

Implementing the linear noise approximation in Stan

Jon Fintzi

June 12, 2017

Summary

Stochastic epidemic models (SEMs) describe the transmission dynamics of infectious diseases. These models are often represented as density dependent Markov jump processes (MJPs), which have systems of ordinary differential equations as functional limits (as the population size tends to infinity). The canonical example of a SEM is the susceptible-infected-recovered (SIR) model for a closed and homogeneously mixing population. This model, presented above, is so called because an epidemic is represented in terms of the disease histories of individuals as they stochastically transition between susceptible, infected, and recovered states (model compartments). In this model, per-contact infectivity rate is given by β , so with S susceptibles and I infecteds the rate at which a new infection occurs is βSI . Similarly, each infected (and hence infectious) individual recovers with rate μ , and thus the rate at which a recovery occurs is given by μI . A loose explanation for what we mean by density dependent in characterizing the MJP is that the dynamics of the process are driven by the fraction of the population in each compartment (i.e., we can equivalently represent the model in terms of concentrations and maintain the dynamics).

Let $X = (S, I, R)$ be the vector of SIR compartment counts, $N = (N_{SI}, N_{IR})$ be the counting process for the cumulative numbers of infection and recovery events, and $\phi(t, X) = (\beta SI, \mu I)$ be the vector of infection and recovery rates for the SIR model. It will be necessary in working with incidence data to reparameterize the compartment counts, and hence the vector of rates, ϕ , in terms of the counting process with respect to the initial compartment counts $X_0 = (S_0, I_0, R_0)$. Let

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

and note that we can express the compartment concentrations at time t in terms of the initial compartment counts and the cumulative numbers of infections and recoveries using the relation, $X(t) = X(0) + A'N(t)$. Thus, we can write the rates, ϕ , as

$$\phi(t, N) = \begin{pmatrix} \beta(S_0 - N_{SI})(I_0 + N_{SI} - N_{IR}) \\ \mu(I_0 + N_{SI} - N_{IR}) \end{pmatrix}.$$

In the infinite population limit, the counting process for the numbers of infections and recoveries evolves according to the deterministic path obtained by integrating the following system of ODEs:

$$\frac{dN(t)}{dt} = \phi(t, N), \quad N_0 = (0, 0), \quad X_0 = (S_0, I_0, R_0).$$

Clearly, the dynamics of the limiting ODE do not reflect the stochastic aspects of the time-evolution of the MJP. However, in the large population limit, it is possible to approximate the transition density of the MJP with the following Ito diffusion:

$$dN(t) = \phi(t, N)dt + \sqrt{\Phi(t, N)}dW(t), \quad s.t., N(0) = 0, \quad X_0 = (S_0, I_0, R_0).$$

where $\Phi(t, N) = \text{diag}(\phi(t, N))$, and $W(t)$ is a 2-dimensional (as we have two elementary transition events, infections and recoveries) standard Brownian motion.

Since the counting process N is positive and monotonically increasing, it will be advantageous to work with the log-transformed SDE for the log-transformed process, $Z = (Z_{SI}, Z_{IR}) = (\log(N_{SI}), \log(N_{IR}))$. However, we should proceed with caution as there is an obvious problem with the initial condition for N_0 , as $\log(0) = -\infty$.

We can resolve this by setting $N_0 = (1, 1)$ and “borrowing” events from X_0 in the following way: Note that $X(t) = X(0) + A'N(t) = X(0) + A'(N(t) - \mathbf{1} + \mathbf{1}) = (X(0) - A'\mathbf{1}) + A'(N(t) + \mathbf{1})$ for all t . Thus, an equivalent set of initial conditions for the SDE is given by $N_0 = (1, 1)$, $X_0 = (S_0, I_0, R_0) - A'\mathbf{1} \equiv \tilde{X}_0$, as long as $\tilde{X}_0 \geq 0$. In our context, this is not an unreasonable condition as it implies, for example, that there are not more infection events than there are initial susceptibles.

By Ito’s lemma the SDE for the log-transformed process, Z , subject to initial conditions $Z_0 = (0, 0)$, $X_0 = \tilde{X}_0$, is given by

$$dZ(t) = \left[\text{diag}(\exp(-Z(t))) - \frac{1}{2} \text{diag}(\exp(-2Z(t))) \right] \phi(t, \exp(Z(t))) + \text{diag}(\exp(-Z(t))) \sqrt{\Phi(t, \exp(Z(t)))} dW(t).$$

The LNA is derived by Taylor expanding the drift and diffusion terms in the above SDE about the deterministic limit of the MJP. The resulting approximation decomposes the Ito SDE into a non-linear deterministic drift and a linear SDE for the fluctuations about the drift. This linear SDE has an analytic solution that is given by a multivariate Gaussian, the moments of which are obtained by solving a coupled, non-autonomous system of ODEs. The LNA transition density over the interval $[t_{\ell-1}, t_\ell]$ thus takes the form,

$$Z(t_\ell) | Z(t_{\ell-1}) = \boldsymbol{\eta}(t_\ell) + \mathbf{m}(Z(t_{\ell-1}) - \boldsymbol{\eta}(t_{\ell-1})) + \boldsymbol{\epsilon}_\ell, \quad \boldsymbol{\epsilon}_\ell \sim N(0, \boldsymbol{\Sigma}(t_\ell)),$$

where

- $\boldsymbol{\eta}(t_\ell)$ is the solution to $\frac{d\boldsymbol{\eta}(t)}{dt} = [\text{diag}(e^{-\boldsymbol{\eta}(t)}) - \frac{1}{2} \text{diag}(e^{-2\boldsymbol{\eta}(t)})] \boldsymbol{\Phi}(e^{\boldsymbol{\eta}(t)}) \equiv \boldsymbol{\xi}(\boldsymbol{\eta}(t))$.
- $\mathbf{m}(Z(t_{\ell-1}) - \boldsymbol{\eta}(t_{\ell-1}))$ is the solution to $\frac{d\mathbf{m}(t)}{dt} = \mathbf{F}(t)\mathbf{m}(t)$, where $\mathbf{F}(t) = \left(\frac{\partial \boldsymbol{\xi}(Z(t))}{\partial \mathbf{Z}(t)} \right) \Big|_{\boldsymbol{\eta}(t)}$.
- $\boldsymbol{\Sigma}(t)$ is the solution to $\frac{d\boldsymbol{\Sigma}(t)}{dt} = \mathbf{F}(t)\boldsymbol{\Sigma}(t) + \boldsymbol{\Sigma}(t)\mathbf{F}(t)' + \text{diag}(e^{-\boldsymbol{\eta}(t)}) \boldsymbol{\Phi}(\boldsymbol{\eta}(t)) \text{diag}(e^{-\boldsymbol{\eta}(t)})$.
- The above ODEs are subject to the initial conditions: $\mathbf{X}(t_{\ell-1}) = \mathbf{x}_0 - A'\mathbf{1} + A' \exp(Z(t_{\ell-1}))$, $\boldsymbol{\eta}(t_{\ell-1}) = \mathbf{0}$, $\mathbf{m}(t_{\ell-1}) = \mathbf{0}$ ($\implies \mathbf{m}(t_\ell) = \mathbf{0}$), and $\boldsymbol{\Sigma}(t_{\ell-1}) = \mathbf{0}$.

In practice, we will use a non-centered parameterization for the LNA representation of the latent epidemic process in which we map a vector of independent Gaussian draws, $\mathbf{N}_{\text{raw}} \stackrel{i.i.d.}{\sim} N(0, 1)$, deterministically to a latent LNA path by noting that $Z(t) \stackrel{\mathcal{L}}{=} \boldsymbol{\eta}(t_\ell) + \boldsymbol{\Sigma}(t_\ell) N_{\text{raw}}(t_\ell)$. Furthermore, we reset the initial conditions for the compartment volumes to reflect the cumulative numbers of infections and recoveries up to the beginning of each observation interval, and reset the counting process for the infections and recoveries. The LNA moments for the log transformed process for the next interval are then computed, and we obtain positive increments in the numbers of infections and recoveries by exponentiating the log-LNA values that were obtained via the deterministic mapping.

Note that the latent LNA process is Markov. Thus, the observed data likelihood, conditional on a latent LNA path, decomposes into a product of emission probabilities, which in the SIR example in this post is a product of negative binomial probabilities.