Likelihood-Based Inference for Partially Observed Spatiotemporal Dynamics: Measles as a Case Study

Advances in Spatio-Temporal Epidemiology
JSM2017, 7/30/2017, 4:55 - 5:20 PM, CC-310

Edward Ionides and Joonha Park
University of Michigan, Department of Statistics

Slides are online at
`http://dept.stat.lsa.umich.edu/~ionides/talks/jsm17.pdf`

We seek a “mechanistic” approach to modeling and inference for dynamic systems. What does this mean?

- Write down equations, based on scientific understanding of a dynamic system, which describe how it evolves with time.
- Further equations describe the relationship of the state of the system to available observations on it.
- Data analysis via mechanistic models concerns drawing inferences from the available data about the hypothesized equations.
- Questions of general interest: Are the data consistent with a particular model? If so, for what range of values of model parameters? Does one mechanistic model describe the data better than another?
- A defining principle: the model structure should be chosen based on scientific considerations, rather than statistical convenience.

Statistical challenges for nonlinear mechanistic modeling in ecology and epidemiology

1. Combining measurement noise and process noise.
2. Including covariates in mechanistically plausible ways.
3. Continuous time models.
4. Modeling and estimating interactions in coupled systems.
5. Dealing with unobserved variables.
6. **Spatiotemporal data and models.**
7. Inferences from genetic sequence data.

(1–6) were enumerated by Bjornstad and Grenfell (*Science*, 2001).

(1–5) are now routinely solved using modern methods for nonlinear partially observed Markov process (POMP) models (Ionides et al., 2015; King et al., 2016).

(7) was described by Grenfell et al (Science, 2004) and a general POMP solution was shown by Smith et al (*Molecular Biology & Evolution*, 2017).

Overview of sequential Monte Carlo

- **Sequential Monte Carlo (SMC)**, a.k.a. **the particle filter**, is a standard tool for fitting mechanistic dynamic models to nonlinear non-Gaussian time series.
- SMC allows full-information statistical inference. Standard MCMC methods struggle with POMP models; many other methods involve information loss or approximations.
- SMC struggles with a **curse of dimensionality** preventing the use of the basic algorithm when the dimension of the dynamic system gets large (in practice, say, more than 10 dimensions).
- Theoretical results suggest that, in some situations, this curse can avoided (Rebeschini and van Handel, 2015).
- We have a method that partially avoids the curse and is practical on some problems with 80 latent dynamic dimensions and 20 measured dimensions: Susceptible-Exposed-Infected-Recovered (SEIR) dynamics for measles in 20 connected cities.

A modified SMC algorithm for spatiotemporal data

- The **Guided intermediate resampling filter (GIRF)** breaks up the information in the data into small pieces that are used incrementally to inform the particles and “guide” them toward the next observation.
- Observations $y^*_{1:N} = (y^*_1, \ldots, y^*_N)$ are collected at times $t_{1:N}$.
- We require a latent Markov process $\{X(t), t_0 \le t \le t_N\}$ to be defined in continuous time. We assess particles at intermediate times

  $$t_{n,s} = \left(1 - \frac{s}{S}\right) t_n + \frac{s}{S} t_{n+1}$$

  using a reweighting function

  $$u_{n,s}(x).$$

- Our algorithm works asymptotically for almost any $u_{n,s}(x)$, but gains numerical efficiency if this reweighting function approximates the conditional likelihood of subsequence measurements.

Input and output for the guided intermediate resampling filter (GIRF)

input:

Simulator for latent process initial density, $f_{X_0}(x_0\,;\theta)$

Simulator for transition density, $f_{X(t)|X(s)}(\cdot\,|\,\cdot\,;\theta)$, $t_0 \le s < t \le t_N$

Evaluator for measurement density, $f_{Y_n|X_n}(\cdot\,|\,\cdot\,;\theta)$, $n \in 1\!:\!N$

Data, $y^*_{1:N}$. Parameter vector, θ. Number of particles, J.

Number of intermediate reweighting steps, S.

Evaluator for reweighting function, $u_t(x_t)$.

output:

Filtered particles, $\{X_N^{F,j}, j \in 1\!:\!J\}$.

Log likelihood estimate, $\hat{\ell} \approx \log f_{Y_{1:N}}(y^*_{1:N}\,;\theta)$.

Algorithms based on a simulator of the dynamic model are **plug-and-play**. This property ensures broad applicability.

GIRF has the structure of a particle filter with the addition of the **reweighting function** evaluated at the **intermediate reweighting steps**.

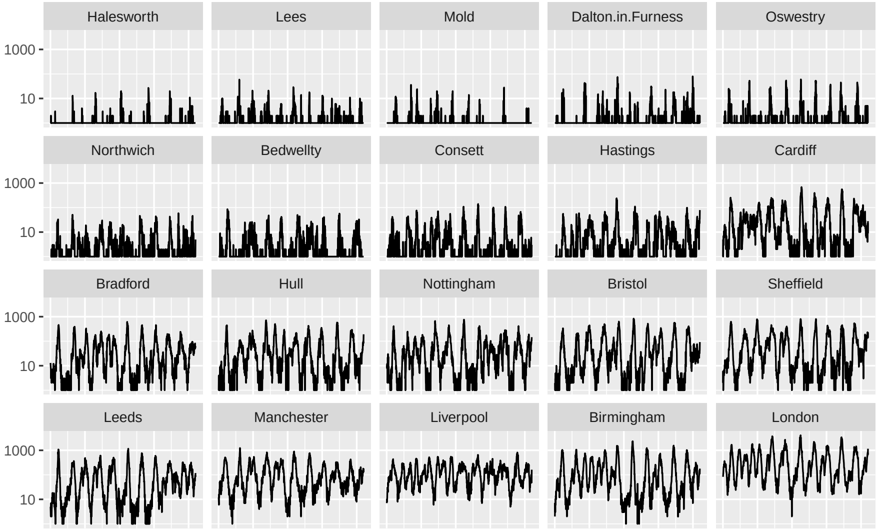
The guided intermediate resampling filter (GIRF)

```
Initialize: $\hat{\ell} = 0$,   $X^{F,j}_{0,0} \sim f_{X_0}(x_0\,;\theta)$,   $u_j = 1$ for $j$ in $1\!:\!J$.
For $n$ in $0\!:\!N-1$
    if $n \geq 1$ then $u^j = u^j \big/ f_{Y_n|X_n}(y^*_n \,|\, X^{F,j}_{n,0}\,;\theta)$
    For $s$ in $1\!:\!S$
        $X^{P,j}_{n,s} \sim f_{X_{n,s}|X_{n,s-1}}(x_{n,s} \,|\, X^{F,j}_{n,s-1};\theta)$ for $j$ in $1\!:\!J$
        $v^j = u_{n,s}(X^{P,j}_{n,s})$
        $w^j = v^j \big/ u_j$
        $\hat{\ell} = \hat{\ell} + \log\left\{ \frac{1}{J}\sum_{j=1}^{J} w^j \right\}$
        Draw $a_{1:J}$ with $\mathbb{P}(a_j = i) = w^m_{n,i} \Big/ \sum_{i'=1}^{J} w^m_{n,i'}$
        $X^{F,j}_{n,s} = X^{P,a_j}_{n,s}$ and $w^j = v^{a_j}$ for $j$ in $1\!:\!J$
    End For
    $X^{F,j}_{n+1,0} = X^{F,j}_{n,S}$
End For
```

From filtering to inference

- Filtering estimates latent states given data and a model. It doesn't estimate parameters.
- Iterated filtering algorithms (Ionides et al., 2015) filter repeatedly with perturbed parameters to maximize the likelihood.
- Iterated filtering tends to be more effective on large POMP models than alternative methods (e.g., Particle Markov chain Monte Carlo, Monte Carlo Expectation-Maximization).
- GIRF can be used within an iterated filtering algorithm.

Measles in 20 UK cities, 1944–1965



Motivation for studying measles

- Standard models have arisen from the extensive study of pre-vaccination measles. To demonstrate the new methodology, we chose a situation where basic models for transmission within a city are well established.
- Coupling between cities is less well understood. We analyze the ten largest plut ten other randomly selected UK cities.
- Coupling between neighborhoods of a single city, or aggretated at county or state level gives rise to similar numbers of spatial units.
- Measles remains a major cause of morbidity and mortality globally. It may be an upcoming target for global eradication.

A spatiotemporal model for measles

- We start with the Susceptible-Exposed-Infected-Recovered (SEIR) measles model of He et al. (2010) and add spatial interaction.
- For each city k, the population dynamics satisfy a set of equations,

$$\begin{aligned} \frac{dS_k}{dt} &= -\frac{dN_{SE,k}(t)}{dt} - \mu S_k(t) + r_k(t) \\ \frac{dE_k}{dt} &= \frac{dN_{SE,k}(t)}{dt} - \frac{dN_{EI,k}(t)}{dt} - \mu E_k(t) \qquad\qquad k = 1, \cdots, d, \\ \frac{dI_k}{dt} &= \frac{dN_{EI,k}(t)}{dt} - \frac{dN_{IR,k}(t)}{dt} - \mu I_k(t), \end{aligned}$$

where, $N_{SE,k}(t), N_{EI,k}(t), N_{IR,k}(t)$ denote the cumulative number of transitions between the compartments up to time t in city k, μ is the per-capita mortality rate, and r_k the susceptible recruitment rate.

The cumulative transitions were modelled as negative binomial processes, following the construction of Bretó et al. (2009). Specifically,

$$\begin{aligned}&\mathbb{E}\left[N_{SE,k}(t+dt) - N_{SE,k}(t)\right] \\ &\qquad = \beta(t) \cdot S_k(t) \cdot \left[\left(\frac{I_k}{P_k}\right)^{\alpha} + \sum_m \frac{v_{km}}{P_k}\left(\left(\frac{I_m}{P_m}\right)^{\alpha} - \left(\frac{I_k}{P_k}\right)^{\alpha}\right)\right] dt + o(dt)\end{aligned}$$

where $\beta(t)$ is transmission coefficient with time dependence due to seasonality, α is a mixing coefficient, P_k is the population at city k, and v_{kl} the number of travelers from city k to l.

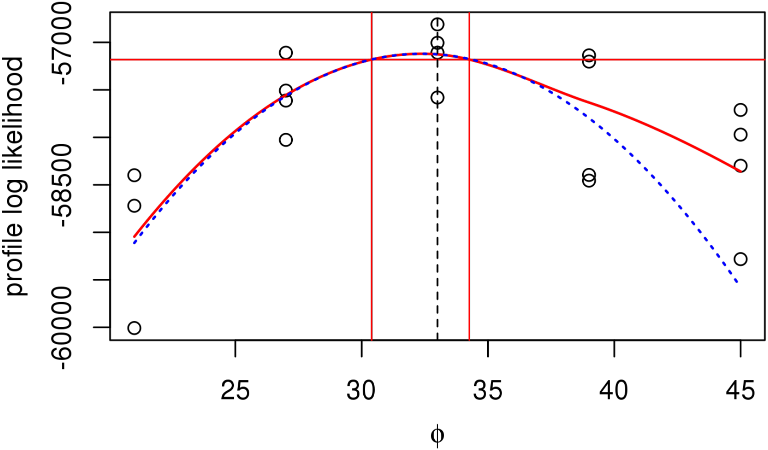
The gravity model of Xia et al. (2004) describes the number of travelers:

$$v_{kl} = G \cdot \frac{\bar{d}}{\bar{P}^2} \cdot \frac{P_k \cdot P_l}{d_{kl}}.$$

Here, the gravitation constant G was rescaled using the average population of all 20 cities, $\bar{P}$, and the average distance of all pairs of cities, $\bar{d}$.

Data are weekly cumulative reported cases, modeled using an overdispersed binomial distribution.

Coupled measles SEIR in 20 cities: profiling contact rate



Monte Carlo adjusted profile (MCAP) methodology of (Ionides et al., 2017) indicates a cutoff of 61.6, rather than the usual 1.92, for the confidence interval construction. Here, Monte Carlo variability is larger than statistical uncertainty.

Conclusions

- The guided intermediate resampling filter (GIRF) can permit statistical inference for coupled nonlinear partially observed stochastic dynamic systems of moderate dimension.
- GIRF enables likelihood-based inference for a spatiotemporal SEIR model with 20 coupled cities.
- GIRF is plug-and-play, though it needs a tractable approximation to a conditional transition density.
- Techniques assisting the use of a Monte Carlo filter for parameter estimation and hypothesis testing include:
  - Iterated filtering methodology to adapt a successful filter for maximum likelihood estimation.
  - Monte Carlo adjusted profile (MCAP) methodology to enable proper inference despite non-negligible Monte Carlo error.

References I

Bjørnstad, O. N. and Grenfell, B. T. (2001). Noisy clockwork: Time series analysis of population fluctuations in animals. *Science*, 293:638–643.

Bretó, C., He, D., Ionides, E. L., and King, A. A. (2009). Time series analysis via mechanistic models. *Ann. Appl. Stat.*, 3:319–348.

He, D., Ionides, E. L., and King, A. A. (2010). Plug-and-play inference for disease dynamics: Measles in large and small towns as a case study. *J. R. Soc. Interface*, 7:271–283.

Ionides, E. L., Breto, C., Park, J., Smith, R. A., and King, A. A. (2017). Monte Carlo profile confidence intervals for dynamic systems. *Journal of the Royal Society Interface*, To appear.

Ionides, E. L., Nguyen, D., Atchadé, Y., Stoev, S., and King, A. A. (2015). Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. *Proc. Natl. Acad. Sci. USA*.

References II

King, A. A., Nguyen, D., and Ionides, E. L. (2016). Statistical inference for partially observed Markov processes via the R package pomp. *Journal of Statistical Software*, 69:1–43.

Rebeschini, P. and van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *The Annals of Applied Probability*, 25(5):2809–2866.

Smith, R. A., Ionides, E. L., and King, A. A. (2017). Infectious disease dynamics inferred from genetic data via sequential Monte Carlo. *Molecular Biology and Evolution*, pre-published online, doi:10.1093/molbev/msx124.

Xia, Y., Bjørnstad, O. N., and Grenfell, B. T. (2004). Measles metapopulation dynamics: A gravity model for epidemiological coupling and dynamics. *Am. Nat.*, 164(2):267–281.

Theoretical results

- Assumption 1. The predictive likelihood can be closely approximated.
- Assumption 2. The length of subinterval is sufficiently small.
- Assumption 3. The POMP possesses conditional mixing property given data.

Sketch of Theorem: Under assumptions 1, 2, and 3, for any h with $\|h\|_\infty \leq 1$,

$$\left| \frac{1}{J} \sum_{j=1}^{J} h(X_{t_N}^{F,j}) - \mathbb{E}\big[h(X_N) | Y_{1:N} = y_{1:N}^*\big] \right| \leq a_1 + \frac{a_2(d)}{\sqrt{J}}$$

with high probability. The constant $a_2(d)$ increases slowly in d and is dependent on the accuracy of $u_{n,s}$ as an approximation to the predictive likelihood.