Inference for stochastic differential random effects models

Gavin Whitaker

4th December 2013

Supervisors: RJB and AG

SDE models

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

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

  - $\boldsymbol{\alpha}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the drift
  - $\boldsymbol{\beta}(\boldsymbol{X}_t, \boldsymbol{\theta})$ is the diffusion coefficient
  - $\boldsymbol{W}_t$ is standard Brownian motion

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

Random effects SDE models

- Consider the case where we have ℓ subjects and that each individual can be represented by the same SDE
- Common parameters $\boldsymbol{\theta}$
- Different parameters $\boldsymbol{\phi}^{(i)}$, $i = 1, \dots, \ell$
- This gives us a stochastic differential random effects model:

$$dX_t^{(i)} = \boldsymbol{\alpha}\left(\boldsymbol{X}_t^{(i)}, \boldsymbol{\theta}, \boldsymbol{\phi}^{(i)}\right) dt + \sqrt{\boldsymbol{\beta}\left(\boldsymbol{X}_t^{(i)}, \boldsymbol{\theta}, \boldsymbol{\phi}^{(i)}\right)} d\boldsymbol{W}_t^{(i)}$$

  for $i = 1, \dots, \ell$

- Suppose we have data available at times $t_0^{(i)}, t_1^{(i)}, \dots, t_{n_i}^{(i)}$ for each individual i
- Note n_i may be different for each individual
- We implement a data augmentation approach

- Consider $\left[t_j^{(i)}, t_{j+1}^{(i)}\right]$ and introduce a partition

$$t_j^{(i)} = \tau_{j,0}^{(i)} < \underbrace{\tau_{j,1}^{(i)} < \ldots < \tau_{j,m_j^{(i)}-1}^{(i)}}_{\text{latent times}} < \tau_{j,m_j^{(i)}}^{(i)} = t_{j+1}^{(i)}$$

- Time step between observations

$$\Delta_{t_j}^{(i)} = \frac{t_{j+1}^{(i)} - t_j^{(i)}}{m_j^{(i)}}$$

- Allows for irregularly spaced data for each individual

- Formulate joint posterior for parameters and latent values
- For individual i

$$\boldsymbol{d}^{(i)} = \left(\boldsymbol{x}_{t_0}^{(i)}, \boldsymbol{x}_{t_1}^{(i)}, \ldots, \boldsymbol{x}_{t_{n_i}}^{(i)}\right)$$

$$\boldsymbol{x}^{(i)} = \left(\boldsymbol{x}_{\tau_{0,1}}^{(i)}, \boldsymbol{x}_{\tau_{0,2}}^{(i)}, \ldots, \boldsymbol{x}_{\tau_{0,m_0^{(i)}-1}}^{(i)}, \boldsymbol{x}_{\tau_{1,1}}^{(i)}, \ldots, \boldsymbol{x}_{\tau_{n_i-1,m_{n-1}^{(i)}-1}}^{(i)}\right)$$

- $\boldsymbol{x}^{(i)}$ is the values of the skeleton path at times $\tau_{0,1}^{(i)}, \tau_{0,2}^{(i)}$ etc
- Putting these together for ℓ individuals

$$\boldsymbol{d} = \left(\boldsymbol{d}^{(1)}, \boldsymbol{d}^{(2)}, \ldots, \boldsymbol{d}^{(\ell)}\right)$$

$$\boldsymbol{x} = \left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(\ell)}\right)$$

Formulate joint posterior for parameters and latent data as

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

$$\propto \underbrace{\pi(\boldsymbol{\theta})\pi(\boldsymbol{\phi}|\boldsymbol{\theta})}_{\text{prior}} \times \prod_{i=1}^{\ell}\prod_{j=0}^{n_i-1}\prod_{k=0}^{m_j^{(i)}-1} \underbrace{\pi\left(\boldsymbol{x}^{(i)}_{\tau_{j,(k+1)}}\middle|\boldsymbol{x}^{(i)}_{\tau_{j,k}},\boldsymbol{\theta},\boldsymbol{\phi}^{(i)}\right)}_{\text{Euler density}}$$

where

$$\pi\left(\boldsymbol{x}^{(i)}_{\tau_{j,(k+1)}}\middle|\boldsymbol{x}^{(i)}_{\tau_{j,k}},\boldsymbol{\theta},\boldsymbol{\phi}^{(i)}\right) = \phi\left(\boldsymbol{x}^{(i)}_{\tau_{j,(k+1)}}\middle|\boldsymbol{x}^{(i)}_{\tau_{j,k}} + \boldsymbol{\alpha}^{(i)}_{j,k}\Delta^{(i)}_{t_j},\boldsymbol{\beta}^{(i)}_{j,k}\Delta^{(i)}_{t_j}\right)$$

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

Note $\boldsymbol{\alpha}^{(i)}_{j,k} = \boldsymbol{\alpha}\left(\boldsymbol{x}^{(i)}_{\tau_{j,k}},\boldsymbol{\theta},\boldsymbol{\phi}^{(i)}\right)$ and $\boldsymbol{\beta}^{(i)}_{j,k} = \boldsymbol{\beta}\left(\boldsymbol{x}^{(i)}_{\tau_{j,k}},\boldsymbol{\theta},\boldsymbol{\phi}^{(i)}\right)$

- The posterior distribution is typically analytically intractable
- Use a Gibbs sampler, alternating between draws of
  - $\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{\phi}$
  - $\boldsymbol{\phi}|\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{\theta}$
  - $\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{\phi}$
- We can update $\boldsymbol{\phi}$ componentwise, for each individual

$$\pi(\boldsymbol{\phi}|\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{\theta}) = \prod_{i=1}^{\ell} \pi\left(\boldsymbol{\phi}^{(i)}|\boldsymbol{x}^{(i)}, \boldsymbol{d}^{(i)}, \boldsymbol{\theta}\right)$$

- Similarly

$$\pi(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{\phi}) = \prod_{i=1}^{\ell} \pi\left(\boldsymbol{x}^{(i)}|\boldsymbol{\theta}, \boldsymbol{d}^{(i)}, \boldsymbol{\phi}^{(i)}\right)$$

- Typically Metropolis within Gibbs updates are needed

Example: orange tree growth

- Picchini and Ditlevsen (2011) discuss a model for the growth of orange trees incorporating random effects. We use an equivalent reparameterisation, written as

$$dX_t^{(i)} = \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} X_t^{(i)} \left(\phi_1^{(i)} - X_t^{(i)}\right) dt + \sigma \sqrt{X_t^{(i)}} dW_t^{(i)}, \quad i = 1, \ldots, 100$$

- $\phi_1^{(i)} \sim N(\phi_1, \sigma^2_{\phi_1})$ and $\phi_2^{(i)} \sim N(\phi_2, \sigma^2_{\phi_2})$
- X_t is the circumference (mm)
- t is the number of days since December 31^{st} 1968
- $\phi_1^{(i)}$ is the asymptotic circumference
- $\phi_2^{(i)}$ is the rate of change parameter

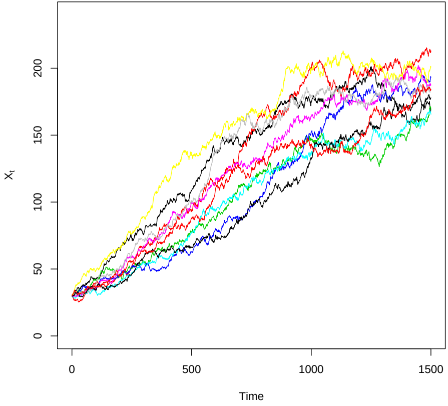
$$dX_t^{(i)} = \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} X_t^{(i)} \left(\phi_1^{(i)} - X_t^{(i)}\right) dt + \sigma\sqrt{X_t^{(i)}} dW_t^{(i)}, \quad i = 1, \ldots, 100$$

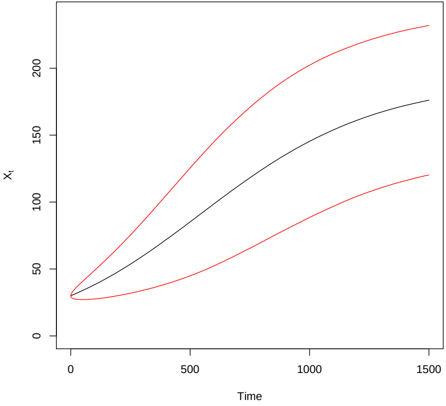
- Common parameters: ϕ_1, ϕ_2, σ_{ϕ_1}, σ_{ϕ_2} and σ
- Tree specific parameters: $\phi_1^{(i)}$ and $\phi_2^{(i)}$
- We repeat Picchini and Ditlevsen's simulation study for 100 trees with $x_0^{(i)} = 30$,

$$(\phi_1, \phi_2, \sigma_{\phi_1}, \sigma_{\phi_2}, \sigma) = (195, 350, 25, 52.5, 0.08),$$

  $\phi_1^{(i)} \sim N(195, 25^2)$ and $\phi_2^{(i)} \sim N(350, 52.5^2)$
- This gives us 16 observations at intervals of 100 days on 100 trees

Figure: **left:** 10 simulated skeleton paths
right: mean, upper and lower 2.5 percentiles for 1k realisations





Possible MCMC scheme

- Recall we wish to sample
  - $\boldsymbol{\theta}|\boldsymbol{x},\boldsymbol{d},\boldsymbol{\phi}$
  - $\boldsymbol{\phi}|\boldsymbol{x},\boldsymbol{d},\boldsymbol{\theta}$
  - $\boldsymbol{x}|\boldsymbol{\theta},\boldsymbol{d},\boldsymbol{\phi}$
- We can formulate the joint posterior as

$$\begin{aligned}\pi\left(\boldsymbol{\theta},\boldsymbol{\phi},\boldsymbol{x}|\boldsymbol{d}\right) \propto\ & \pi(\phi_1)\pi(\phi_2)\pi(\sigma_{\phi_1})\pi(\sigma_{\phi_2})\pi(\sigma) \\ & \times \pi(\boldsymbol{\phi}_1|\phi_1,\sigma_{\phi_1})\pi(\boldsymbol{\phi}_2|\phi_2,\sigma_{\phi_2}) \\ & \times \prod_{i=1}^{100} \underbrace{\pi\left(\boldsymbol{x}^{(i)}|\phi_1^{(i)},\phi_2^{(i)},\sigma,\boldsymbol{d}^{(i)}\right)}_{\text{Euler density}}\end{aligned}$$

- Sample common parameters $\boldsymbol{\theta}$ via Gibbs updates
- Sample each $\boldsymbol{\phi}^{(i)}$ using random normal proposals
- Use fairly uninformative priors
    - ϕ_1 and ϕ_2 are Normally distributed
    - $\sigma_{\phi_1}^{-2}$, $\sigma_{\phi_2}^{-2}$ and σ^{-2} follow Gamma distributions
- Recall

$$dX_t^{(i)} = \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} X_t^{(i)} \left(\phi_1^{(i)} - X_t^{(i)}\right) dt + \sigma\sqrt{X_t^{(i)}} dW_t^{(i)}, \quad i = 1, \ldots, 100$$

  with $\phi_1^{(i)}$ and $\phi_2^{(i)}$ Normally distributed
- Sample $\boldsymbol{x}$ by updating each path $\boldsymbol{x}^i$ separately, conditional on $\phi_1^{(i)}$, $\phi_2^{(i)}$ and σ
- See my SBSSB talk on 5th December 2012 for details on the path update

(Modified) Innovation scheme for random effects SDEs

- As discussed in previous SBSSB talks, this type of scheme suffers from intolerably poor mixing as $m \to \infty$
- Caused by relation between the parameters and the path in the quadratic variation $\Rightarrow$ acceptance probability $\to 0$ as $m \to \infty$
- The (Modified) Innovation scheme conditions on the Brownian increments

$$\boldsymbol{w} = \left(\boldsymbol{w}^{(1)}, \boldsymbol{w}^{(2)}, \dots, \boldsymbol{w}^{(\ell)}\right)$$

  (which drive the D&G bridges) to overcome the dependence between the parameters and the path
- Insight: the quadratic variation of $\boldsymbol{w}$ does not itself determine any of the model parameters and should therefore be effective in decoupling the problematic dependence

  $\Rightarrow$ the scheme should not become degenerate as $m \to \infty$

- We now alternate between draws of
  - $\boldsymbol{\theta}|\boldsymbol{w}, \boldsymbol{d}, \boldsymbol{\phi}$
  - $\boldsymbol{\phi}|\boldsymbol{w}, \boldsymbol{d}, \boldsymbol{\theta}$
  - $\boldsymbol{w}|\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{\phi}$
- Even easier for orange tree growth as $\beta\left(\cdot, \boldsymbol{\theta}, \boldsymbol{\phi}^{(i)}\right)$ only depends on σ
- Therefore

$$\begin{aligned}\pi\left(\boldsymbol{\theta}\backslash\sigma|\boldsymbol{w}, \boldsymbol{d}, \boldsymbol{\phi}\right) &= \pi\left(\boldsymbol{\theta}\backslash\sigma|\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{\phi}\right)\\ \pi\left(\boldsymbol{\phi}|\boldsymbol{w}, \boldsymbol{d}, \boldsymbol{\theta}\right) &= \pi\left(\boldsymbol{\phi}|\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{\theta}\right)\\ \pi\left(\boldsymbol{w}|\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{\phi}\right) &= \pi\left(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{d}, \boldsymbol{\phi}\right)\end{aligned}$$

- The only modification to our previous scheme is to now sample $\sigma|\boldsymbol{w}, \boldsymbol{d}, \boldsymbol{\phi}$

The linear noise approximation (LNA)

- Various SBSSB talks on this method or see Golightly and Gillespie (2013)
- Let

$$\boldsymbol{X}_t = \boldsymbol{Z}_t + \boldsymbol{M}_t$$

- Taylor expand $\boldsymbol{\alpha}\left(\boldsymbol{Z}_t + \boldsymbol{M}_t, \boldsymbol{\theta}, \boldsymbol{\phi}\right)$ and $\sqrt{\boldsymbol{\beta}\left(\boldsymbol{Z}_t + \boldsymbol{M}_t, \boldsymbol{\theta}, \boldsymbol{\phi}\right)}$ around $\boldsymbol{Z}_t$

$$\boldsymbol{\alpha}(\boldsymbol{Z}_t + \boldsymbol{M}_t, \boldsymbol{\theta}, \boldsymbol{\phi}) = \boldsymbol{\alpha}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi}) + \boldsymbol{F}_t \boldsymbol{M}_t + \cdots$$

$$\sqrt{\boldsymbol{\beta}(\boldsymbol{Z}_t + \boldsymbol{M}_t, \boldsymbol{\theta}, \boldsymbol{\phi})} = \sqrt{\boldsymbol{\beta}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi})} + \cdots$$

where $\boldsymbol{F}_t$ is the Jacobian matrix with i, j^{th} element

$$(\boldsymbol{F}_t)_{i,j} = \frac{\partial \boldsymbol{\alpha}_i(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi})}{\partial Z_{j,t}}$$

- We assume that the drift dominates the diffusion
- For fixed or Gaussian initial conditions, $\boldsymbol{M}_{t_0} \sim N_d(\boldsymbol{m}_{t_0}, \boldsymbol{V}_{t_0})$ and $\boldsymbol{M}_t \sim N_d(\boldsymbol{m}_t, \boldsymbol{V}_t)$
- The LNA is characterised by the system of ODEs

$$\begin{aligned}
\frac{d\boldsymbol{Z}_t}{dt} &= \boldsymbol{\alpha}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi}) \\
\frac{d\boldsymbol{m}_t}{dt} &= \boldsymbol{F}_t \boldsymbol{m}_t \\
\frac{d\boldsymbol{V}_t}{dt} &= \boldsymbol{F}_t \boldsymbol{V}_t + \sqrt{\boldsymbol{\beta}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi})} \sqrt{\boldsymbol{\beta}(\boldsymbol{Z}_t, \boldsymbol{\theta}, \boldsymbol{\phi})}^T + \boldsymbol{V}_t \boldsymbol{F}_t^T
\end{aligned}$$

- Solving this system gives $\boldsymbol{X}_t | \boldsymbol{X}_{t_0} \sim N_d(\boldsymbol{Z}_t + \boldsymbol{m}_t, \boldsymbol{V}_t)$

- The accuracy of the LNA can become poor over time
- Restart the LNA at each simulation time (Fearnhead *et al*, 2012)
- We solve the system of ODEs over each interval $[t, t + \Delta t]$ where $\boldsymbol{Z}_t = \boldsymbol{X}_t$ and $\boldsymbol{V}_t = \boldsymbol{0}$
- Note using this restart means that $\boldsymbol{m}_t$ is $\boldsymbol{0}$ for all t and as such the second equation need not be solved
- Advantageous as it is generally more tractable than the CLE
- Will not suffer from mixing issues (see previous SBSSB talks)

The LNA: orange tree growth

- Recall for $i = 1, \ldots, 100$

$$dX_t^{(i)} = \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} X_t^{(i)} \left(\phi_1^{(i)} - X_t^{(i)}\right) dt + \sigma \sqrt{X_t^{(i)}} dW_t^{(i)}$$

- The LNA for a particular tree i is characterised by

$$\begin{aligned}
\frac{dZ_t^{(i)}}{dt} &= \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} Z_t^{(i)} \left(\phi_1^{(i)} - Z_t^{(i)}\right) \\
\frac{dm_t^{(i)}}{dt} &= \frac{1}{\phi_1^{(i)}\phi_2^{(i)}} \left(\phi_1^{(i)} - 2Z_t^{(i)}\right) m_t^{(i)} \\
\frac{dV_t^{(i)}}{dt} &= \frac{2}{\phi_1^{(i)}\phi_2^{(i)}} \left(\phi_1^{(i)} - 2Z_t^{(i)}\right) V_t^{(i)} + \sigma^2 Z_t^{(i)}
\end{aligned}$$

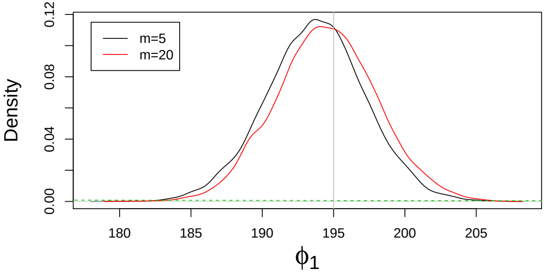
- This system can be solved analytically

Application: orange tree growth

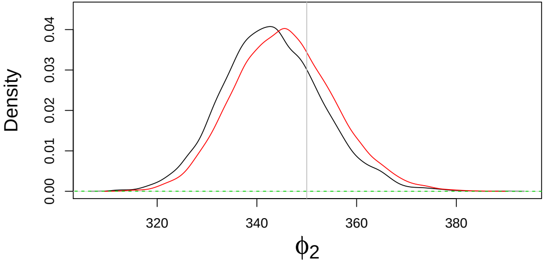
- Recall we have 16 observations on 100 trees at intervals of 100 days
- Parameter choice:

$$(\phi_1, \phi_2, \sigma_{\phi_1}, \sigma_{\phi_2}, \sigma) = (195, 350, 25, 52.5, 0.08)$$
$$\phi_1^{(i)} \sim N(195, 25^2)$$
$$\phi_2^{(i)} \sim N(350, 52.5^2)$$

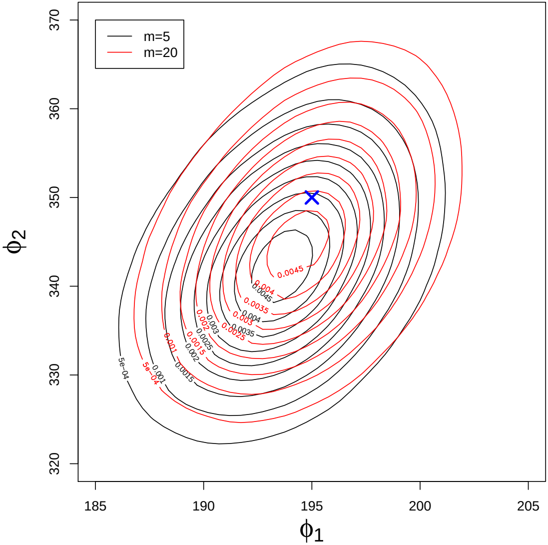
- We run the (Modified) Innovation scheme with $m = 20$
- We run the (Modified) Innovation scheme and the LNA for
  - 1 million iterations
  - thin of 100
  - burn in of 1000



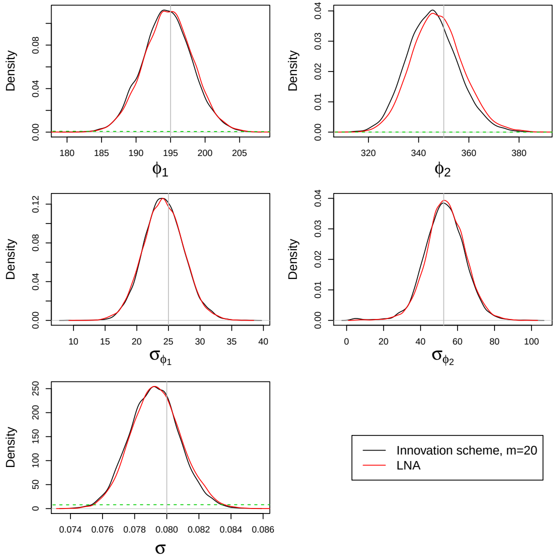
Density
0.00
0.04
0.08
0.12
m=5
m=20
180
185
190
195
200
205
ϕ_1



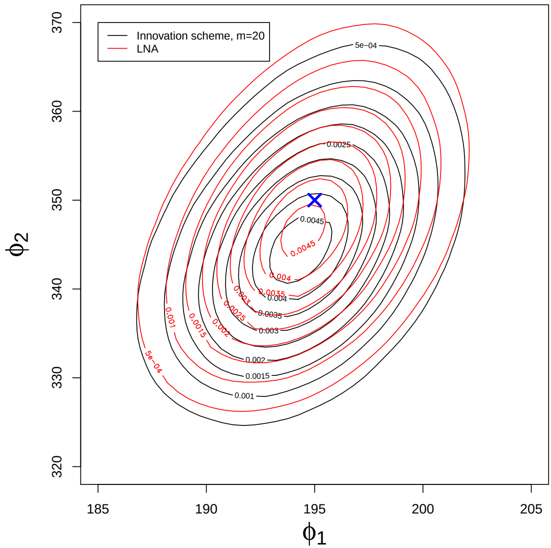
Density
0.00
0.01
0.02
0.03
0.04
320
340
360
380
ϕ_2

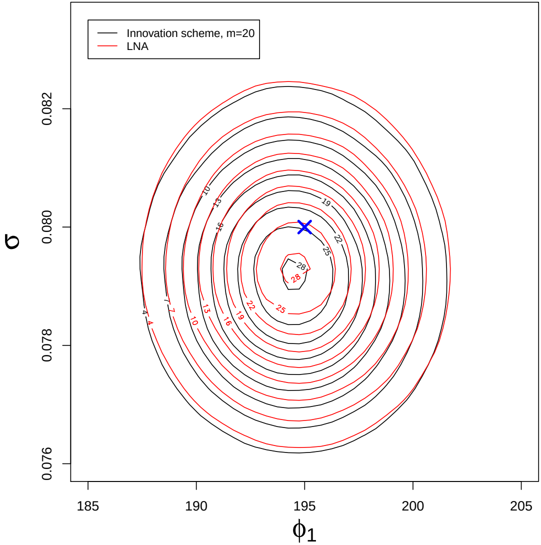

m=5
m=20
ϕ_1
ϕ_2
185
190
195
200
205
320
330
340
350
360
370

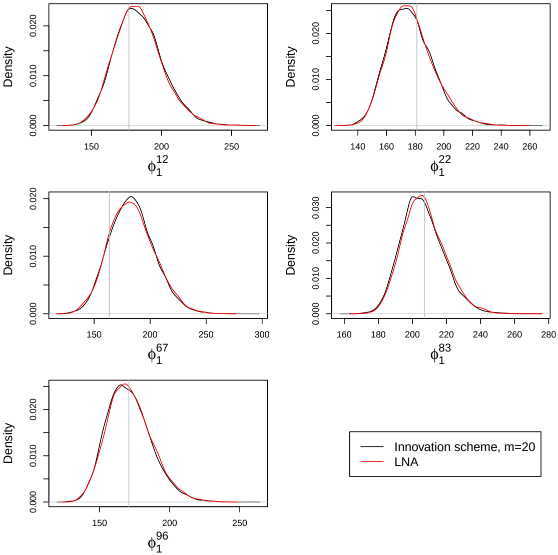
Results

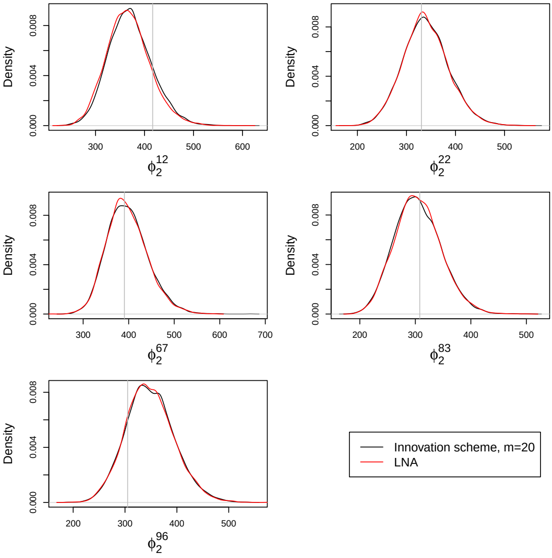


Density
ϕ_1
ϕ_2
σ_{ϕ_1}
σ_{ϕ_2}
σ
Innovation scheme, m=20
LNA

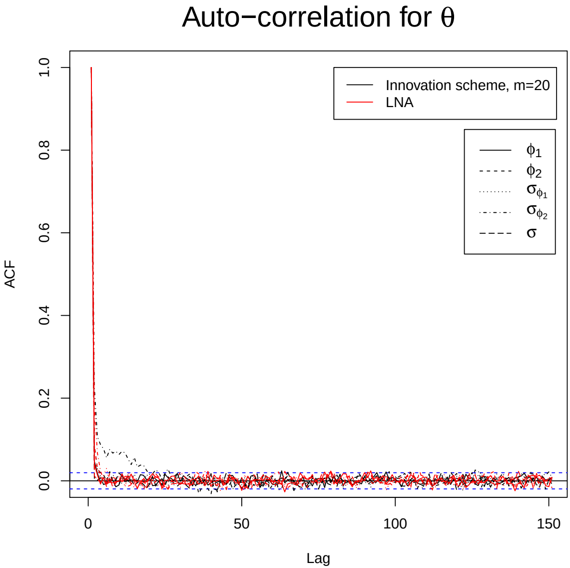

Innovation scheme, m=20
LNA
5e−04
0.001
0.0015
0.002
0.0025
0.003
0.0035
0.004
0.0045
185
190
195
200
205
320
330
340
350
360
370
ϕ_1
ϕ_2


Innovation scheme, m=20
LNA
σ
0.082
0.080
0.078
0.076
185
190
195
200
205
ϕ_1
4
7
10
13
16
19
22
25
28

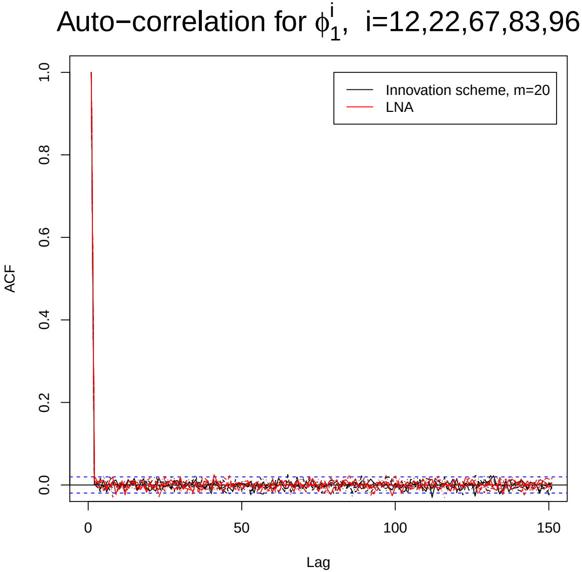

Density
0.000
0.010
0.020
150
200
250
ϕ_1^{12}
140
160
180
200
220
240
260
ϕ_1^{22}
300
ϕ_1^{67}
0.030
280
ϕ_1^{83}
ϕ_1^{96}
Innovation scheme, m=20
LNA

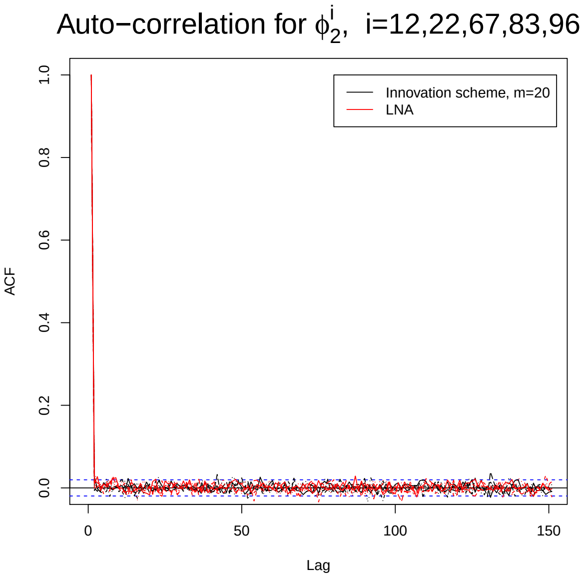


Density
0.000
0.004
0.008
300
400
500
600
ϕ_2^{12}
200
ϕ_2^{22}
700
ϕ_2^{67}
ϕ_2^{83}
ϕ_2^{96}
Innovation scheme, m=20
LNA



Auto−correlation for θ
Innovation scheme, m=20
LNA
ϕ_1
ϕ_2
σ_{ϕ_1}
σ_{ϕ_2}
σ
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
Lag


Auto−correlation for ϕ_1^i, i=12,22,67,83,96
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
Lag
Innovation scheme, m=20
LNA

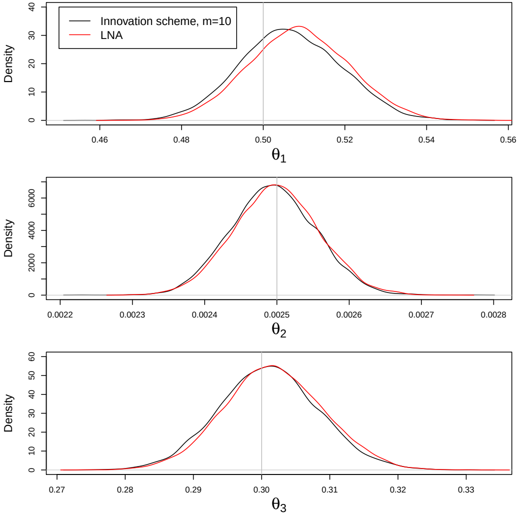

Auto−correlation for ϕ_2^i, i=12,22,67,83,96
Innovation scheme, m=20
LNA
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150

Comparison of the schemes

- Comparing the LNA against the (Modified) Innovation scheme
  - LNA: ESS/sec $= 1.178$
  - (Modified) Innovation scheme: ESS/sec $= 0.057$

  which gives a comparable ESS/sec $= 20.667$
- The gain by using the LNA is slightly exaggerated by the fact that the system of ODEs can be solved analytically
- However it should still be advantageous to use the LNA over the (Modified) Innovation scheme on a system that can't be solved analytically
- What about the Lotka-Volterra model?
  Look at fixed effects: LNA ODEs can't be solved analytically
  Using the (Modified) Innovation scheme with $m = 10$
  Comparable ESS/sec $= 1.776$

Figure: 1 million iterations with a thin of 100



Future work

- Extend these methods to data where we only have partial observations
- Examine the case where we have observations observed with error, typically Gaussian error
- Apply these schemes to a more challenging example e.g. larger model
- Compare these methods with pMCMC
- Apply these schemes to lemon trees

References

- Boys, R. J. Wilkinson, D. J. and Kirkwood, T. B. L. *Bayesian inference for a discretely observed stochastic kinetic model.* Statistics and Computing, 18 (2) 125-135, 2008
- Dargatz, C. *Bayesian Inference for Diffusion Processes with Applications in Life Sciences.* PhD thesis, Ludwig–Maximilians–Universität München, 2010
- Fearnhead, P. Giagos, V. and Sherlock, C. *Inference for reaction networks using the linear noise approximation.* http://arxiv.org/abs/1205.6920, 2012
- Golightly, A. and Gillespie, C.S. *Simulation of stochastic kinetic models.* In Silico Systems Biology, 169-187, 2013
- Picchini, U. and Ditlevsen, S. *Practical estimation of high dimensional stochastic differential mixed-effects models.* Computational Statistics and Data Analysis, 55 (E3) 1426-1444, 2011
- Wilkinson, D. J. and Golightly, A. *Markov chain Monte Carlo algorithms for SDE parameter estimation.* Learning and Inference in Computational Systems Biology, 2010