, we can derive a similar relation for incidence:

$$
\begin{aligned}
I(t + t_{\text{rep}}) &= I(t) + \int_{t}^{t+t_{\text{rep}}} \frac{dI(s)}{ds} ds \\
&= I(t) + \int_{t}^{t+t_{\text{rep}}} \beta(s)S\frac{I}{N} ds - \int_{t}^{t+t_{\text{rep}}} \\
&\approx \int_{t}^{t+t_{\text{rep}}} \beta(s)S\frac{I}{N} ds
\end{aligned}
\tag{21}
$$

Here, we assumed that individuals leaving infected class I from natural mortality is negligible over the reporting period.

[**SWP:** *TODO*]

12   Fitting continuous models

12.1   Trajectory matching

Trajectory matching is one of the simplest fitting methods. As its name suggests, the goal is to find a set of parameters such that the deterministic trajectory of the ODE model best matches the observed data. This method assumes that there is only measurement error and no process error. The measurement model can be written as:

$$y_t \sim \text{NegBinomial}\left(\mu_t = \int_{t-\Delta t}^{t} \beta(s)S\frac{\hat{I}}{N} ds, \phi\right), \tag{22}$$

where ρ is reporting rate and ϕ is dispersion parameter. We prefer this model compared to the binomial or beta-binomial model because fitting binomial or beta-binomial models may not work if there are any irregular patterns in the data.
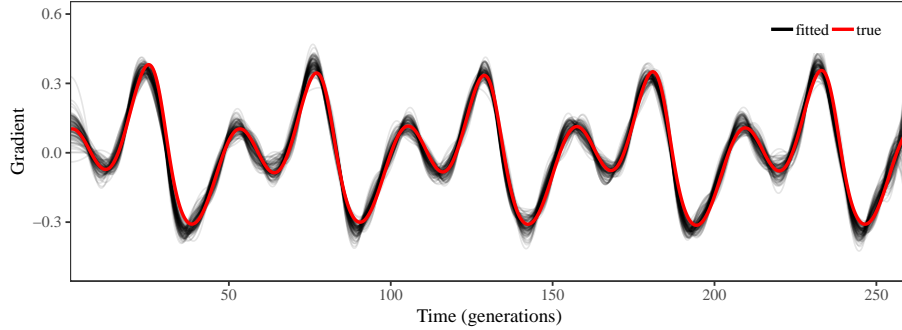
Figure 4: This is not a caption.

12.2   Gradient matching

When *all* state variables are observed, we can fit smooth curves to the data and obtain an estimate of the gradients of the ODE by taking the derivatives of the estimated smooth functions. Gradient matching methods seek to estimate the rate equation using nonparametric regression. However, for most epidemics, it is impossible to have data of all state variables. When only incidence is provided, we have no direct observations of any state variables. If we can reconstruct the dynamics of the susceptible population and assume that incidence is sufficiently similar to prevalence, we can still be able to apply gradient mathing methods.

Observe that

$$\frac{d \log \hat{I}}{dt} = \beta(t) \frac{S}{N} - (\gamma + \mu), \tag{23}$$

If we can obtain a reasonable estimate of the gradient $Z_i = d \log \hat{I}/dt(t_i)$, we can write

$$Z_i + (\gamma + \mu) = \exp\left(\log \beta(t_i) + \log S_i - \log N\right) + \epsilon_i \tag{24}$$

and estimate β, given that we can reconstruct the susceptible dynamics S_i. Estimate of β can be done using nonparametric regression by treating $\log S_i$ and $\log N$ as offset terms.

First, we want to test whether gradients can be reliably estimated from incidence data alone by simulating a deterministic SIR model with measurement error. We use natural cubic spline with knots placed every 6 biweeks. It works generally well but there is systematic bias. When gradients are at their local maxima (and minima), we tend to overestimate (and underestimate) the gradients. This is because incidence precedes prevalence???

If we simply construct confidence intervals based on equation ?? we underestimate uncertainty significantly by ignoring the uncertainty in the estimated gradients Z_i.

Conditions:

- See Jost and Ellner (???)

18

- sampled sufficiently

- not too much error

- Brunel Parameter estimation of ODE's via nonparametric estimators

12.3   Spectral matching

What is this? Read Reuman et al. 2006

References

Althaus, C. L. (2014). Estimating the reproduction number of ebola virus (ebov) during the 2014 outbreak in west africa. *PLoS currents 6*.

Bjørnstad, O. N., B. F. Finkenstädt, and B. T. Grenfell (2002). Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecological Monographs 72*(2), 169–184.

Bolker, B. M. (2008). *Ecological models and data in R.* Princeton University Press.

Chowell, G., N. W. Hengartner, C. Castillo-Chavez, P. W. Fenimore, and J. M. Hyman (2004). The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *Journal of theoretical biology 229*(1), 119–126.

Dalziel, B. D., O. N. Bjørnstad, W. G. van Panhuis, D. S. Burke, C. J. E. Metcalf, and B. T. Grenfell (2016). Persistent chaos of measles epidemics in the prevaccination United States caused by a small change in seasonal transmission patterns. *PLoS computational biology 12*(2), e1004655.

Earn, D. J., P. Rohani, B. M. Bolker, and B. T. Grenfell (2000). A simple model for complex dynamical transitions in epidemics. *Science 287*(5453), 667–670.

Ellner, S. P., Y. Seifu, and R. H. Smith (2002). Fitting population dynamic models to time-series data by gradient matching. *Ecology 83*(8), 2256–2270.

Finkenstadt, B. F. and B. T. Grenfell (2000). Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 49*(2), 187–205.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics 22*(4), 403–434.

Glass, K., Y. Xia, and B. Grenfell (2003). Interpreting time-series analyses for continuous-time biological models—measles as a case study. *Journal of theoretical biology 223*(1), 19–25.

He, D., E. L. Ionides, and A. A. King (2009). Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface*.

Hooker, G., S. P. Ellner, L. D. V. Roditi, and D. J. Earn (2010). Parameterizing state–space models for infectious disease dynamics by generalized profiling: measles in Ontario. *Journal of The Royal Society Interface 8*(60), 961–974.

Ionides, E. L., A. Bhadra, Y. Atchadé, A. King, et al. (2011). Iterated filtering. *The Annals of Statistics 39*(3), 1776–1802.

Ionides, E. L., D. Nguyen, Y. Atchadé, S. Stoev, and A. A. King (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proceedings of the National Academy of Sciences 112*(3), 719–724.

King, A. A., D. Nguyen, and E. L. Ionides (2015). Statistical inference for partially observed markov processes via the r package pomp. *arXiv preprint arXiv:1509.00503*.

Li, M., J. Dushoff, and B. M. Bolker (2018). Fitting mechanistic epidemic models to data: a comparison of simple Markov chain Monte Carlo approaches. *Statistical methods in medical research 27*(7), 1956–1967.

Morris, W. F. (1997). Disentangling effects of induced plant defenses and food quantity on herbivores by fitting nonlinear models. *The American Naturalist 150*(3), 299–327.

Riley, S., C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L.-M. Ho, T.-H. Lam, T. Q. Thach, et al. (2003). Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science 300*(5627), 1961–1966.

13   Appendix

13.1   Derivation of growth rate in discrete-time SIR model

Derivation:

$$
\begin{aligned}
S_{t+\Delta t} &= S_t - S_t(1 - \exp(-\beta I_t \Delta t/N)) \\
I_{t+\Delta t} &= S_t(1 - \exp(-\beta I_t \Delta t/N)) + I_t - I_t(1 - \exp(-\gamma \Delta t))
\end{aligned}
\tag{25}
$$

Assuming tat S_t is approximately equal to N, we get

$$
I_{t+\Delta t} = N_t(1 - \exp(-\beta I_t \Delta t/N)) + I_t \exp(-\gamma \Delta t)
\tag{26}
$$

When $I_t \approx 0$, $\exp(-\beta I_t \Delta t/N) \approx 1 - \beta I_t \Delta t/N$ and so

$$
\begin{aligned}
I_{t+\Delta t} &\approx I_t \beta \Delta t + I_t \exp(-\gamma \Delta t) \\
&= I_t(\beta \Delta t + \exp(-\gamma \Delta t))
\end{aligned}
\tag{27}
$$

Substituting $\hat{\gamma}$, we get

$$
I_{t+\Delta t} = I_t \left(1 + \beta \Delta t - \frac{\Delta t}{\mu} \right),
\tag{28}
$$

Therefore,

$$
I_t = I_0 \left(1 + \beta \Delta t - \frac{\Delta t}{\mu} \right)^{t/\Delta t}
\tag{29}
$$

and the initial growth rate is given by

$$
r = \frac{1}{\Delta t} \log \left(1 + \beta \Delta t - \frac{\Delta t}{\mu} \right).
\tag{30}
$$

13.2   Probability of infection

In both the TSIR model and the hazard-based model, transition from the susceptible comparment, S, to the infected compartment, I, is represented as a product of number of susceptible individuals and probability of infection between two time steps t and $t + \Delta t$. The TSIR model assumes that the probability of infection is a linear function of prevalence I_t. This formulation is based on the Euler approximation to the solution of the ordinary differential equation. On the other hand, the hazard-based model assumes that the probability of infection is an inverse exponential function of prevalence I_t. Besides their differences in the relationship between prevalence and probability of infection, there are two more assumptions that we need to consider: (1) force of infection remains constant over two time steps and (2) new susceptible individuals have zero probability of infection between the two time steps.

13.3 Infectious period

Geometric distribution with probability of $1 - \exp(-\gamma \Delta t)$ and unit of Δt. Then, mean infectious period and generation interval is given by

$$\frac{\Delta t}{1 - \exp(-\gamma \Delta t)}. \tag{31}$$

An important but often overlooked component is the variance of the distribution. Squared coefficient of variation for this distribution is equal to $\exp(-\gamma \Delta t)$, which necessarily depends on pre-specified time-step Δt. If we want to match the mean of the distribution to a fixed value μ regardless of our choices of Δt, we obtain

$$\hat{\gamma} = -\frac{1}{\Delta t} \log \left(1 - \frac{\Delta t}{\mu} \right) \tag{32}$$

Then,

$$\mathrm{CV}_{\text{Infectious period}}(\Delta t)^2 = 1 - \frac{\Delta t}{\mu}. \tag{33}$$

We expect changes in variation in infectious period to affect variation in stochstic realizations.

This means that the relationship between r and $\mathcal{R}$ changes. For this generation-interval distribution, we have

$$\mathcal{R} = 1 + \frac{\exp(r) - 1}{(1 - \exp(-\gamma \Delta t))} \tag{34}$$

Substituting

$$r = \frac{1}{\Delta t} \log \left(1 + \beta \Delta t - \frac{\Delta t}{\mu} \right), \tag{35}$$

as well as $\hat{\gamma}$, we get

$$\mathcal{R} = 1 + \frac{\mu}{\Delta t} \left(1 + \beta \Delta t - \frac{\Delta t}{\mu} \right)^{1/\Delta t} \tag{36}$$

Then, we can obtain a relationship between $\mathcal{R}$ of a discrete-time system and that of a corresponding continuous time system, assuming that contact rate β and mean generation-time μ are the same:

$$\mathcal{R}_{\text{discrete}} = 1 + \frac{\mu}{\Delta t} \left(1 + \frac{(\mathcal{R}_{\text{continuous}} - 1)\Delta t}{\mu} \right)^{1/\Delta t} \tag{37}$$

On the other hand, we may want to match choose β and μ to match true r and $\mathcal{R}$.