

Bayesian Estimation of Diffusion Process

Xiaowei Wu Hongxiao Zhu
April 2008 Rice University

Outline

Part I

- Review of Diffusion Process
- The Problem
- Review Estimation Methods of Diffusion Process
- Bayesian Modeling

Part II

- Applications to Stochastic Kinetic Models
 - Chemical Reaction System
 - Diffusion approximation
 - MCMC algorithm
 - Results

Diffusion Process--Review

- A **diffusion process** is the solution X_t to the stochastic differential equation:

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dB_t$$

- X_t is a continuous-time Markov process with continuous sample paths.
- Existence and Uniqueness of the solution. (Theorem 9.1 Steele)

Lipschitz condition: $|\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K|x - y|^2$

Linear growth condition: $|\mu(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2)$

Motivation

- Consider the SDE:

$$dY_t = \mu(Y_t; \theta)dt + \sigma(Y_t; \theta)dW_t$$

- $\mu(\cdot)$ and $\sigma(\cdot)$ are treated as known function. If θ is also known, one can solve the SDE analytically for some cases (e.g. GBM, OU), or at least one can simulate paths of Y_t recursively from SDE.
- If θ is unknown, but one has data observed from Y_t , estimating θ is necessary to understand the characteristics of the process.

Goal: Estimate diffusion parameters based on discretized samples of Y_t .

Estimation Methods of Diffusion Process--Review

- Based on discretization:

$$dY_t = \mu(Y_t; \theta)dt + \sigma(Y_t; \theta)dW_t$$



$$\Delta Y_t = \mu(Y_t; \theta)\Delta t + \sigma(Y_t; \theta)\Delta W_t$$

- Frequentist tools

Maximum Likelihood Estimation and related.

Method of Moments and related.

- We focus on Bayesian modeling.

MCMC Analysis proposed by Eraker (2001),

Application on Stochastic Kinetic models by Golightly & Wilkinson (2005) 5

Bayesian Modeling of Diffusion Process

- Consider: $dY_t = \mu(Y_t; \theta)dt + \sigma(Y_t; \theta)dW_t$

$$\mu : \mathfrak{R}^d \times \Theta \rightarrow \mathfrak{R}^d$$

$$\sigma : \mathfrak{R}^d \times \Theta \rightarrow \mathfrak{R}^{d \times d}$$

$$\text{compact } \Theta \subseteq \mathfrak{R}^k$$

- Assumption:
 1. Lipschitz and linear growth condition.
 2. $\sigma(x; \theta)\sigma(x; \theta)'$ is positive definite for all $x \in \mathfrak{R}^d$.
 3. We have a sample of observations

$$Y_{t,j}, t = 1, \dots, T, j = 1, 2, \dots, d_1$$

where $d_1 \leq d$

Bayesian Modeling of Diffusion Process

- Denote $Y_t = (X_t, Z_t)'$,
where X_t is d_1 dimensional, Z_t is d_2 dimensional, $d = d_1 + d_2$.
- Discretize Y_t on $[0, T]$: $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$

$$t_i = 0 \quad 1/m \quad \dots \quad m/m \quad (m+1)/m \quad \dots \quad n/m = T$$

$$\widehat{Y} = \begin{bmatrix} X_{1,t_0} & \widehat{X}_{1,t_1} & \dots & X_{1,t_m} & \widehat{X}_{1,t_{m+1}} & \dots & X_{1,t_n} \\ X_{2,t_0} & \widehat{X}_{2,t_1} & & X_{2,t_m} & \widehat{X}_{2,t_{m+1}} & \dots & X_{2,t_n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ X_{d_1,t_0} & \widehat{X}_{d_1,t_1} & \dots & X_{d_1,t_m} & \widehat{X}_{d_1,t_{m+1}} & \dots & X_{d_1,t_n} \\ \widehat{Z}_{1,t_0} & \widehat{Z}_{1,t_1} & \dots & \widehat{Z}_{1,t_m} & \widehat{Z}_{1,t_{m+1}} & & \widehat{Z}_{1,t_n} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ \widehat{Z}_{d_2,t_0} & \widehat{Z}_{d_2,t_1} & \dots & \widehat{Z}_{d_2,t_m} & \widehat{Z}_{d_2,t_{m+1}} & \dots & \widehat{Z}_{d_2,t_n} \end{bmatrix} \quad \Delta t = 1/m$$

Bayesian Modeling of Diffusion Process

- Discretized Model

$$\Delta Y_t = \mu(Y_t; \theta) \Delta t + \sigma(Y_t; \theta) \Delta W_t$$

$$\Delta W \sim \mathbf{N}(0, I_d \Delta t)$$

- Joint Posterior

$$\pi(\hat{Y}, \theta) \propto \prod_{i=1}^n p(\hat{Y}_i | \theta) p(\theta)$$

$|\hat{Y}_{i-1}, \theta$

$$p(\hat{Y}_i | \theta) = |\sigma_{i-1}^{-2}|^{1/2} \exp \left\{ -\frac{1}{2} \left\| (\Delta \hat{Y}_i - \mu_{i-1} \Delta t) \sigma_{i-1}^{-1} (\Delta t)^{-1/2} \right\|^2 \right\}$$

$$\Delta \hat{Y}_i := \hat{Y}_i - \hat{Y}_{i-1} \qquad \mu_i := \mu(\hat{Y}_i; \theta)$$

$$\sigma_{i-1}^{-2} := (\sigma_{i-1} \sigma'_{i-1})^{-1} \qquad \sigma_i := \sigma(\hat{Y}_i; \theta)$$

Bayesian Modeling of Diffusion Process

- MCMC: Metropolis-Hastings within Gibbs. Update θ and \hat{Y} alternatively.

- Step 0: Set initial values for θ and latent/missing data.
- Step 1: For $i=1, \dots, n$, conditional on θ , update \hat{Y}_i

$$\hat{Y}_i^{(h)} \sim \pi(\hat{Y}_i | \hat{Y}_{i-1}^{(h)}, \hat{Y}_{i+1}^{(h-1)}; \theta)$$

- Step 2: Conditional on \hat{Y} , update θ .

$$\theta^{(h+1)} \sim \pi(\theta | \hat{Y}^{(h)})$$

MCMC--step 1: update \hat{Y}_i

- Conditional posterior for missing data.

$$\pi(\hat{Y}_i | \hat{Y}_{\setminus i}, \theta) \propto p(\hat{Y}_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta)$$

$$p(\hat{Y}_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta) = |\sigma_{i-1}^{-2}|^{1/2} |\sigma_i^{-2}|^{1/2}$$

$$(7) \quad \times \exp \left\{ -\frac{1}{2} \left\| (\Delta \hat{Y}_i - \mu_{i-1} \Delta t) \sigma_{i-1}^{-1} (\Delta t)^{-1/2} \right\|^2 - \frac{1}{2} \left\| (\Delta \hat{Y}_{i+1} - \mu_i \Delta t) \sigma_i^{-1/2} (\Delta t)^{-1} \right\|^2 \right\}$$

- The unnormalized density (7) can be bimodal.

MCMC--step 1: update \hat{Y}_i

- Actually we have

$$p(\hat{Y}_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta) = N(\tilde{\mu}, \tilde{V})$$

$$\tilde{\mu} = (\sigma_{i-1}^{-2} + \sigma_i^{-2})^{-1} [\sigma_{i-1}^{-2} (\hat{Y}_{i-1} + \mu_{i-1} \Delta t) + \sigma_i^{-2} (\hat{Y}_{i+1} - \mu_i \Delta t)]$$

$$\tilde{V} = (\sigma_{i-1}^{-2} + \sigma_i^{-2})^{-1} \Delta t \quad (\sigma_{i-1}^{-2} := (\sigma_{i-1} \sigma'_{i-1})^{-1}).$$

This is a one-mode normal if both $\mu(\cdot)$ and $\sigma(\cdot)$ does not depend on Y_t ,
Otherwise, it can be of more complex form.

- The author propose Metropolis-Hastings to sample from (7) .

MCMC--step 1: update \hat{Y}_i

- If the X_i part of Y_i is observed, then:

$$p(Z_{1,i}, \dots, Z_{d_2,i} \mid \hat{Y}_{i-1}, \hat{Y}_{i+1}, X_{1,i}, X_{2,i}, \dots, X_{d_1,i}; \theta) \propto p(\hat{Y}_i \mid \hat{Y}_{i-1}, \hat{Y}_{i+1}; \theta)$$

- Metropolis-Hastings algorithm the author used:

Un-normalized posterior density : $p(\theta_i)$

Proposal density: $q(\theta_i)$

(Note: proposal density does not depend on current state of Markov Chain)

And $\text{supp}(q) \subseteq \text{supp}(p)$

Detailed manipulation:

Sample θ^* from q , and accept it with prob. $\min\{1, [p(\theta_i^*)]/[q(\theta_i^*)c]\}$

Note that this is different from MH we used in general.

MCMC--step 1: update \hat{Y}_i

- Determining the proposal density.

Proposition 1. For a scalar process, $Y_t \in \mathfrak{R}$ with constant drift and diffusion functions,

$$\hat{Y}_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta \sim N\left(\frac{1}{2}(\hat{Y}_{i-1} + \hat{Y}_{i+1}), \frac{1}{2}\sigma^2 \Delta t\right).$$

Proposition 2. For drift and diffusion functions satisfying Assumptions A1–A4,

$$\Delta t^{-1/2} \left(\hat{Y}_i - \frac{1}{2}(\hat{Y}_{i-1} + \hat{Y}_{i+1}) \right) \Rightarrow N\left(0, \frac{1}{2}\sigma_{i-1}^2\right)$$

as $\Delta t \rightarrow 0$.

- Therefore for Δt small: $\hat{Y}_i | \hat{Y}_{i-1}, \hat{Y}_{i+1} \sim N\left(\frac{1}{2}(\hat{Y}_{i-1} + \hat{Y}_{i+1}), \frac{1}{2}\sigma_{i-1}^2 \Delta t\right)$

This is the proposal density

MCMC--step 1: update \hat{Y}_i

- Determining the c, pick arbitrary m_i .

$$C = \frac{p(m_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta)}{q(m_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta)}$$
$$= p(m_i | \hat{Y}_{i-1}, \hat{Y}_{i+1}, \theta) \left| (2\pi)^d \frac{1}{2} \sigma(\hat{Y}_{i-1}) \sigma(\hat{Y}_{i-1})' \Delta t \right|^{-1/2}$$

If the p is approximately has the shape of q, the ratio of the density should be constant for all x, thus for $x = m_i$.

Side note: since we update the θ^* with prob. $\min\{1, [p(\theta_i^*)]/[q(\theta_i^*)c]\}$

This way to find C implies acceptance probability is near 1.

MCMC--step 2: update θ

- Conditional on $\widehat{Y}^{(h)}$, update

$$\theta^{(h+1)} \sim \pi(\theta \mid \widehat{Y}^{(h)})$$

- Depend on the form of θ , i.e. the form of SDE.
- Two examples given:

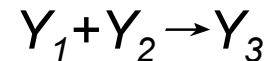
- Constant elasticity of variance (CEV) model: $dY_t = (\theta_r + \kappa_r Y_t)dt + \sigma Y_t^\beta dW_t$
- Continuous-Time stochastic volatility (SV) model:

$$dr_t = (\theta_r + \kappa_r r_t)dt + \exp\left(\frac{1}{2}Z_t\right)r_t^\beta dW_{1,t}$$
$$dZ_t = (\theta_z + \kappa_z Z_t)dt + \sigma_z dW_{2,t},$$

Applications to Stochastic Kinetic Models

Molecular Approach to Kinetics

- Bimolecular reaction example:



This reaction will occur when a molecule of Y_1 collides with a molecule of Y_2 while molecules move around randomly, driven by Brownian motion.

- Hazard h

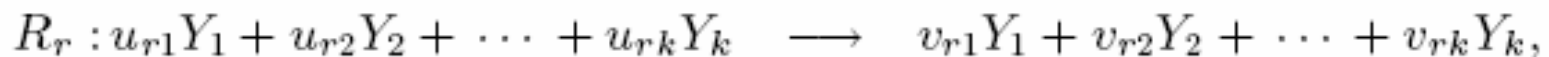
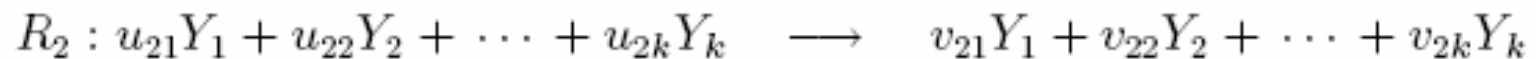
In a small, fixed volume and assuming thermal equilibrium, it has been shown that the hazard of molecules colliding is constant (Gillespie, 1992).

- Law of mass action:

h proportional to $Y_1 Y_2$

Chemical Reaction System

- Consider a system of reactions involving k species Y_1, Y_2, \dots, Y_k and r reactions R_1, R_2, \dots, R_r in thermal equilibrium inside some fixed volume V .



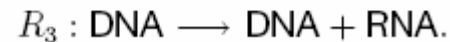
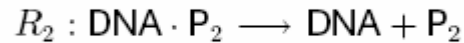
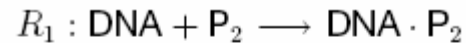
where u_{ij} is the stoichiometry associated with the j th reactant of the i th reaction and v_{ij} is the stoichiometry associated with the j th product of the i th reaction. Each reaction, R_i , has a stochastic rate constant, c_i , and a rate law or hazard, $h_i(Y, c_i)$, where $Y = (Y_1, Y_2, \dots, Y_k)$ is the current state of the system and each hazard is determined by the order of reaction R_i under an assumption of mass action kinetics.

Chemical Reaction System

- Write in a compact form: $UY \rightarrow VY$, where $U = (u_{ij})$ and $V = (v_{ij})$ are $r \times k$ dimensional matrices (obtained from the stoichiometry of the system).
- Consider change of number of molecules in reaction i and species j ,
 $a_{ij} = v_{ij} - u_{ij}$.
- The reaction network can then be represented by the net effect reaction matrix: $A = V - U$.

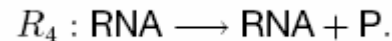
Example: Prokaryotic Autoregulatory Gene Network

- Repression, transcription reaction

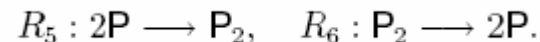


- Translation reaction (including binding ribosome to mRNA, translation of mRNA and folding of polypeptide chain)

- Dimerization



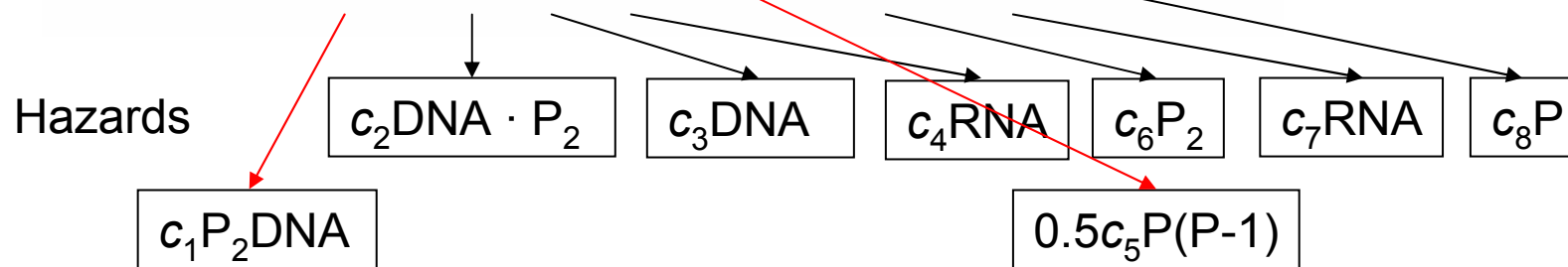
- mRNA and protein degradation



Net Effect Reaction Matrix

$$Y' = (RNA, P, P_2, DNA \cdot P_2, DNA)$$

$$A' = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$



Diffusion Approximation

- Master equation

$$P(Y; t + \Delta t) = \sum_{i=1}^r h_i(Y - A_i, c_i) P(Y - A_i; t) \Delta t$$
$$+ \left\{ 1 - \sum_{i=1}^r h_i(Y, c_i) \Delta t \right\} P(Y; t),$$
$$\frac{\partial}{\partial t} P(Y; t) = \sum_{i=1}^r \{ h_i(Y - A_i, c_i) P(Y - A_i; t)$$
$$- h_i(Y, c_i) P(Y; t) \}.$$

Expand $P(Y-A_j; t)$ using 2nd order Taylor expansion

Diffusion Approximation

- Fokker-Planck Equation

$$\frac{\partial}{\partial t} P(Y; t) = - \sum_{i=1}^k \frac{\partial}{\partial Y_i} \{ \mu_i(Y) P(Y; t) \} \\ + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \frac{\partial^2}{\partial Y_i \partial Y_j} \{ \beta_{ij}(Y) P(Y; t) \},$$

$$\mu_i(Y) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} E[\{Y_i(t + \Delta t) - Y_i(t)\} | Y(t) = Y]$$

$$\beta_{ij}(Y) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \text{Cov}[\{Y_i(t + \Delta t) - Y_i(t)\}, \\ \{Y_j(t + \Delta t) - Y_j(t)\} | Y(t) = Y].$$

- Itô diffusion

$$dY(t) = \mu(Y) dt + \beta^{1/2}(Y) dW(t),$$

$\mu(Y)$ is the column vector of $\mu_i(Y)$

$\beta^{1/2}(Y)$ is any matrix satisfying $\beta^{1/2}(\beta^{1/2})' = [\beta_{ij}(Y)] = \beta(Y)$

$W(t)$ is k -dimensional Brownian motion

Calculating Coefficients

- Suppose at time t , the state of the system is $Y(t)=(Y_1(t), \dots, Y_k(t))=Y$ so that the hazards of R_1, R_2, \dots, R_r are $h_1(Y, c_1), h_2(Y, c_2), \dots, h_r(Y, c_r)$.
- Let N_i denote the number of type- i reactions occurring in the interval $(t, t + \Delta t]$. Then for “small” time Δt , $N_i \approx \text{Poisson}(h_i(Y, c_i) \Delta t)$ and the change in the number of molecules of Y_j is given by

$$Y_j(t + \Delta t) - Y_j(t) = a_{1j}N_1 + a_{2j}N_2 + \dots + a_{rj}N_r$$

- The diffusion equation then becomes

$$dY(t) = \mu(Y, \Theta) dt + \beta^{\frac{1}{2}}(Y, \Theta) dW(t),$$

with drift and diffusion functions

$$\mu(Y, \Theta) = A'h(Y, \Theta), \quad \beta(Y, \Theta) = A'\text{diag}\{h(Y, \Theta)\}A.$$

Here, μ and β depend explicitly on Y and the parameter vector $\Theta = (c_1, c_2, \dots, c_r)$. A is the net effect matrix and $h(Y, \Theta)$ is the column vector of hazards $h_i(Y, c_i)$.

Drift $\mu (Y, \Theta)$

$$A' = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 2 & 0 & -1 \\ -1 & 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{not full rank}$$

Add row 4 to row 5, get DNA·P2 + DNA = K, remove row 4

$$\begin{pmatrix} c_3 \text{DNA} - c_7 \text{RNA} \\ c_4 \text{RNA} + 2c_6 \text{P}_2 - c_5 \text{P}(\text{P} - 1) - c_8 \text{P} \\ c_2(K - \text{DNA}) + 0.5c_5 \text{P}(\text{P} - 1) - c_1 \text{P}_2 \text{DNA} - c_6 \text{P}_2 \\ c_2(K - \text{DNA}) - c_1 \text{P}_2 \text{DNA} \end{pmatrix}$$

Diffusion $\beta(Y, \Theta) = BB'$

$$B' = \begin{pmatrix} 0 & 0 & -\sqrt{c_1 P_2 \text{DNA}} & -\sqrt{c_1 P_2 \text{DNA}} \\ 0 & 0 & \sqrt{c_2 (K - \text{DNA})} & \sqrt{c_2 (K - \text{DNA})} \\ \sqrt{c_3 \text{DNA}} & 0 & 0 & 0 \\ 0 & \sqrt{c_4 \text{RNA}} & 0 & 0 \\ 0 & -2\sqrt{0.5c_5 P(P-1)} & \sqrt{0.5c_5 P(P-1)} & 0 \\ 0 & 2\sqrt{c_6 P_2} & -\sqrt{c_6 P_2} & 0 \\ -\sqrt{c_7 \text{RNA}} & 0 & 0 & 0 \\ 0 & -\sqrt{c_8 P} & 0 & 0 \end{pmatrix}$$

Inference for Nonlinear Diffusion Models

- Discretize SDE by Euler approximation

$$\Delta Y(t) = \mu(Y(t), \Theta)\Delta t + \beta^{\frac{1}{2}}(Y(t), \Theta)\Delta W(t),$$

where $\Delta W(t)$ is a d -dimensional i.i.d. $N(0, I \Delta t)$ random vector.

- Augment with latent data points

Suppose we have measurements x^i at evenly spaced times with intervals of length Δ^* , refine grid by $\Delta t = \Delta^*/m$, where m is sufficiently large to ensure that the discretization bias associated with the Euler approximation is arbitrarily small.

Augmented data

$$\hat{Y} = \begin{pmatrix} x_1(t_0) & X_1(t_1) & \cdots & x_1(t_m) & X_1(t_{m+1}) & \cdots & x_1(t_n) \\ x_2(t_0) & X_2(t_1) & \cdots & x_2(t_m) & X_2(t_{m+1}) & \cdots & x_2(t_n) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{d_1}(t_0) & X_{d_1}(t_1) & \cdots & x_{d_1}(t_m) & X_{d_1}(t_{m+1}) & \cdots & x_{d_1}(t_n) \\ Z_1(t_0) & Z_1(t_1) & \cdots & Z_1(t_m) & Z_1(t_{m+1}) & \cdots & Z_1(t_n) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ Z_{d_2}(t_0) & Z_{d_2}(t_1) & \cdots & Z_{d_2}(t_m) & Z_{d_2}(t_{m+1}) & \cdots & Z_{d_2}(t_n) \end{pmatrix}$$

Note, \hat{Y} is $(d_1+d_2) \times (n+1)$ matrix, we have $d_1(n-(n/m+1)) + d_2(n+1)$ missing values

Bayesian Estimation

- Joint posterior density

$$\pi(\hat{Y}, \Theta) \propto \pi(\Theta)\pi(Z^0) \prod_{i=1}^n f(Y^i | Y^{i-1}, \Theta),$$

where $\pi(\Theta)$ is the prior density of Θ , $\pi(Z^0)$ is the prior density of Z^0 , and $f(Y^i | Y^{i-1}, \Theta)$

$$= |\beta_{i-1}^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\Delta Y^i - \mu_{i-1} \Delta t)' (\Delta t \beta_{i-1})^{-1} \right. \\ \left. \times (\Delta Y^i - \mu_{i-1} \Delta t) \right\}.$$

$\Delta Y^i = Y^i - Y^{i-1}$, $\mu_i = \mu(Y^i, \Theta)$, and $\beta_i = \beta(Y^i, \Theta)$. f denotes the (unnormalized) transition density obtained from the Euler discretization.

MCMC

- Metropolis-within-Gibbs
 - Initialize all unknowns. Use linear interpolation to initialize X^i and set $Z^i = 0.0$ for all i . Set $g = 1$.
 - For all $i = 0, 1, \dots, n$ at iteration g draw Y^i from its full conditional. When i is not a multiple of m , we use an MH step with proposal density $q(\cdot | Y^{i-1}, Y^{i+1}, \Theta) = N(1/2(Y^{i-1} + Y^{i+1}), 1/2 \Delta t \beta(Y^{i-1}, \Theta))$. If i is a multiple of m , only simulate the d_2 elements, Z^i , using an MH step with proposal density $q(\cdot | Y^{i-1}, Y^{i+1}, \Theta)$ further conditioned on x^i .
 - Draw $\Theta^{(g)}$ using an MH step with a Gaussian random walk update (on $\log(\Theta)$).

Simulation Study

Table 1

Posterior means and standard deviations for parameters estimated on three replicate length-50 data sets (D_1 , D_2 , and D_3) from the fully observed model with $m = 5$. The estimation results are based on the final 900,000 iterations of a single run of 1,000,000 MCMC iterations.

	c_1	c_2	c_1/c_2	c_3	c_4	c_5	c_6	c_5/c_6	c_7	c_8
	0.1	0.7	0.143	0.35	0.2	0.1	0.9	0.111	0.3	0.1
					True values					
					D_1					
Mean	0.064	0.474	0.141	0.360	0.252	0.043	0.475	0.094	0.288	0.143
SD	0.022	0.148	0.035	0.125	0.079	0.013	0.154	0.025	0.099	0.044
					D_2					
Mean	0.058	0.363	0.157	0.372	0.240	0.048	0.477	0.105	0.285	0.121
SD	0.020	0.120	0.090	0.131	0.071	0.014	0.154	0.047	0.095	0.039
					D_3					
Mean	0.052	0.346	0.153	0.416	0.213	0.044	0.488	0.092	0.321	0.115
SD	0.020	0.120	0.046	0.151	0.061	0.011	0.145	0.021	0.108	0.036

Table 2

Posterior means and standard deviations for parameters estimated using data sets D_1 , D_4 , and D_5 from the fully observed model. The estimation results are based on the final 900,000 iterations of a single run of 1,000,000 MCMC iterations.

m		c_1	c_2	c_1/c_2	c_3	c_4	c_5	c_6	c_5/c_6	c_7	c_8
		0.1	0.7	0.143	0.35	0.2	0.1	0.9	0.111	0.3	0.1
		True values									
		50 observations									
2	Mean	0.049	0.370	0.137	0.333	0.235	0.030	0.308	0.100	0.269	0.135
	SD	0.016	0.110	0.039	0.116	0.079	0.008	0.084	0.031	0.092	0.044
5	Mean	0.066	0.475	0.140	0.361	0.253	0.042	0.468	0.093	0.286	0.143
	SD	0.022	0.150	0.032	0.124	0.079	0.012	0.150	0.018	0.095	0.044
8	Mean	0.074	0.524	0.142	0.373	0.258	0.053	0.630	0.087	0.295	0.143
	SD	0.027	0.175	0.027	0.122	0.075	0.017	0.226	0.014	0.093	0.041
10	Mean	0.076	0.531	0.143	0.403	0.265	0.060	0.741	0.084	0.316	0.146
	SD	0.025	0.165	0.027	0.141	0.076	0.019	0.273	0.013	0.105	0.041
		100 observations									
2	Mean	0.103	0.661	0.157	0.285	0.240	0.051	0.571	0.090	0.224	0.105
	SD	0.024	0.142	0.028	0.082	0.061	0.010	0.126	0.015	0.055	0.029
5	Mean	0.096	0.663	0.147	0.286	0.246	0.057	0.593	0.097	0.228	0.110
	SD	0.018	0.119	0.027	0.054	0.055	0.013	0.151	0.013	0.048	0.025
8	Mean	0.101	0.687	0.148	0.295	0.250	0.076	0.856	0.091	0.235	0.110
	SD	0.020	0.132	0.021	0.066	0.051	0.018	0.233	0.010	0.046	0.024
10	Mean	0.102	0.691	0.149	0.296	0.257	0.096	0.967	0.086	0.236	0.110
	SD	0.020	0.134	0.021	0.066	0.052	0.023	0.235	0.009	0.047	0.023
		500 observations									
2	Mean	0.092	0.597	0.155	0.327	0.214	0.101	0.925	0.110	0.222	0.091
	SD	0.010	0.062	0.022	0.041	0.026	0.009	0.082	0.011	0.031	0.016
5	Mean	0.098	0.622	0.158	0.331	0.213	0.113	1.028	0.110	0.226	0.092
	SD	0.010	0.063	0.021	0.039	0.025	0.010	0.093	0.010	0.029	0.015
8	Mean	0.113	0.824	0.138	0.330	0.216	0.144	1.230	0.114	0.225	0.094
	SD	0.013	0.077	0.016	0.040	0.025	0.013	0.126	0.009	0.030	0.016
10	Mean	0.110	0.773	0.143	0.330	0.214	0.137	1.180	0.112	0.226	0.093
	SD	0.012	0.073	0.017	0.038	0.024	0.012	0.114	0.009	0.032	0.016

Table 3

Posterior means and standard deviations for parameters estimated using data sets D_1 , D_4 , and D_5 from the partially observed model. Discretization is set at $m = 5$ and the estimation results are based on the final 6,000,000 iterations of a single run of 10,000,000 MCMC iterations.

	c_1	c_2	c_1/c_2	c_3	c_4	c_5	c_6	c_5/c_6	c_7	c_8
	0.1	0.7	0.143	0.35	0.2	0.1	0.9	0.111	0.3	0.1
	True values									
	50 observations									
Mean	0.049	0.442	0.116	0.310	0.012	0.062	0.603	0.103	0.265	0.011
SD	0.015	0.131	0.033	0.080	0.023	0.018	0.183	0.013	0.062	0.014
	100 observations									
Mean	0.077	0.941	0.090	0.255	0.270	0.097	0.761	0.120	0.280	0.125
SD	0.020	0.253	0.022	0.050	0.122	0.027	0.214	0.012	0.048	0.061
	500 observations									
Mean	0.105	0.574	0.180	0.370	0.187	0.112	1.021	0.110	0.218	0.107
SD	0.016	0.076	0.049	0.062	0.073	0.009	0.084	0.008	0.024	0.041

References

- Eraker, B (2001). MCMC analysis of diffusion models with applications to finance. *Journal of Business and Economic Statistics*. **19**, 177-191.
- Golightly, A and Wilkinson D.J (2005). Bayesian inference for stochastic kinetic models using a diffusion approximation. *Biometrics*. **61**, 781-788.