Bayesian inference for stochastic differential mixed effects models - initial steps

Gavin Whitaker

2nd May 2012

Supervisors: RJB and AG

Outline

- Mixed Effects Stochastic Differential Equations (SDEs)
- Bayesian inference for SDEs
- Toy model (Roberts and Stramer 2001 paper)

SDE Models

- Consider an Itô process $\{\boldsymbol{X}_t, t \geq 0\}$ satisfying

$$d\boldsymbol{X}_t = \boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})dt + \sqrt{\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})}d\boldsymbol{W}_t$$

  - $\boldsymbol{X}_t$ is the value of the process at time t
  - $\boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the drift
  - $\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the diffusion coefficient
  - $\boldsymbol{W}_t$ is standard Brownian motion
  - $\boldsymbol{X}_0$ is the vector of initial conditions

SDE Models

- Consider an Itô process $\{\boldsymbol{X}_t, t \geq 0\}$ satisfying

$$d\boldsymbol{X}_t = \boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})dt + \sqrt{\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})}d\boldsymbol{W}_t$$

  - $\boldsymbol{X}_t$ is the value of the process at time t
  - $\boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the drift
  - $\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the diffusion coefficient
  - $\boldsymbol{W}_t$ is standard Brownian motion
  - $\boldsymbol{X}_0$ is the vector of initial conditions

- Seek a numerical solution via (for example) the Euler-Maruyama approximation

$$\Delta\boldsymbol{X}_t \equiv \boldsymbol{X}_{t+\Delta t} - \boldsymbol{X}_t = \boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})\Delta t + \sqrt{\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})}\Delta\boldsymbol{W}_t$$

where $\Delta\boldsymbol{W}_t \sim N(\boldsymbol{0}, \boldsymbol{I}\Delta t)$

CIR Model

$$dX_t = (\theta_1 - \theta_2 X_t)dt + \theta_3 \sqrt{X_t} dW_t$$

- Used to model short term interest rates
- The process is mean reverting

CIR Model

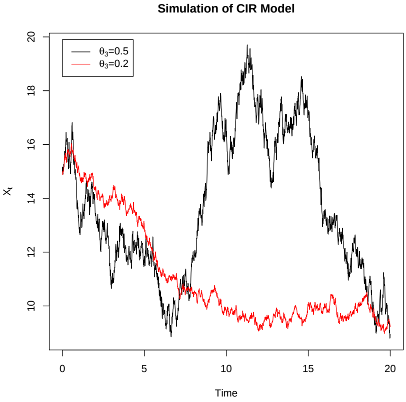

Figure: Numerical solution for CIR model, $\theta_1 = 1$, $\theta_2 = 0.1$, $X_0 = 15$

Aphid Growth Model

- Also known as plant lice, or greenfly
- They are small sap sucking insects
- Some species of ants "farm" aphids, for the honeydew they release. These "dairying ants", "milk" the aphids by stroking them

Aphid Growth Model

$$\begin{pmatrix} dN_t \\ dC_t \end{pmatrix} = \begin{pmatrix} \lambda N_t - \mu N_t C_t \\ \lambda N_t \end{pmatrix} dt + \begin{pmatrix} \lambda N_t + \mu N_t C_t & \lambda N_t \\ \lambda N_t & \lambda N_t \end{pmatrix}^{\frac{1}{2}} d\boldsymbol{W}_t$$

- N_t is the aphid population size at time t
- C_t is the cumulative population at time t
- This model is an SDE approximation to an underlying stochastic kinetic model
- Birth rate of λN_t and a death rate of $\mu N_t C_t$

Aphid Growth Model

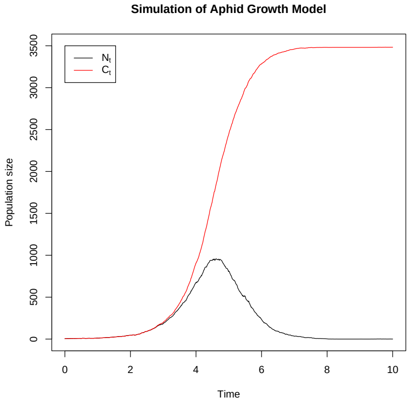

Figure: Numerical solution for Aphid model, $\lambda = 1.75$, $\mu = 0.001$

Mixed Effects SDE Models

- What if experimental units are not identical?
- Suppose the units have common parameters $\boldsymbol{\theta}$ but different parameters $\boldsymbol{b}^i$
- We treat the $\boldsymbol{b}^i$ as random effects with a population profile

Mixed Effects SDE Models

- What if experimental units are not identical?
- Suppose the units have common parameters $\boldsymbol{\theta}$ but different parameters $\boldsymbol{b}^i$
- We treat the $\boldsymbol{b}^i$ as random effects with a population profile
- This gives us a stochastic differential mixed-effects model for the experimental units:

$$d\boldsymbol{X}_t^i = \boldsymbol{\alpha}(\boldsymbol{X}_t^i, \boldsymbol{\theta}, \boldsymbol{b}^i)dt + \sqrt{\boldsymbol{\beta}(\boldsymbol{X}_t^i, \boldsymbol{\theta}, \boldsymbol{b}^i)}d\boldsymbol{W}_t^i, \quad i = 1, \ldots, M$$

- Differences between units are down to different realisations of the Brownian motion paths $\boldsymbol{W}_t^i$ and the random effects $\boldsymbol{b}^i$
- Allows us to split the total variation between within- and between-individual components

Bayesian inference for SDEs

- Problematic due to the intractability of the transition density characterising the process
- In other words, we typically can't solve an SDE analytically
- So we could just work with the Euler approximation
- Given data $\boldsymbol{d}$ at equidistant times $t_0, t_1, \ldots, t_n$, the Euler approximation may be unsatisfactory for $\Delta t = t_{i+1} - t_i$
- We therefore adopt a data augmentation scheme

Bayesian inference for SDEs

- Introduce a partition of $[t_i, t_{i+1}]$ as

$$t_i = \tau_{im} < \tau_{im+1} < \ldots < \tau_{(i+1)m} = t_{i+1}$$

where

$$\Delta\tau \equiv \tau_{im+1} - \tau_{im} = \frac{t_{i+1} - t_i}{m}$$

- Apply Euler approximation over each interval of width $\Delta\tau$
- Introduces $m-1$ latent values between every pair of observations

- Formulate joint posterior for parameters and latent values

$$
\begin{aligned}
\rightarrow \quad \boldsymbol{d} &= (\boldsymbol{x}_{t_0}, \boldsymbol{x}_{t_1}, \ldots, \boldsymbol{x}_{t_n}) \\
\rightarrow \quad \boldsymbol{x} &= (\boldsymbol{x}_{\tau_1}, \boldsymbol{x}_{\tau_2}, \ldots, \boldsymbol{x}_{\tau_{m-1}}, \boldsymbol{x}_{\tau_{m+1}}, \ldots, \ldots, \boldsymbol{x}_{\tau_{nm-1}}) \\
&= \text{latent path} \\
\rightarrow \quad (\boldsymbol{x}, \boldsymbol{d}) &= (\boldsymbol{x}_{\tau_0}, \boldsymbol{x}_{\tau_1}, \ldots, \boldsymbol{x}_{\tau_m}, \boldsymbol{x}_{\tau_{m+1}}, \ldots, \ldots, \boldsymbol{x}_{\tau_{nm}}) \\
&= \text{augmented path}
\end{aligned}
$$

Bayesian inference for SDEs

Formulate joint posterior for parameters and latent data as

$$\begin{aligned}\pi(\boldsymbol{\theta},\boldsymbol{x}|\boldsymbol{d}) &\propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{x},\boldsymbol{d}|\boldsymbol{\theta}) \\ &\propto \pi(\boldsymbol{\theta})\prod_{i=0}^{nm-1}\pi(\boldsymbol{x}_{\tau_{i+1}}|\boldsymbol{x}_{\tau_i},\boldsymbol{\theta})\end{aligned}$$

where

$$\pi(\boldsymbol{x}_{\tau_{i+1}}|\boldsymbol{x}_{\tau_i},\boldsymbol{\theta}) = \phi\left(\boldsymbol{x}_{\tau_{i+1}}\,;\,\boldsymbol{x}_{\tau_i} + \boldsymbol{\alpha}(\boldsymbol{x}_{\tau_i}\,,\,\boldsymbol{\theta})\Delta t, \boldsymbol{\beta}(\boldsymbol{x}_{\tau_i},\boldsymbol{\theta})\Delta t\right)$$

and $\phi(\cdot\,;\,\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes the Gaussian density with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$

Formulate joint posterior for parameters and latent data as

$$\begin{aligned}\pi(\boldsymbol{\theta},\boldsymbol{x}|\boldsymbol{d}) &\propto \pi(\boldsymbol{\theta})\pi(\boldsymbol{x},\boldsymbol{d}|\boldsymbol{\theta}) \\ &\propto \pi(\boldsymbol{\theta})\prod_{i=0}^{nm-1}\pi(\boldsymbol{x}_{\tau_{i+1}}|\boldsymbol{x}_{\tau_i},\boldsymbol{\theta})\end{aligned}$$

where

$$\pi(\boldsymbol{x}_{\tau_{i+1}}|\boldsymbol{x}_{\tau_i},\boldsymbol{\theta}) = \phi\left(\boldsymbol{x}_{\tau_{i+1}}\,;\,\boldsymbol{x}_{\tau_i}+\boldsymbol{\alpha}(\boldsymbol{x}_{\tau_i}\,,\,\boldsymbol{\theta})\Delta t,\boldsymbol{\beta}(\boldsymbol{x}_{\tau_i},\boldsymbol{\theta})\Delta t\right)$$

and $\phi(\cdot\,;\,\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes the Gaussian density with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$

- ▷ The posterior distribution is typically analytically intractable

A Gibbs sampling approach

- We therefore sample via an MCMC scheme
- E.g a Gibbs sampler, alternating between draws of
  - $\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{d}$
  - $\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{d}$

A Gibbs sampling approach

- We therefore sample via an MCMC scheme
- E.g a Gibbs sampler, alternating between draws of
  - $\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{d}$
  - $\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{d}$
- The last step can be done (for example) in blocks of length $m-1$ between observations
- Metropolis within Gibbs updates may be needed
- Problem: the mixing is poor for large m

Toy model

Consider the SDE

$$dX_t = \frac{1}{\sqrt{\theta}} dW_t$$

- Suppose that we have observations $X_0 = x_0 = 0$ and $X_1 = x_1$
- Set $\tau_i = i/m$ for $i = 0, 1, \ldots, m$ so that

$$(x, d) = x_0, x_{\frac{1}{m}}, x_{\frac{2}{m}}, \ldots, x_{\frac{m-1}{m}}, x_1$$

Toy model

Consider the SDE

$$d\boldsymbol{X}_t = \frac{1}{\sqrt{\theta}} d\boldsymbol{W}_t$$

- Suppose that we have observations $X_0 = x_0 = 0$ and $X_1 = x_1$
- Set $\tau_i = i/m$ for $i = 0, 1, \ldots, m$ so that

$$(\boldsymbol{x}, \boldsymbol{d}) = \underset{\text{obs}}{\underset{\uparrow}{x_0}}, \underbrace{x_{\frac{1}{m}}, x_{\frac{2}{m}}, \ldots, x_{\frac{m-1}{m}}}_{\text{path (bridge)}}, \underset{\text{obs}}{\underset{\uparrow}{x_1}}$$

Toy model

Consider the SDE

$$dX_t = \frac{1}{\sqrt{\theta}} dW_t$$

- Suppose that we have observations $X_0 = x_0 = 0$ and $X_1 = x_1$
- Set $\tau_i = i/m$ for $i = 0, 1, \ldots, m$ so that

$$(\boldsymbol{x}, \boldsymbol{d}) = \underset{\substack{\uparrow \\ \text{obs}}}{x_0}, \underbrace{x_{\frac{1}{m}}, x_{\frac{2}{m}}, \ldots, x_{\frac{m-1}{m}}}_{\text{path (bridge)}}, \underset{\substack{\uparrow \\ \text{obs}}}{x_1}$$

- Under the Euler approximation

$$X_{\frac{i}{m}} | x_{\frac{(i-1)}{m}}, \theta \sim N\left(x_{\frac{(i-1)}{m}}, \frac{1}{m\theta}\right)$$

Toy model

- Hence

$$\begin{aligned}\pi(\boldsymbol{x},\boldsymbol{d}|\theta) &\propto \prod_{i=1}^{m}\frac{\sqrt{m\theta}}{\sqrt{2\pi}}\exp\left\{-\frac{m\theta\left(x_{\frac{i}{m}}-x_{\frac{(i-1)}{m}}\right)^2}{2}\right\}\\ &\propto \theta^{m/2}\exp\left\{-\frac{1}{2}m\theta\sum_{i=1}^{m}\left(x_{\frac{i}{m}}-x_{\frac{(i-1)}{m}}\right)^2\right\}\end{aligned}$$

Toy model

- Hence

$$\begin{aligned}\pi(\boldsymbol{x},\boldsymbol{d}|\theta) &\propto \prod_{i=1}^{m} \frac{\sqrt{m\theta}}{\sqrt{2\pi}} \exp\left\{-\frac{m\theta\left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}}\right)^2}{2}\right\} \\ &\propto \theta^{m/2} \exp\left\{-\frac{1}{2}m\theta \sum_{i=1}^{m}\left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}}\right)^2\right\}\end{aligned}$$

- Take prior $\theta \sim \text{Exp}(1)$
- The full conditional for θ is

$$\pi(\theta|\boldsymbol{x},\boldsymbol{d}) \propto \pi(\theta)\pi(\boldsymbol{x},\boldsymbol{d}|\theta)$$

Toy model

$$
\begin{aligned}
&\propto \quad \theta^{m/2} \exp\left\{-\frac{1}{2} m\theta \sum_{i=1}^{m}\left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}}\right)^2 - \theta\right\} \\
&\propto \quad \theta^{m/2} \exp\left\{-\theta\left(\frac{m\Sigma_X}{2} + 1\right)\right\}
\end{aligned}
$$

where

$$\Sigma_X = \sum_{i=1}^{m}\left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}}\right)^2$$

Toy model

$$\begin{aligned}
&\propto \quad \theta^{m/2} \exp\left\{ -\frac{1}{2} m\theta \sum_{i=1}^{m} \left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}} \right)^2 - \theta \right\} \\
&\propto \quad \theta^{m/2} \exp\left\{ -\theta \left(\frac{m\Sigma_X}{2} + 1 \right) \right\}
\end{aligned}$$

where

$$\Sigma_X = \sum_{i=1}^{m} \left(x_{\frac{i}{m}} - x_{\frac{(i-1)}{m}} \right)^2$$

- Therefore

$$\theta | \boldsymbol{x}, \boldsymbol{d} \sim \Gamma\left(\frac{m}{2} + 1, \frac{m\Sigma_X}{2} + 1 \right)$$

Toy model

- Under the linear Gaussian structure of the simple SDE, the full conditional $\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{d}$ can be sampled using

$$X_{\frac{i}{m}}|x_1,\theta = \frac{ix_1}{m} + \frac{1}{\sqrt{\theta}} Z_{\frac{i}{m}}, \quad i = 1, 2, ..., m$$

where $\{Z_t, 0 \leq t \leq 1\}$ is a standard Brownian bidge, that is a standard Brownian motion conditioned to hit 0 at time 0, at time 1

Toy model

Simulated data

- Take $x_0 = 0$, $\theta = 1$
- Simulate x_1 using Euler scheme
- Get $x_1 = -0.6947 \Rightarrow \boldsymbol{d} = (0, -0.6947)$

Toy model

Simulated data

- Take $x_0 = 0$, $\theta = 1$
- Simulate x_1 using Euler scheme
- Get $x_1 = -0.6947 \Rightarrow \boldsymbol{d} = (0, -0.6947)$

MCMC scheme

We perform a run of 1000 iterations of the scheme with no thin. Initialise with $\theta^{(0)} = 1$, the prior mean

Step 1 Update the discretised Brownian bridge which hits x_1 at $t = 1$

Step 2 Draw θ from its full conditional distribution,

$$\theta | \boldsymbol{x}, \boldsymbol{d} \sim \Gamma\left(\frac{m}{2} + 1, \frac{m\Sigma_X}{2} + 1\right)$$

Toy model – Results

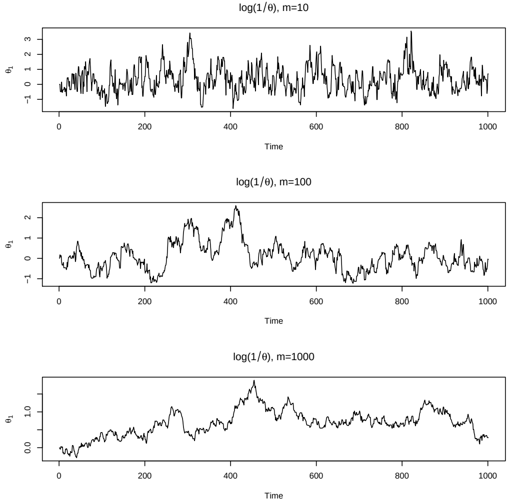

Figure: Trace plots for $\log(\frac{1}{\theta})$

Toy model – Results

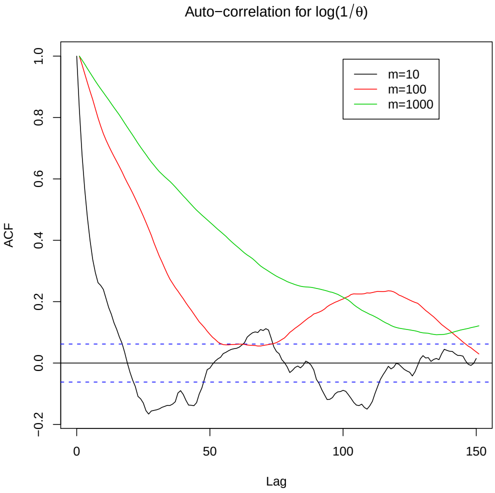

Figure: Auto-correlation plots for $\log(\frac{1}{\theta})$

Toy model – Results

- Mixing gets even worse for larger m
- Why is this happening?
- Try and quantify the mixing time by considering the parameter update
- The parameter update can be rewritten in a clever way ...

Toy model – What's going wrong?

- Recall that

$$X_{\frac{i}{m}} | x_1, \theta = \frac{i x_1}{m} + \frac{1}{\sqrt{\theta}} Z_{\frac{i}{m}}, \quad i = 1, 2, ..., m$$

Toy model – What's going wrong?

- Recall that

$$X_{\frac{i}{m}}|x_1,\theta = \frac{ix_1}{m} + \frac{1}{\sqrt{\theta}} Z_{\frac{i}{m}}, \quad i = 1, 2, ..., m$$

- Now

$$\begin{aligned}
\Sigma_X &= \sum_{i=1}^{m} \left(\frac{i}{m} x_1 + \frac{1}{\sqrt{\theta}} Z_{\frac{i}{m}} - \left[\frac{(i-1)}{m} x_1 + \frac{1}{\sqrt{\theta}} Z_{\frac{(i-1)}{m}} \right] \right)^2 \\
&= \sum_{i=1}^{m} \left(\frac{1}{\sqrt{\theta}} \left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}} \right] + \frac{i}{m} x_1 - \frac{(i-1)}{m} x_1 \right)^2 \\
&= \sum_{i=1}^{m} \left(\frac{1}{\sqrt{\theta}} \left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}} \right] + \frac{x_1}{m} \right)^2
\end{aligned}$$

- Expanding out

$$
\begin{aligned}
\Sigma_X &= \sum_{i=1}^{m}\left(\frac{1}{\theta}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]^2 + \frac{x_1^2}{m^2} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]\right) \\
&= \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\sum_{i=1}^{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]
\end{aligned}
$$

- Now

$$
\begin{aligned}
\sum_{i=1}^{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right] &= \left(Z_{\frac{1}{m}} - Z_0\right) + \left(Z_{\frac{2}{m}} - Z_{\frac{1}{m}}\right) + \ldots \\
&\quad \ldots + \left(Z_{\frac{(m-1)}{m}} - Z_{\frac{(m-2)}{m}}\right) + \left(Z_1 - Z_{\frac{(m-1)}{m}}\right)
\end{aligned}
$$

- Expanding out

$$\Sigma_X = \sum_{i=1}^{m}\left(\frac{1}{\theta}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]^2 + \frac{x_1^2}{m^2} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]\right)$$

$$= \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\sum_{i=1}^{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right]$$

- Now

$$\sum_{i=1}^{m}\left[Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right] = \left(\cancel{Z_{\frac{1}{m}}} - Z_0\right) + \left(\cancel{Z_{\frac{2}{m}}} - \cancel{Z_{\frac{1}{m}}}\right) + \cancel{\cdots}$$

$$\cancel{\cdots} + \left(\cancel{Z_{\frac{(m-1)}{m}}} - \cancel{Z_{\frac{(m-2)}{m}}}\right) + \left(Z_1 - \cancel{Z_{\frac{(m-1)}{m}}}\right)$$

Toy model – What's going wrong?

- So
$$\Sigma_X = \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\left[Z_1 - Z_0\right]$$
But $Z_0 = Z_1 = 0$ since Z is a standard Brownian bridge

Toy model – What's going wrong?

- So

$$\Sigma_X = \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\left[\cancelto{0}{Z_1 - Z_0}\right]$$

But $Z_0 = Z_1 = 0$ since Z is a standard Brownian bridge

- Thus

$$\Sigma_X = \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m}$$

where

$$\Sigma_Z = \sum_{i=1}^{m}\left(Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right)^2$$

Toy model – What's going wrong?

- So
$$\Sigma_X = \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m} + \frac{2}{\sqrt{\theta}}\frac{x_1}{m}\left[\cancelto{0}{Z_1 - Z_0}\right]$$
But $Z_0 = Z_1 = 0$ since Z is a standard Brownian bridge
- Thus
$$\Sigma_X = \frac{1}{\theta}\Sigma_Z + \frac{x_1^2}{m}$$
where
$$\Sigma_Z = \sum_{i=1}^{m}\left(Z_{\frac{i}{m}} - Z_{\frac{(i-1)}{m}}\right)^2$$
- Using properties of Gaussian random variables we have
$$\Sigma_Z \sim \frac{\chi^2_{m-1}}{m}$$

Toy model – What's going wrong?

- Now

$$\theta|\boldsymbol{x},\boldsymbol{d} \sim \Gamma\left(\frac{m}{2}+1, \frac{m\Sigma_X}{2}+1\right)$$

- If $H \sim \Gamma\left(\frac{m}{2}+1, 1\right)$ then

$$\begin{aligned} \theta_{\text{new}} &= \frac{H}{\frac{m\Sigma_X}{2}+1} \\ &= \frac{H}{\frac{m}{2}\left(\frac{x_1^2}{m}+\frac{\chi^2_{m-1}}{m\theta_{\text{old}}}\right)+1} \end{aligned}$$

Toy model – What's going wrong?

- Now

$$\theta|\boldsymbol{x},\boldsymbol{d} \sim \Gamma\left(\frac{m}{2}+1, \frac{m\Sigma_X}{2}+1\right)$$

- If $H \sim \Gamma\left(\frac{m}{2}+1, 1\right)$ then

$$\begin{aligned}\theta_{\text{new}} &= \frac{H}{\frac{m\Sigma_X}{2}+1} \\ &= \frac{H}{\frac{x_1^2}{2}+\frac{\chi^2_{m-1}}{2\theta_{\text{old}}}+1}\end{aligned}$$

Toy model – What's going wrong?

- For large m, approximate H and χ^2_{m-1} with Normal random variables
- Roberts and Stramer then use a suitable Taylor expansion of the expression for θ_{new} to give

$$\theta_{\text{new}} \approx \theta_{\text{old}} \left\{ 1 + \left(\frac{2}{m}\right)^{\frac{1}{2}} (W_1 - W_2) - \frac{2 + x_1^2}{m}\theta_{\text{old}} + \frac{W_2^2}{m} \right\}$$

where W_1 and W_2 are independent $N(0,1)$ random variables

Toy model – What's going wrong?

- For large m, approximate H and χ^2_{m-1} with Normal random variables
- Roberts and Stramer then use a suitable Taylor expansion of the expression for θ_{new} to give

$$\theta_{\text{new}} \approx \theta_{\text{old}} \left\{ 1 + \left(\frac{2}{m}\right)^{\frac{1}{2}} (W_1 - W_2) - \frac{2 + x_1^2}{m}\theta_{\text{old}} + \frac{W_2^2}{m} \right\}$$

  where W_1 and W_2 are independent $N(0,1)$ random variables
- $\theta_{\text{new}} = \theta_{\text{old}} \left\{ 1 + O(m^{-1}) \right\}$

Toy model – What's going wrong?

- For large m, approximate H and χ^2_{m-1} with Normal random variables
- Roberts and Stramer then use a suitable Taylor expansion of the expression for θ_{new} to give

$$\theta_{\text{new}} \approx \theta_{\text{old}} \left\{ 1 + \left(\frac{2}{m} \right)^{\frac{1}{2}} (W_1 - W_2) - \frac{2 + x_1^2}{m} \theta_{\text{old}} + \frac{W_2^2}{m} \right\}$$

  where W_1 and W_2 are independent $N(0,1)$ random variables
- $\theta_{\text{new}} = \theta_{\text{old}} \left\{ 1 + O(m^{-1}) \right\}$

$\Rightarrow$ Mixing time is $O(m)$

Future work

- ▷ Construct MCMC schemes for arbitrary nonlinear diffusion processes
  - Naive schemes with a block update for the path
  - "Better" schemes that use a reparameterisation
  - Joint update of path and parameters (pMCMC)
- ▷ Application to mixed effects SDEs
  - Aphid model, real data examples

References

- Roberts. G. O. and Stramer. O.
  On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm.
  Biometrika, 88 (3) 603-621, 2001
- Gillespie, C. S. and Golightly, A.
  Bayesian inference for generalized stochastic population growth models with application to aphids.
  JRSS Series C, Applied Statistics, 59(2):341-357, 2010