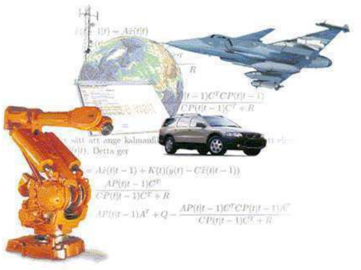
Speeding up the particle Metropolis-Hastings algorithm for Bayesian parameter inference

University of Newcastle, Newcastle, Australia, November 28, 2014.



Johan Dahlin
johan.dahlin@liu.se

Division of Automatic Control,
Linköping University,
Sweden.

This is collaborative work with my supervisors!

Dr. Fredrik Lindsten (University of Cambridge, United Kingdom)
Prof. Thomas B. Schön (Uppsala University, Sweden)

Outline

Pseudo-marginal Metropolis-Hastings

- Pseudo-marginal algorithms.
- Particle filtering for likelihood estimation.

Particle Metropolis-Hastings using gradients and Hessians

- Fixed-lag particle smoothing for estimating gradients and Hessians.
- Extending the proposal with gradient and Hessian information.

Particle Metropolis-Hastings with intractable likelihoods

- Particle filtering using approximate Bayesian computations.

Nonlinear state space model

Consider a nonlinear state space model

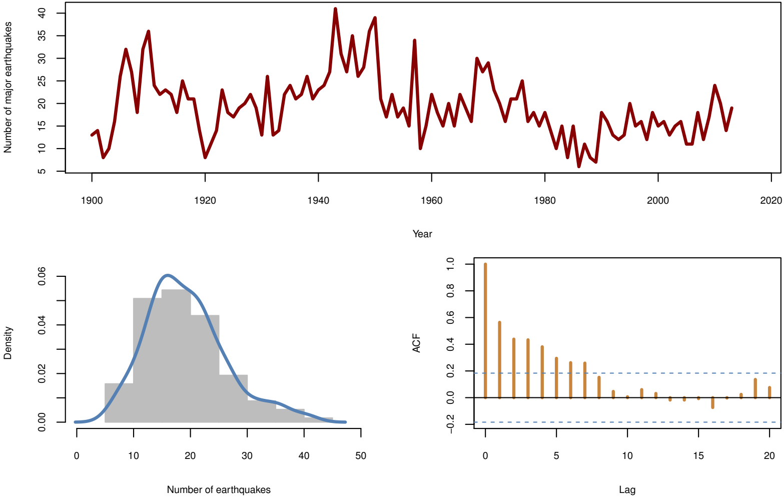
$$
\begin{aligned}
x_0 &\sim \mu(x_0), \\
x_{t+1} | x_t &\sim f_{\boldsymbol{\theta}}(x_{t+1} | x_t), \\
y_t | x_t &\sim g_{\boldsymbol{\theta}}(y_t | x_t),
\end{aligned}
$$

where $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$, $x_t \subset \mathbb{R}^n$ and $y_t \subset \mathbb{R}^m$.

Inference: compute estimates of $x_{0:T}$ and θ given $y_{1:T}$.

Example: Earthquakes between 1900 and 2013

Example: Earthquakes between 1900 and 2013



Example: A simple model of annual earthquake counts

$$x_{t+1}|x_t \sim \mathcal{N}\Big(x_{t+1}; \boldsymbol{\phi} x_t, \boldsymbol{\sigma}^2\Big),$$
$$y_t|x_t \sim \mathcal{P}\Big(y_t; \boldsymbol{\beta} \exp(x_t)\Big),$$

where the parameters describe:

- $\boldsymbol{\phi}$: persistence of intensity.
- $\boldsymbol{\sigma}$: standard deviation of innovation in intensity.
- $\boldsymbol{\beta}$: *nominal* number of annual earthquakes.

Task: Estimate $\boldsymbol{\theta} = \{\boldsymbol{\phi}, \boldsymbol{\sigma}, \boldsymbol{\beta}\}$ and $x_{0:T}$ given $y_{1:T}$.

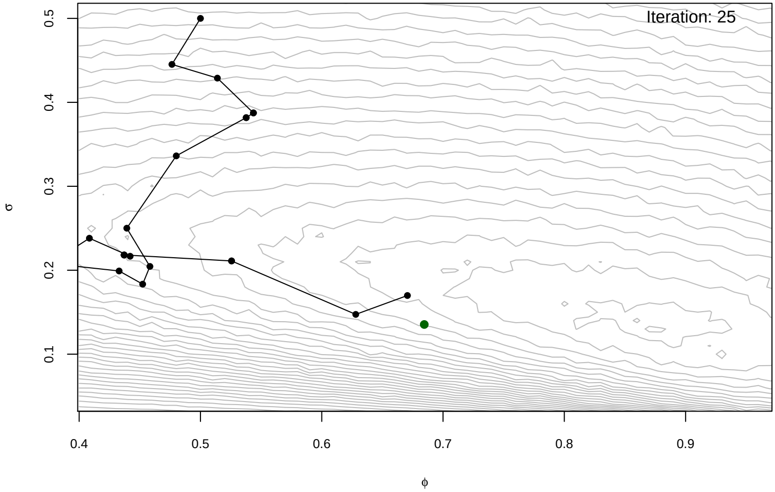
Bayesian parameter inference

Consider the *parameter posterior*

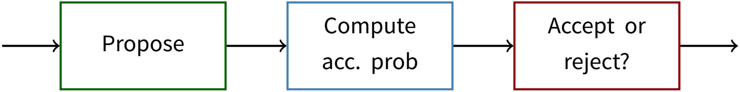
$$\pi(\theta) = p(\theta | y_{1:T}) = \frac{p_\theta(y_{1:T}) p(\theta)}{p(y_{1:T})} \propto p_\theta(y_{1:T}) p(\theta).$$

posterior = *prior* + *information in data* + *model*.

Metropolis-Hastings algorithm



Metropolis-Hastings algorithm



- Propose: $\theta' \sim q(\theta'|\theta_{k-1})$.
- Compute acceptance probability:

$$\alpha(\theta', \theta_{k-1}) = \min\left\{1, \frac{p(\theta')}{p(\theta_{k-1})} \frac{p_{\theta'}(y_{1:T})}{p_{\theta_{k-1}}(y_{1:T})} \frac{q(\theta_{k-1}|\theta')}{q(\theta'|\theta_{k-1})}\right\}.$$

- Accept or reject? $\theta' \to \theta_k$ w.p. $\alpha(\theta', \theta_{k-1})$.

Particle Metropolis-Hastings algorithm

Problem

We cannot compute $p_\theta(y_{1:T})$ in closed form.

Idea

Replace the likelihood with an unbiased estimate $\widehat{p}_\theta(y_{1:T}|u)$.

Implementation

Run a particle filter to estimate the likelihood and $\alpha(\theta'', \theta')$.

Exact approximations

Keeps the Markov chain invariant.
The marginal of the stationary distribution is $\pi(\theta)$.

Particle Metropolis-Hastings algorithm

The target distribution is given by the parameter proposal

$$\pi(\theta) = \frac{p_\theta(y_{1:T})p(\theta)}{p(y_{1:T})}.$$

An unbiased estimator of the likelihood is given by

$$\mathbb{E}_m\left[\widehat{p}_\theta(y_{1:T}|u)\right] = \int \widehat{p}_\theta(y_{1:T}|u) m_\theta(u)\, \mathrm{d}u = p_\theta(y_{1:T}).$$

An extended target is given by

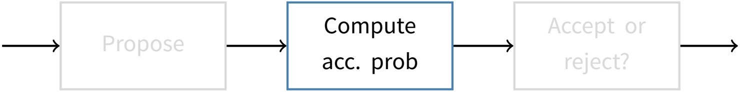
$$\pi(\theta, u) = \frac{\widehat{p}_\theta(y_{1:T}|u) m_\theta(u) p(\theta)}{p(y_{1:T})} = \frac{\widehat{p}_\theta(y_{1:T}|u) m_\theta(u) \pi(\theta)}{p_\theta(y_{1:T})}.$$

Particle Metropolis-Hastings algorithm (cont.)

$$\int \pi(\theta, u)\,\mathrm{d}u = \int \frac{\widehat{p}_\theta(y_{1:T}|u) m_\theta(u) \pi(\theta)}{p_\theta(y_{1:T})}\,\mathrm{d}u$$
$$= \frac{\pi(\theta)}{p_\theta(y_{1:T})} \underbrace{\int \widehat{p}_\theta(y_{1:T}|u) m_\theta(u)\,\mathrm{d}u}_{=p_\theta(y_{1:T})},$$
$$= \pi(\theta).$$

That is, the marginal is the desired target distribution and the Markov chain is kept invariant.

Particle Metropolis-Hastings algorithm (cont.)

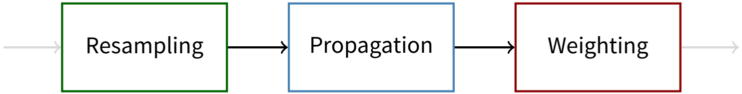


- Propose: $\theta' \sim q(\theta' | \theta_{k-1}, u')$ and $u' \sim \text{PF}(\theta')$.
- Compute $\widehat{p}_{\theta'}(y_{1:T} | u')$ and the acceptance probability:

$$\alpha(\theta', \theta_{k-1}) = 1 \wedge \frac{p(\theta')}{p(\theta_{k-1})} \frac{\widehat{p}_{\theta'}(y_{1:T} | u')}{\widehat{p}_{\theta_{k-1}}(y_{1:T} | u_{k-1})} \frac{q(\theta_{k-1} | \theta', u')}{q(\theta' | \theta_{k-1}, u_{k-1})}.$$

- Accept or reject? $\theta' \to \theta_k$ and $u' \to u_k$ w.p. $\alpha(\theta', \theta_{k-1})$.

Particle filtering



- Resampling: $\mathbb{P}(a_t^{(i)} = j) = \widetilde{w}_{t-1}^{(j)}$ and set $\widetilde{x}_{t-1}^{(i)} = x_{t-1}^{a_t^{(i)}}$.
- Propagation: $x_t^{(i)} \sim R_\theta\left(x_t \middle| \widetilde{x}_{t-1}^{(i)}\right) = f_\theta(x_t | \widetilde{x}_{t-1}^{(i)})$.
- Weighting: $w_t^{(i)} = W_\theta\left(x_t^{(i)}, \widetilde{x}_{t-1}^{(i)}\right) = g_\theta(y_t | x_t^{(i)})$.

Particle filtering (cont.)

Given the particle system (the random variables u)

$$u = \left\{ x_{1:T}^{(i)}, w_{1:T}^{(i)} \right\}_{i=1}^{N},$$

we can obtain (consistent) estimates of:

- the likelihood $p_\theta(y_{1:T})$.
- the first and second order information of $\pi(\theta)$.

Particle filtering (cont.)

The likelihood for an SSM can be decomposed by

$$\mathcal{L}(\theta) = p_\theta(y_{1:T}) = p_\theta(y_1) \prod_{t=2}^{T} p_\theta(y_t | y_{1:t-1}),$$

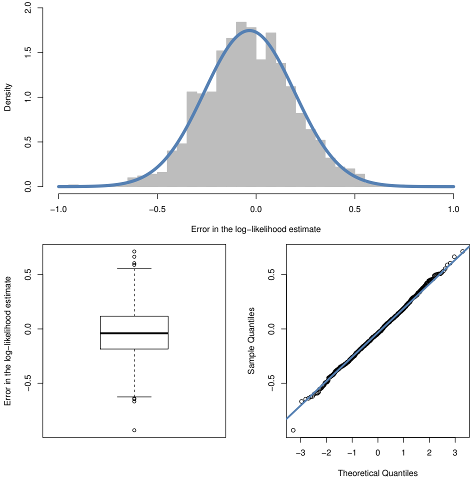
where the one-step ahead predictor can be computed by

$$p_\theta(y_t | y_{1:t-1}) \approx \frac{1}{N} \sum_{i=1}^{N} \int W_\theta(x_t | x_{t-1}) \delta_{x_t^{(i)}, \widetilde{x}_{t-1}^{(i)}} \, \mathrm{d}x_t = \frac{1}{N} \sum_{i=1}^{N} w_t^{(i)}.$$

Likelihood estimator

$$\widehat{\mathcal{L}}(\theta) = \widehat{p}_\theta(y_{1:T} | u) = \frac{1}{N^T} \prod_{t=1}^{T} \sum_{i=1}^{N} w_t^{(i)}.$$

$$\sqrt{N}\Big(\mathcal{L}(\theta) - \widehat{\mathcal{L}}(\theta)\Big) \xrightarrow{d} \mathcal{N}\left(0, \sigma_l^2\right).$$



Zeroth order proposal (PMH0)

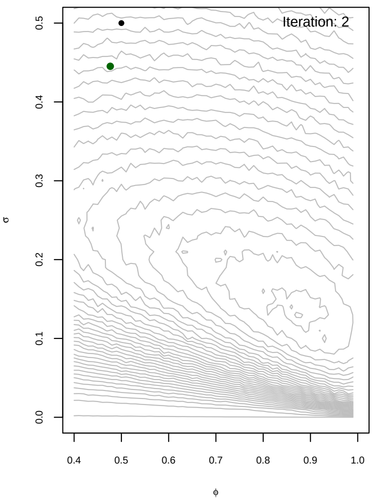
Gaussian random walk

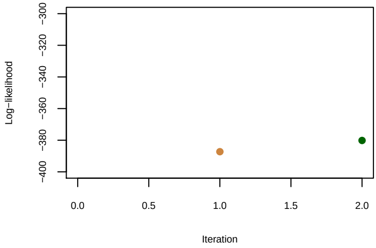
$$\theta'' = \theta' + \epsilon z, \quad z \sim \mathcal{N}(z; 0, 1).$$

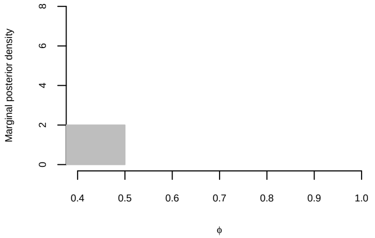
gives the zeroth order (marginal) proposal

$$q(\theta''|\theta', u') = \mathcal{N}\left(\theta''; \theta', \epsilon^2 I_d\right).$$

Example: Parameter inference in earthquake model



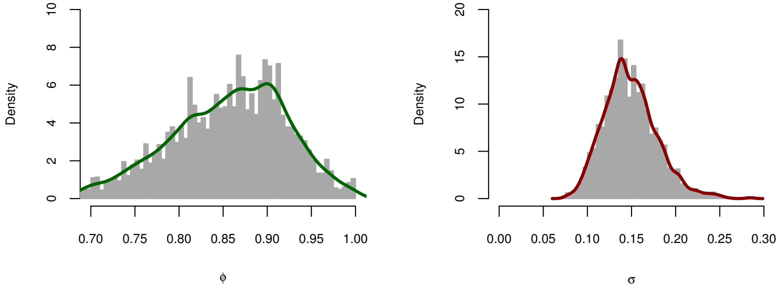




Example: Parameter inference in earthquake model

Example: Parameter inference in earthquake model

Example: Parameter inference in earthquake model



	ϕ	σ
Posterior mean	0.86	0.15
Posterior median	0.86	0.15
Posterior mode	0.90	0.14

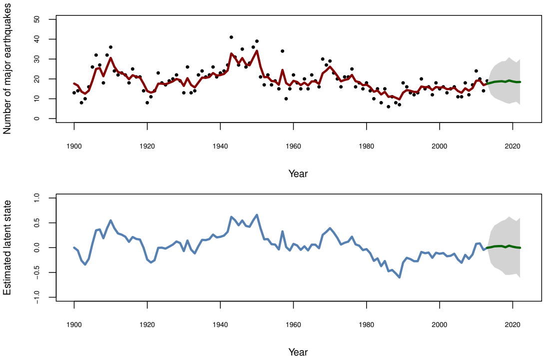
Example: State inference in the earthquake model

$$x_{t+1} | x_t \sim \mathcal{N}\Big(x_{t+1}; \phi x_t, \sigma^2\Big),$$

$$y_t | x_t \sim \mathcal{P}\Big(y_t; \beta \exp(x_t)\Big),$$

with parameters:

$\phi = 0.88$ (persistence.)
$\sigma = 0.15$ (sd. of innovation.)
$\beta = 17.65$ (*nominal* number.)



Conclusions

Results

Algorithm for Bayesian inference in nonlinear SSMs.
Reasonable performance on small models.

Methods

Pseudo-marginal version of Metropolis-Hastings.
Particle filtering.

References

C. Andrieu and G. O. Roberts. **The pseudo-marginal approach for efficient Monte Carlo computations**. The Annals of Statistics, 37(2):697-725, 2009.

C. Andrieu, A. Doucet, and R. Holenstein. **Particle Markov chain Monte Carlo methods**. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 72(3):269-342, 2010.

M. K. Pitt, R. S. Silva, P. Giordani, and R. Kohn. **On some properties of Markov chain Monte Carlo simulation methods based on the particle filter**. Journal of Econometrics, 171(2):134-151, 2012.

Outline

Pseudo-marginal Metropolis-Hastings

- Pseudo-marginal algorithms.
- Particle filtering for likelihood estimation.

Particle Metropolis-Hastings using gradients and Hessians

- Fixed-lag particle smoothing for estimating gradients and Hessians.
- Extending the proposal with gradient and Hessian information.

Particle Metropolis-Hastings with intractable likelihoods

- Particle filtering using approximate Bayesian computations.

Fixed-lag particle smoothing

The first and second order information can be estimated using

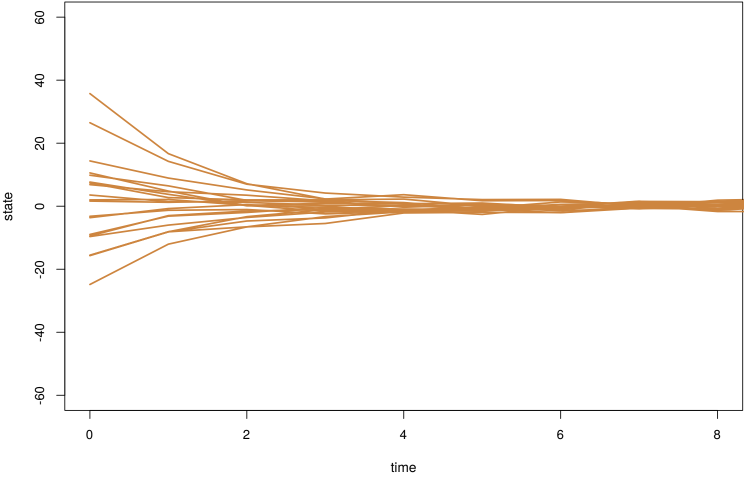
$$u = \left\{ x_{1:T}^{(i)}, w_{1:T}^{(i)} \right\}_{i=1}^{N},$$

and the fixed-lag particle smoother approximation,

$$\widehat{p}_\theta(\mathrm{d}x_{t:t-1} | y_{1:T}) \approx \widehat{p}_\theta(\mathrm{d}x_{t:t-1} | y_{1:\kappa_t}), \quad \kappa_t = \min\{T, t + \Delta\},$$

with no additional computational complexity.

Fixed-lag particle smoothing (cont.)



Fixed-lag particle smoothing (cont.)

Assume that

$$\widehat{p}_\theta(\mathrm{d}x_{t:t-1}|y_{1:T}) \approx \widehat{p}_\theta(\mathrm{d}x_{t:t-1}|y_{1:\kappa_t}), \quad \kappa_t = \min\{T, t+\Delta\},$$

for some $0 \leq \Delta \leq T$. It follows that

$$\widehat{p}_\theta(x_{t-1:t}|y_{1:T}) = \sum_{i=1}^{N} \widetilde{w}_{\kappa_t}^{(i)} \delta_{\widetilde{x}_{t-1:t,\kappa_t}^{(i)}}(\mathrm{d}x_{t-1:t}).$$

Fixed-lag particle smoothing (cont.)

The score can be estimated using Fisher's identity given by

$$\begin{aligned}\nabla_\theta \log p_\theta(y_{1:T})\big|_{\theta=\theta'} &= \int \nabla_\theta \log p_\theta(x_{1:T}, y_{1:T}) p_{\theta'}(x_{1:T}|y_{1:T}) \mathrm{d}x_{1:T} \\ &\approx \int \nabla_\theta \log p_\theta(x_{1:T}, y_{1:T}) \widehat{p}_{\theta'}(x_{1:T}|y_{1:T}) \mathrm{d}x_{1:T}\end{aligned}$$

We also know that

$$\nabla_\theta \log p_\theta(x_{1:T}, y_{1:T}) = \sum_{t=1}^{T} \underbrace{\left[\nabla_\theta \log f_\theta(x_t|x_{t-1}) + \nabla_\theta \log g_\theta(y_t|x_t)\right]}_{\triangleq \eta(x_t, x_{t-1})},$$

which gives

$$\nabla_\theta \log p_\theta(y_{1:T})\big|_{\theta=\theta'} \approx \sum_{t=1}^{T} \sum_{i=1}^{N} \widetilde{w}_{\kappa_t}^{(i)} \eta(\widetilde{x}_{t-1,\kappa_t}^{(i)}, \widetilde{x}_{t,\kappa_t}^{(i)}).$$

First order proposal (PMH1)

Noisy gradient-based ascent update

$$\theta'' = \theta' + \frac{\epsilon^2}{2}\mathcal{S}(\theta') + \epsilon z, \quad z \sim \mathcal{N}(z; 0, 1),$$

with the first order information

$$\mathcal{S}(\theta') = \nabla_\theta \log \pi(\theta)\big|_{\theta=\theta'},$$

gives the first order proposal

$$q(\theta'' | \theta', u') = \mathcal{N}\left(\theta''; \theta' + \frac{\epsilon^2}{2}\widehat{\mathcal{S}}(\theta' | u'), \epsilon^2 I_d\right).$$

Second order proposal (PMH2)

Noisy Newton update

$$\theta'' = \theta' + \frac{\epsilon^2}{2}\big[\mathcal{J}(\theta')\big]^{-1}\mathcal{S}(\theta') + \epsilon\big[\mathcal{J}(\theta')\big]^{-1/2} z, \quad z \sim \mathcal{N}(z; 0, 1),$$

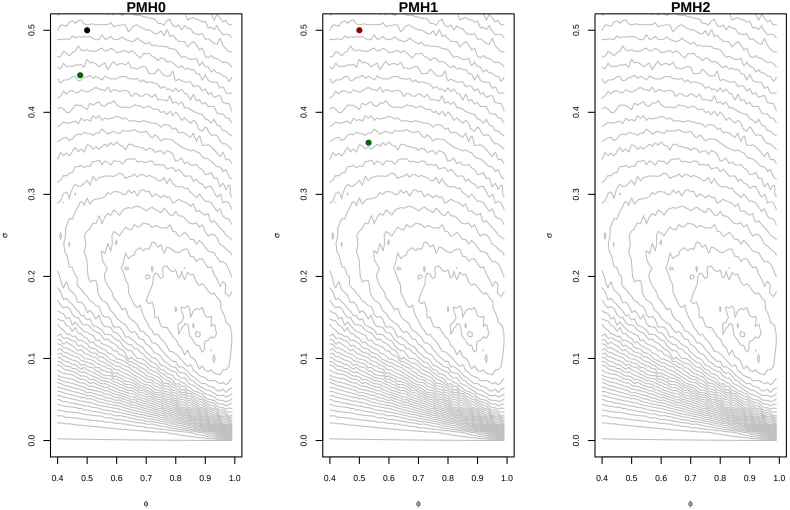
with the second order information

$$\mathcal{J}(\theta') = -\nabla^2_\theta \log \pi(\theta)\big|_{\theta=\theta'},$$

gives the second order proposal

$$q(\theta''|\theta', u') = \mathcal{N}\left(\theta''; \theta' + \frac{\epsilon^2}{2}\widehat{\mathcal{S}}(\theta'|u')\big[\widehat{\mathcal{J}}(\theta'|u')\big]^{-1}, \epsilon^2\big[\widehat{\mathcal{J}}(\theta'|u')\big]^{-1}\right).$$

Example: Parameter inference in earthquake model



Example: Parameter inference in earthquake model

Mixing

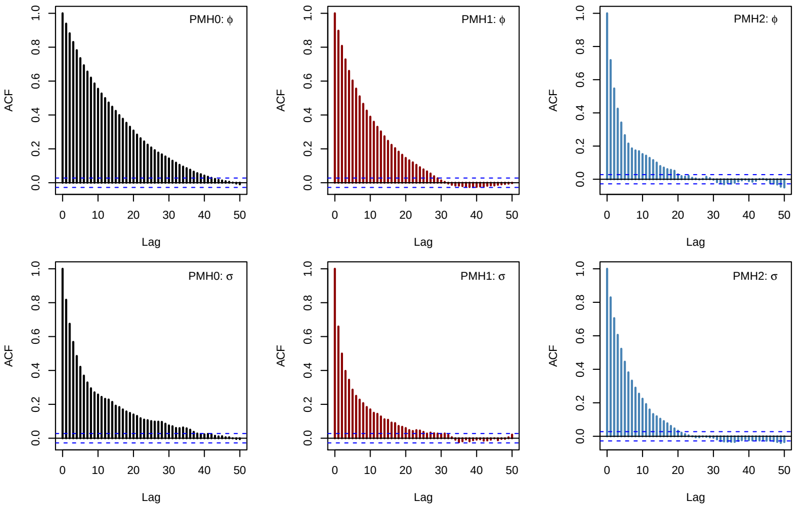
Let $\varphi(\theta)$ denote a test function, then

$$\sqrt{M}\left[\widehat{\varphi}_{\mathsf{MH}} - \mathbb{E}[\varphi(\theta)]\right] \xrightarrow{d} \mathcal{N}(0, \sigma_\varphi^2).$$

Here, σ_φ^2 depends on the integrated autocorrelation time (IACT)

$$\mathsf{IACT}(\theta_{1:M}) = 1 + 2\sum_{k=1}^{\infty} \rho_k(\theta_{1:M}).$$

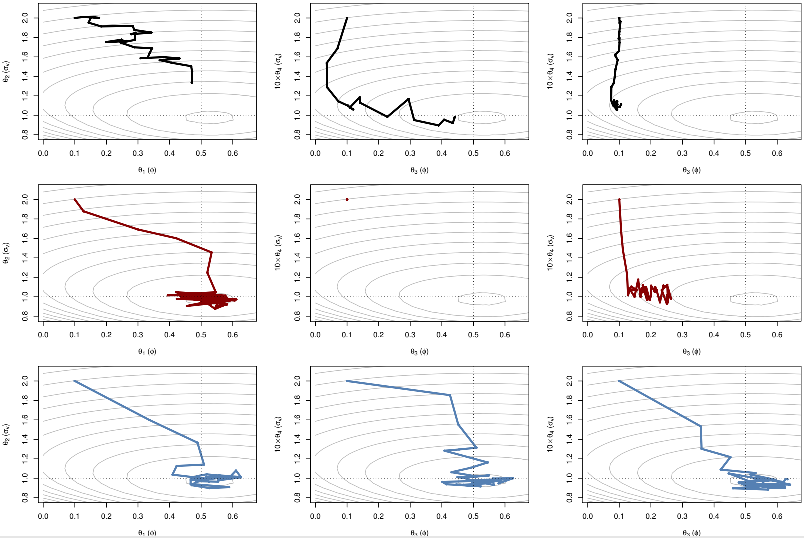
Integrated autocorrelation time



Integrated autocorrelation time (cont.)

	Acceptance rate	max IACT
standard PMH0	0.30	2639
pre-conditioned PMH0	0.45	129
standard PMH1	0.82	2875
pre-conditioned PMH1	0.70	1480
standard PMH2	0.28	139
hybrid PMH2	0.49	**23**

Scale-invariance property



Conclusions

Results

Shorter burn-in phase and increased mixing.
Simplified tuning due to scale invariance property.

Future work

Better estimation of the information matrix.
Non-reversible Markov chains and Hamiltonian Monte Carlo.

References

J. Dahlin, F. Lindsten and T. B. Schön, **Particle Metropolis-Hastings using gradient and Hessian information**. Statistics and Computing (MCMSki 2014 special issue), Springer, 2014.

J. Olsson, O. Cappé, R. Douc, and E. Moulines. **Sequential Monte Carlo smoothing with application to parameter estimation in nonlinear state space models**. Bernoulli, 14(1):155-179, 2008.

M. Girolami and B. Calderhead. **Riemann manifold Langevin and Hamiltonian Monte Carlo methods**. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(2):1-37, 2011.

Outline

Pseudo-marginal Metropolis-Hastings

- Pseudo-marginal algorithms.
- Particle filtering for likelihood estimation.

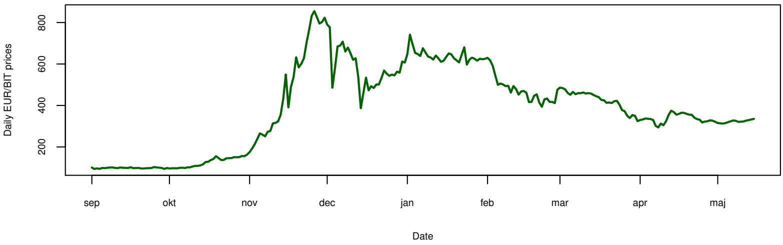
Particle Metropolis-Hastings using gradients and Hessians

- Fixed-lag particle smoothing for estimating gradients and Hessians.
- Extending the proposal with gradient and Hessian information.

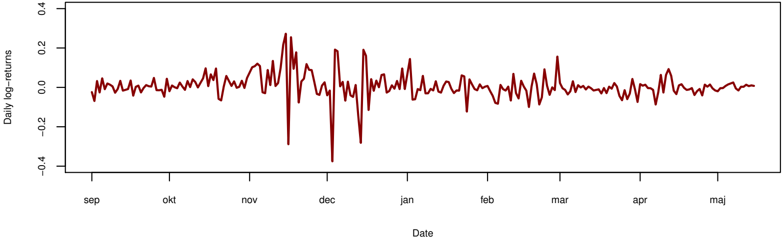
Particle Metropolis-Hastings with intractable likelihoods

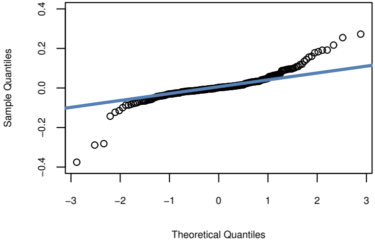
- Particle filtering using approximate Bayesian computations.

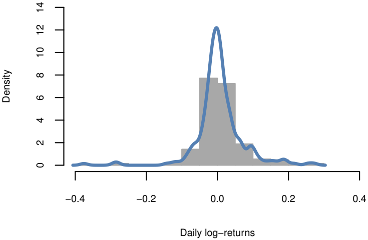
Modelling volatility in Bitcoin returns



Modelling volatility in Bitcoin returns







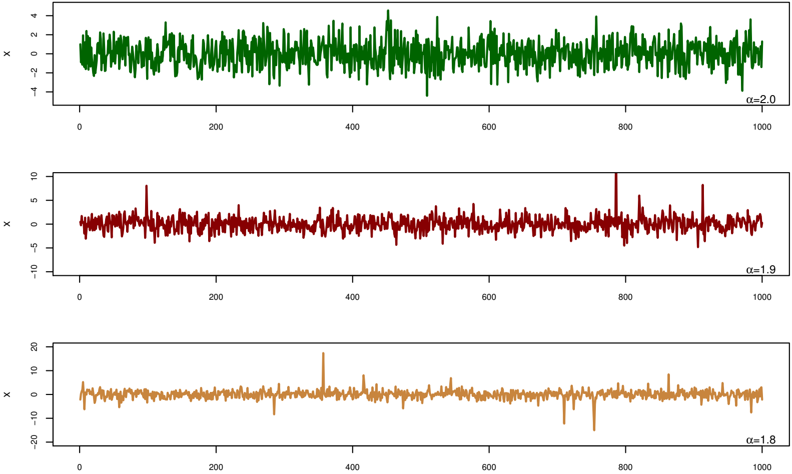
α-stable distributions

$$x \sim \mathcal{A}(x; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\eta}),$$

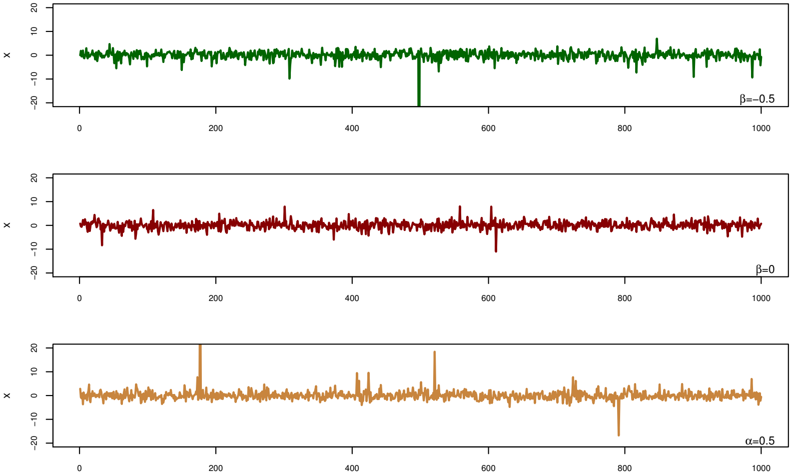
where the parameters describe:

- $\boldsymbol{\alpha} \in [0, 2]$: stability.
- $\boldsymbol{\beta} \in [-1, 1]$: skewness.
- $\boldsymbol{\gamma} \in \mathbb{R}_+$: scale (spread).
- $\boldsymbol{\eta} \in \mathbb{R}$: location.

α-stable distributions



α-stable distributions (cont.)



Stochastic volatility with α-stable returns

$$
\begin{aligned}
x_{t+1}|x_t &\sim \mathcal{N}\Big(x_{t+1}; \boldsymbol{\phi} x_t, \boldsymbol{\sigma}^2\Big), \\
y_t|x_t &\sim \mathcal{A}\Big(y_t; \boldsymbol{\alpha}, 0, \exp(x_t), 0\Big).
\end{aligned}
$$

where the parameters describe:

- ϕ: persistence of volatility.
- $\boldsymbol{\sigma}$: standard deviation of innovation in volatility.
- $\boldsymbol{\alpha}$: stability.

Task: Estimate $x_{0:T}$ given $y_{1:T}$ (requires $\boldsymbol{\theta} = \{\boldsymbol{\phi}, \boldsymbol{\sigma}, \boldsymbol{\alpha}\}$).

Approximate Bayesian computations

Consider the problem of estimating the parameter posterior

$$p(\theta|y_{1:T}) \propto p(\theta)p_\theta(y_{1:T}).$$

Importance sampling (IS)

(i) For $k = 1, \dots, K$

  (a) Sample $\theta^{(k)} \sim p(\theta)$,

  (b) Compute the weight $w^{(k)} = p(\theta^{(k)})p_{\theta^{(k)}}(y_{1:T})$,

(ii) Estimate the parameter posterior mean

$$\mathbb{E}[\theta|y_{1:T}] \approx \sum_{k=1}^{K} \frac{w^{(k)}}{\sum_{l=1}^{K} w^{(l)}} \theta^{(k)}.$$

Approximate Bayesian computations (cont.)

Consider extending the parameter posterior by

$$p(\theta|y_{1:T}) \propto \int \kappa_\epsilon\Big(\rho\{\eta(y_{1:T}), \eta(\tilde{y}_{1:T})\}\Big) p(\theta)\nu_\theta(\tilde{y}_{1:T})\,\mathrm{d}\tilde{y}_{1:T},$$

where the user selects

- The kernel $\kappa_\epsilon(\cdot)$ (Gaussian or uniform).
- The tolerance level (or bandwidth) ϵ.
- The distance $\rho(\cdot)$ (L_1 or L_2 norm).
- The transformation $\eta(\cdot)$ (sufficient or near-sufficient statistics).

Approximate Bayesian computations (cont.)

ABC-IS

(i) For $k = 1, \ldots, K$

   (a) Sample $\theta^{(k)} \sim p(\theta)$,

   (b) Generate $\tilde{y}_{1:T} \sim \nu_{\theta^{(k)}}$.

   (c) Compute the weight

$$w^{(k)} = p(\theta^{(k)}) \kappa_\epsilon \Big(\rho\{\eta(y_{1:T}), \eta(\tilde{y}_{1:T})\} \Big),$$

(ii) Estimate the parameter posterior mean

$$\mathbb{E}[\theta | y_{1:T}] \approx \sum_{k=1}^{K} \frac{w^{(k)}}{\sum_{l=1}^{K} w^{(l)}} \theta^{(k)}.$$

Assume that we can simulate an α-stable r.v. by $\tilde{y}_t \sim \tau_\theta(v_t, x_t)$.

Extend the model to obtain

$$\begin{aligned} x_{t+1}|x_t &\sim \mathcal{N}\Big(x_{t+1}; \phi x_t, \sigma^2\Big), \\ v_t|x_t &\sim \nu_\theta(v_t; x_t), \\ y_t|v_t &\sim h_{\theta,\epsilon}(y_t|v_t) = \frac{1}{\epsilon}\kappa\left(\frac{y_t - \tilde{y}_t}{\epsilon}\right), \end{aligned}$$

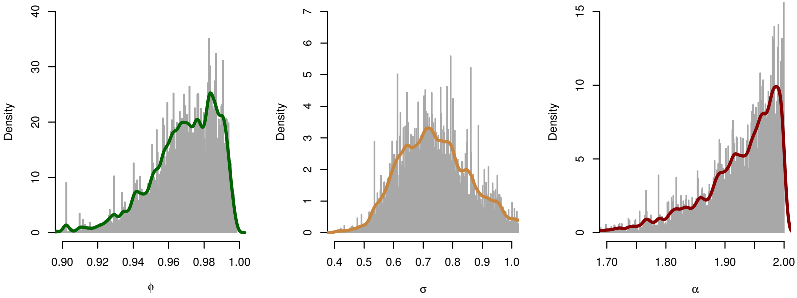
where (x_t, v_t) is the new state vector.

Particle filtering with ABC

- Resampling: $\mathbb{P}(a_t^{(i)} = j) = \widetilde{w}_{t-1}^{(j)}$ and set $\widetilde{x}_{t-1}^{(i)} = x_{t-1}^{a_t^{(i)}}$.
- Propagation: $x_t^{(i)} \sim R_\theta\left(x_t | \widetilde{x}_{t-1}^{(i)}\right) = f_\theta(x_t | \widetilde{x}_{t-1}^{(i)})$.
- Propagation: $v_t^{(i)} \sim \nu_\theta\left(v_t | x_t^{(i)}\right)$.
- Weighting: $w_t^{(i)} = h_{\theta,\epsilon}(y_t, v_t^{(i)}) = \mathcal{K}_\epsilon(y_t, \tilde{y}_t^{(i)})$.

Results in an unbiased likelihood estimator (assumptions).

Modelling volatility in Bitcoin returns



	Bitcoin			OMXS30		
	ϕ	σ	α	ϕ	σ	α
Posterior mean	0.97	**0.75**	1.92	0.96	**0.31**	1.93
Posterior median	0.97	**0.74**	1.94	0.96	**0.30**	1.94
Posterior mode	0.98	**0.72**	1.99	0.97	**0.22**	1.94

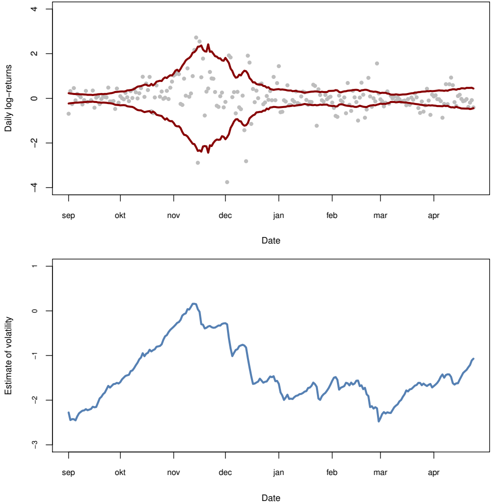
Modelling volatility in Bitcoin returns

$$x_{t+1} | x_t \sim \mathcal{N}\Big(x_{t+1}; \phi x_t, \sigma^2\Big),$$

$$y_t | x_t \sim \mathcal{A}\Big(y_t; \alpha, 0, \exp(x_t), 0\Big),$$

with parameters:

$\phi = 0.965$ (persistence.)
$\sigma = 0.740$ (sd. of innovation.)
$\alpha = 1.925$ (stability.)



Conclusions

Results

Reasonable parameter estimate in a model with intractable likelihood.
Poor mixing in the resulting Markov chain.
Computationally costly as N needs to be large.

Future work

Extend the PMH1 and PMH2 to models with intractable likelihoods.
Avoid computing $\tilde{y}_t^{(i)} - y_t$.

References

Yildirim, S., Singh, S.S., Dean, T., and Jasra, A. **Parameter Estimation in Hidden Markov Models with Intractable Likelihoods Using Sequential Monte Carlo**. Journal of Computational and Graphical Statistics, 2014.

Marin, J.M., Pudlo, P., Robert, C.P., and Ryder, R.J. **Approximate Bayesian computational methods**. Statistics and Computing, 22(6), 1167-1180, 2012.

J. Dahlin, T. B. Schön, and M. Villani. **Approximate inference in state space models with intractable likelihoods using Gaussian process optimisation**. Technical Report LiTH-ISY-R-3075, Department of Electrical Engineering, Linköping University, Linköping, Sweden, April 2014.

Summary

Particle Metropolis-Hastings

Metropolis-Hastings with unbiased estimator of the likelihood.

Sequential Monte Carlo

Estimation of log-likelihood, gradients and Hessians.

Approximate Bayesian computations

For inference in models with intractable likelihoods.

Applications

Earthquake counts and Bitcoin volatility modelling.

Thank you for your attention!

Questions, comments and suggestions are most welcome.

Papers and source code are available at: `http://liu.johandahlin.com/`.

References

C. Andrieu and G. O. Roberts. **The pseudo-marginal approach for efficient Monte Carlo computations**. The Annals of Statistics, 37(2):697-725, 2009.

C. Andrieu, A. Doucet, and R. Holenstein. **Particle Markov chain Monte Carlo methods**. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 72(3):269-342, 2010.

M. K. Pitt, R. S. Silva, P. Giordani, and R. Kohn. **On some properties of Markov chain Monte Carlo simulation methods based on the particle filter**. Journal of Econometrics, 171(2):134-151, 2012.

J. Dahlin, F. Lindsten and T. B. Schön, **Particle Metropolis-Hastings using gradient and Hessian information**. Statistics and Computing (MCMSki 2014 special issue), Springer, 2014.

J. Olsson, O. Cappé, R. Douc, and E. Moulines. **Sequential Monte Carlo smoothing with application to parameter estimation in nonlinear state space models**. Bernoulli, 14(1):155-179, 2008.

M. Girolami and B. Calderhead. **Riemann manifold Langevin and Hamiltonian Monte Carlo methods**. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(2):1-37, 2011.

Yildirim, S., Singh, S.S., Dean, T., and Jasra, A. **Parameter Estimation in Hidden Markov Models with Intractable Likelihoods Using Sequential Monte Carlo**. Journal of Computational and Graphical Statistics, 2014.

Marin, J.M., Pudlo, P., Robert, C.P., and Ryder, R.J. **Approximate Bayesian computational methods**. Statistics and Computing, 22(6), 1167-1180, 2012.

J. Dahlin, T. B. Schön, and M. Villani. **Approximate inference in state space models with intractable likelihoods using Gaussian process optimisation**. Technical Report LiTH-ISY-R-3075, Department of Electrical Engineering, Linköping University, Linköping, Sweden, April 2014.