Particle Markov Chain Monte Carlo with Applications to Partially Observed Markov Processes

John Mattiace

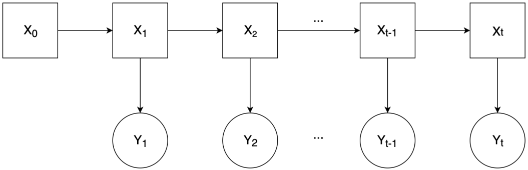
December 10, 2019

Motivation: Modeling Infectious Disease Dynamics

- Put simply, public health is important. Interventions cost social, financial, and political capital to implement effectively, e.g. mass vaccination programs
- Modeling how diseases respond to given conditions helps shape policy, insofar as we can affect those conditions
- Unfortunately, we often do not have access to full, high-quality datasets, privacy concerns notwithstanding
  - Latent period between infection and appearance of symptoms
  - Doctors can misdiagnose illness
  - Mild cases of illness might keep people from visiting a healthcare professional
- In summary, we have some process happening, hidden from us, which we only observe through some mechanism of measurement

Enter POMP

- Underlying (Hidden) Process, X_t:
  - A state-space time series
  - $X_t \sim f_{X_t|X_{(t-1)}}(\cdot|X_{(t-1)};\boldsymbol{\theta})$
  - Assumed to adhere to Markov assumption
- Measurement Process, Y_t:
  - A function such that $Y_t \sim f_{Y_t|X_t}(\cdot|X_t)$
  - This generates our observations



Goal: Want to do inference on some facet of X_t, be it parameters, initial states, prediction, etc.

One Solution: Standard M-H MCMC

From Bayes: $\pi(\theta|Y_t) \propto f(Y_t|\theta)\pi(\theta)$

$$f(Y_2|\theta) = \int f(Y_2|X_2,\theta)f(X_2|X_1,\theta)dx_2$$

Idea: Treat the hidden process as latent variables and "integrate them out"

If time series short enough, can use adaptive MCMC methods to explore $\pi(X,\theta|Y) \propto f(Y|X,\theta)f(X|\theta)\pi(\theta)$

Limitations:

- Hard to come up with proposal distribution, esp. nonlinear, high-dim
- As T increases, sampling the entire time series becomes difficult
- It is unrealistic to assume that our samples of the hidden states are independent

Alternatively, Sequential Monte Carlo

Consider

$$\ell(\theta) = \sum_{t=1}^{T} \ell_{t|1:t-1}(\theta) = \sum_{t=1}^{T} \log f(Y_t|Y_{1:(t-1)}, \theta) = \log f(\mathbf{Y}|\theta)$$

$$\ell_{t|1:t-1}(\theta) = \log \int f_{Y_t|X_t}(Y_t|X_t;\theta) f_{X_t|Y_{1:(t-1)}}(X_t|Y_{1:(t-1)};\theta) dX_t$$

Idea: Easy to consider single time step vs. multiple time steps. Rather than sampling a whole time series, simulate from one step to the next to provide samples of the time series, iteratively

Very useful when likelihood is intractable, as in the case of POMPs

Pseudocode

PMCMC

1. Initialize $\boldsymbol{\theta_0}$, $\hat{\ell}(\theta_0)$
2. Draw $\boldsymbol{\theta}^* \sim q(\cdot|\boldsymbol{\theta})$
3. Run SMC algorithm using $f_{X_t|X_{(t-1)}}(\cdot|X_{(t-1)};\boldsymbol{\theta^*})$
4. Calculate acceptance probability $\alpha = \frac{\pi(\boldsymbol{\theta}^*)\exp(\hat{\ell}(\boldsymbol{\theta}^*))}{\pi(\boldsymbol{\theta_{m-1}})\exp(\hat{\ell}(\boldsymbol{\theta}_{m-1}))}$
5. Draw $U \sim \mathrm{Unif}(0,1)$
6. If $U < \alpha$, set $\hat{\ell}(\boldsymbol{\theta}_m) = \hat{\ell}(\boldsymbol{\theta}^*)$ and $\boldsymbol{\theta}_m = \boldsymbol{\theta}^*$
7. Otherwise, set $\hat{\ell}(\boldsymbol{\theta}_m) = \hat{\ell}(\boldsymbol{\theta}_{m-1})$, $\boldsymbol{\theta}_m = \boldsymbol{\theta}_{m-1}$
8. Repeat 2-7 for M total iterations

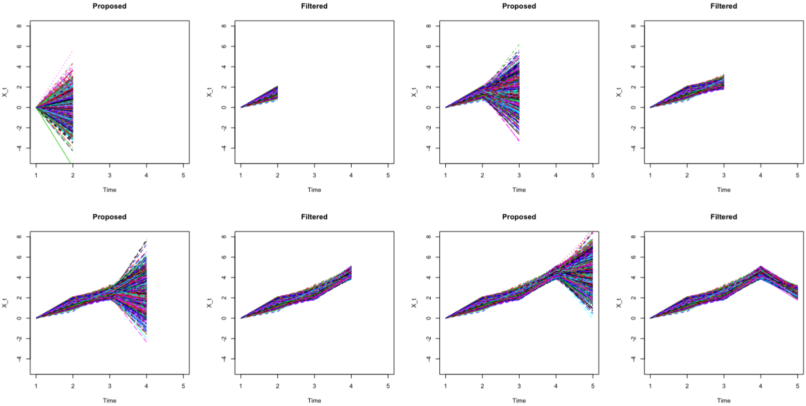
SMC

1. Initialize P particles: $X_{p,0} \sim \pi_{X_0}(\cdot)$
2. Simulate from each particle: $X^*_{p,t} \sim f_{X_t|X_{(t-1)}}(\cdot|X_{p,(t-1)};\boldsymbol{\theta^*})$
3. Weight each simulated value: $w_{p,t} = f_{Y_t|X_t}(Y_t|X^*_{p,t})$
4. Sum weights: $\hat{\ell}_t(\boldsymbol{\theta^*}) = \sum_{p=1}^{P} w_{p,t}$
5. Normalize weights: $\tilde{\mathbf{w}}_\mathbf{t} = \frac{w_{p,t}}{\hat{\ell}_t(\boldsymbol{\theta^*})}$
6. Resample particles with replacement with prob $\tilde{\mathbf{w}}_\mathbf{t}$
7. Repeat 2-6 for T total iterations
8. Return $\hat{\ell}(\boldsymbol{\theta}^*) = \sum_{t=1}^{T} \hat{\ell}_t(\boldsymbol{\theta^*})$

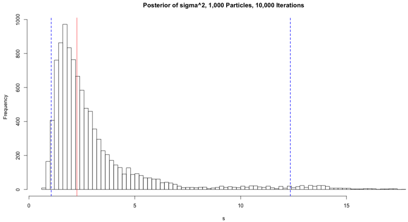
Random Walk Example

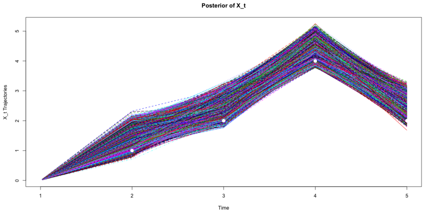
Say we observe the time series $\mathbf{Y} = (0, 1, 2, 4, 2)$; we have a measurement model $Y_t \sim \mathrm{floor}(N(X_t, 0.1))$; and we suspect that $X_t \sim N(X_{t-1}, \sigma^2)$

If we're interested about the process that could generate these data, then we ought investigate σ^2



Random Walk Example





Discussion

Summary:

- Particles sample trajectories of hidden process
- Parameter space explored in MCMC framework
- Performs better than augmented MCMC, which is very slow to explore parameter space among highly correlated parameters and hidden variables

Limitations:

- Potentially high computational load
- Measurement and underlying processes must be tractable
- Finite particles suffer in high-dimension setting

Further considerations:

- Application to hierarchical, compartmental models
- Inference on parameters of measurement model

References

Andrieu, C., Doucet, A., & Holenstein, R. (2010). Particle Markov Chain Monte Carlo methods. Journal of the Royal Statistical Society, 72, 269–342.

Endo, A., van Leeuwen, E., & Baguelin, M. (2019). Introduction to particle Markov-chain Monte Carlo for disease dynamics modellers. Epidemics, 29.

King, A. A., Nguyen, D., & Ionides, E. L. (2016). Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software, 69(12), 1–43.