Bayesian Computation for Stochastic Kinetic Models

Scott Phillips
Durham University

February 25, 2025

Stochastic Kinetic Model

Definition (Stochastic Kinetic Model)

A Stochastic Kinetic Model (SKM) is a mathematical framework used to describe the time evolution of the species in a system where reactions occur randomly.

Reaction Networks

Consider species $X_1, ..., X_U$ and reaction rates $\theta_1, ..., \theta_V$. The corresponding reaction network is as follows:

$$p_{1,1}X_1 + p_{1,2}X_2 + \cdots + p_{1,U}X_U \xrightarrow{\theta_1} q_{1,1}X_1 + q_{1,2}X_2 + \cdots + q_{1,U}X_U$$
$$\vdots$$
$$p_{V,1}X_1 + p_{V,2}X_2 + \cdots + p_{V,U}X_U \xrightarrow{\theta_V} q_{V,1}X_1 + q_{V,2}X_2 + \cdots + q_{V,U}X_U$$

for P and Q ($V \times U$) matrices.

Reaction Networks

Consider species $X_1, ..., X_U$ and reaction rates $\theta_1, ..., \theta_V$. The corresponding reaction network is as follows:

$$p_{1,1}X_1 + p_{1,2}X_2 + \cdots + p_{1,U}X_U \xrightarrow{\theta_1} q_{1,1}X_1 + q_{1,2}X_2 + \cdots + q_{1,U}X_U$$
$$\vdots$$
$$p_{V,1}X_1 + p_{V,2}X_2 + \cdots + p_{V,U}X_U \xrightarrow{\theta_V} q_{V,1}X_1 + q_{V,2}X_2 + \cdots + q_{V,U}X_U$$

for P and Q ($V \times U$) matrices.

Definition (Stoichiometric Matrix)

A stoichiometric matrix S can be defined by a $U \times V$ matrix such that $S = (Q - P)$.

Markov Jump Process

Definition (Markov Jump Process)

A Markov jump process is a stochastic process that evolves with continuous time in a discrete state space with transitions occurring at random times governed by the Markov property.

Markov Jump Process

Definition (Markov Jump Process)

A Markov jump process is a stochastic process that evolves with continuous time in a discrete state space with transitions occurring at random times governed by the Markov property.

Hazard function

For a reaction i taking place in the time interval $[t, t + dt]$, the probability of this reaction is given by

$$h_i(X_t, \theta_i)dt + \mathcal{O}(dt)$$

where the instantaneous rate/hazard can be defined as

$$h_i(X_t, \theta_i) = \theta_i \prod_{j=1}^{U} \binom{X_{j,t}}{p_{i,j}}.$$

Diffusion Approximation

Assumptions

- Large number of molecules.
- Approximate discrete jumps from the MJP process as continuous perturbations.

Diffusion Approximation

Assumptions

- Large number of molecules.
- Approximate discrete jumps from the MJP process as continuous perturbations.

Methodology

- Number of reactions i in a time step dt is distributed by $Po(h_i(X_t, \theta_i) \cdot dt)$.
- Approximation for large X_t gives $dR_t = h(X_t, \theta) \cdot dt + \text{diag}\{\sqrt{h(X_t, \theta)}\} \cdot dW_t$.
- $dW_t \sim N(0, I \cdot dt)$ for I the identity matrix $d \times d$.
- $dX_t = S \cdot dR_t$.
- $dX_t = S \cdot h(X_t, \theta) \cdot dt + \sqrt{S\text{diag}\{h(X_t, \theta)\}S'}dW_t$.

Lotka-Volterra

Let X_1 represent the prey and X_2 represent the predator.

Lotka-Volterra

Let X_1 represent the prey and X_2 represent the predator.

Lotka-Volterra Reaction Network

$$X_1 \xrightarrow{\theta_1} 2X_1, \;\; X_2 \xrightarrow{\theta_2} \emptyset,$$
$$X_1 + X_2 \xrightarrow{\theta_3} 2X_2.$$

Lotka-Volterra

Let X_1 represent the prey and X_2 represent the predator.

Lotka-Volterra Reaction Network

$$X_1 \xrightarrow{\theta_1} 2X_1, \;\; X_2 \xrightarrow{\theta_2} \emptyset,$$
$$X_1 + X_2 \xrightarrow{\theta_3} 2X_2.$$

Lotka-Volterra SDE

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \theta_1 X_1 - \theta_3 X_1 X_2 \\ \theta_3 X_1 X_2 - \theta_2 X_2 \end{bmatrix}}_{\mu(x)} dt +$$
$$\underbrace{\begin{bmatrix} \theta_1 X_1 + \theta_3 X_1 X_2 & \theta_3 X_1 X_2 \\ \theta_3 X_1 X_2 & \theta_3 X_1 X_2 + \theta_2 X_2 \end{bmatrix}^{\frac{1}{2}}}_{\sigma(x)} dW_t$$

Lotka-Volterra

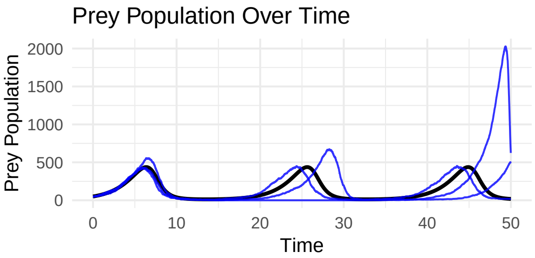
Let X_1 represent the prey and X_2 represent the predator.

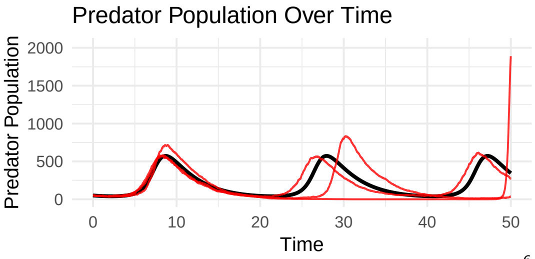
Lotka-Volterra Reaction Network

$$X_1 \xrightarrow{\theta_1} 2X_1, \;\; X_2 \xrightarrow{\theta_2} \emptyset,$$
$$X_1 + X_2 \xrightarrow{\theta_3} 2X_2.$$

Lotka-Volterra SDE

$$\begin{bmatrix} dX_1 \\ dX_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \theta_1 X_1 - \theta_3 X_1 X_2 \\ \theta_3 X_1 X_2 - \theta_2 X_2 \end{bmatrix}}_{\mu(x)} dt +$$
$$\underbrace{\begin{bmatrix} \theta_1 X_1 + \theta_3 X_1 X_2 & \theta_3 X_1 X_2 \\ \theta_3 X_1 X_2 & \theta_3 X_1 X_2 + \theta_2 X_2 \end{bmatrix}^{\frac{1}{2}}}_{\sigma(x)} dW_t$$





Inference Problem

Aim

Produce parameter estimates for θ given some partially observed data.

Inference Problem

Aim

Produce parameter estimates for θ given some partially observed data.

Definition (Euler-Maruyama approximation)

$$X_{t+1}|X_t \sim N(x_{t+1}; x_t + \mu(x_t)dt, \sigma(x_t)\sigma(x_t)'dt)$$

Inference Problem

Aim

Produce parameter estimates for θ given some partially observed data.

Definition (Euler-Maruyama approximation)

$$X_{t+1}|X_t \sim N(x_{t+1}; x_t + \mu(x_t)dt, \sigma(x_t)\sigma(x_t)'dt)$$

Discussion

- Lotka-Volterra SDE intractable.
- Euler-Maruyama approximation too crude for large time.
- Use importance sampling and further partitioning of time intervals to overcome the Euler-Maruyama inadequacies.

Inference Solution

Posterior

Given data $x_0, ..., x_N$, we can find that the posterior takes the form

$$\pi(\theta) \propto \pi_0(\theta) \prod_{t=0}^{T-1} p(x_{t+1}|x_t, \theta).$$

Inference Solution

Posterior

Given data $x_0, \ldots, x_N$, we can find that the posterior takes the form

$$\pi(\theta) \propto \pi_0(\theta) \prod_{t=0}^{T-1} p(x_{t+1}|x_t, \theta).$$

Introduce a partition over $[t, t+1]$ to be $t = t_0 < t_1 < \ldots < t_M = t+1$.

Inference Solution

Posterior

Given data $x_0, \ldots, x_N$, we can find that the posterior takes the form

$$\pi(\theta) \propto \pi_0(\theta) \prod_{t=0}^{T-1} p(x_{t+1}|x_t, \theta).$$

Introduce a partition over $[t, t+1]$ to be $t = t_0 < t_1 < \ldots < t_M = t+1$.

Transition density

$$p(x_{t+1}|x_t, \theta) \approx \mathrm{Pe}^{(M)}(x_{t+1}|x_t, \theta) = \int \prod_{i=0}^{M-1} \mathrm{Pe}(x_{t_{i+1}}|x_{t_i}, \theta) dx_{t_{1:M-1}}.$$

Importance Sampling

Durham and Gallant bridge proposal

$$q(x_{t_1}, \ldots, x_{t_{M-1}} | x_{t_0}, x_{t_M}) = \prod_{i=0}^{M-2} q(x_{t_{i+1}} | x_{t_i}, x_{t_M}).$$

Importance Sampling

Durham and Gallant bridge proposal

$$q(x_{t_1}, \ldots, x_{t_{M-1}} | x_{t_0}, x_{t_M}) = \prod_{i=0}^{M-2} q(x_{t_{i+1}} | x_{t_i}, x_{t_M}).$$

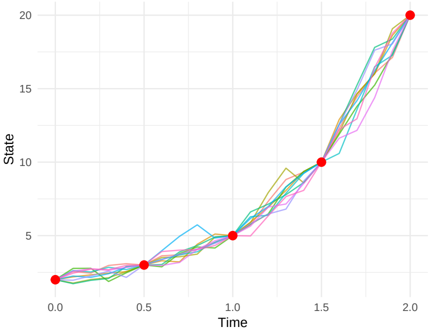


Figure: Durham and Gallant bridges between synthetic observed data points with $M = 10$.

Importance Sampling

Durham and Gallant bridge proposal

$$q(x_{t_1}, \ldots, x_{t_{M-1}} | x_{t_0}, x_{t_M}) = \prod_{i=0}^{M-2} q(x_{t_{i+1}} | x_{t_i}, x_{t_M}).$$

Estimator

$$\mathrm{Pe}^{(M)}(x_{t+1} | x_t, \theta) = \frac{1}{N} \sum_{j=1}^{N} \frac{\prod_{i=0}^{M-1} \mathrm{Pe}(x_{t_{i+1}}^{(j)} | x_{t_i}^{(j)}, \theta)}{\prod_{i=0}^{M-2} q(x_{t_{i+1}}^{(j)} | x_{t_i}^{(j)}, x_{t_M})}.$$

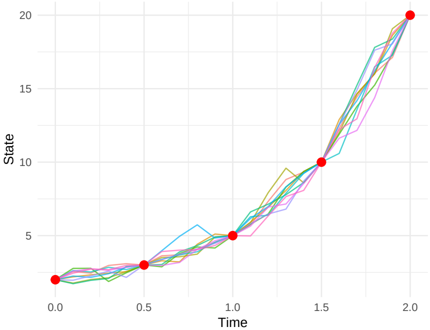


Figure: Durham and Gallant bridges between synthetic observed data points with M = 10.

Pseudo-Marginal Metropolis-Hastings

Outline

- Let u be an auxiliary variable that represents randomness with distribution $g(u)$.
- $\hat{\pi}(\theta, u)$ is the target in the Metropolis-Hastings algorithm.
- $q(\theta^*, u^*|\theta, u)$ is the proposal distribution.

Pseudo-Marginal Metropolis-Hastings

Outline

- Let u be an auxiliary variable that represents randomness with distribution $g(u)$.
- $\hat{\pi}(\theta, u)$ is the target in the Metropolis-Hastings algorithm.
- $q(\theta^*, u^*|\theta, u)$ is the proposal distribution.

Marginalisation

$$\int \hat{\pi}(\theta, u) du \propto \pi_0(\theta) \int \hat{p}(x|\theta) g(u) du \propto \pi_0(\theta) \mathbb{E}_{u \sim g}[\hat{p}(x|\theta)] \propto \pi_0(\theta) p(x|\theta) \propto \pi(\theta).$$

Pseudo-Marginal Metropolis-Hastings

Outline

- Let u be an auxiliary variable that represents randomness with distribution $g(u)$.
- $\hat{\pi}(\theta, u)$ is the target in the Metropolis-Hastings algorithm.
- $q(\theta^*, u^*|\theta, u)$ is the proposal distribution.

Marginalisation

$$\int \hat{\pi}(\theta, u)du \propto \pi_0(\theta) \int \hat{p}(x|\theta)g(u)du \propto \pi_0(\theta)\mathbb{E}_{u\sim g}[\hat{p}(x|\theta)] \propto \pi_0(\theta)p(x|\theta) \propto \pi(\theta).$$

Definition (PMMH algorithm)

Carry out standard Metropolis Hastings algorithm with target $\hat{\pi}(\theta, u)$ and proposal density

$$q(\theta^*, u^*|\theta, u) = q(\theta^*|\theta)g(u^*).$$

Further Work

Outline

- All preceding work has been done under the assumption of partially observed data values for both species.
- Now introduce the idea of only observing one species.
- Implement a proposal mechanism for the unobserved species to produce observations to be accepted or rejected along with the rate parameters.

Thank you for listening!