Pseudo-Marginal Hamiltonian Monte Carlo with Efficient Importance Sampling

Kjartan Kloster Osmundsen[1,*], Tore Selland Kleppe[1] and Roman Liesenfeld[2]

[1]Department of Mathematics and Physics, University of Stavanger, Norway
[2]Institute of Econometrics and Statistics, University of Cologne, Germany
[*] Presenting author. E-mail address: kjartan.k.osmundsen@uis.no

Objective

Simulate from target distributions with strong nonlinear dependencies, for which standard Markov chain Monte Carlo (MCMC) methods are often inadequate.

Introduction

This research focuses on the joint posterior of latent variables and parameters in Bayesian hierarchical models, where such nonlinear dependencies can arise.
More specifically, models on the following form:

$$y_t|x_t, \boldsymbol{\theta} \sim g_t(\cdot|x_t, \boldsymbol{\theta}),$$
$$x_t|x_{t-1}, \boldsymbol{\theta} \sim \mathcal{N}(\cdot|\mu_t(x_{t-1}, \boldsymbol{\theta}), \sigma_t^2(x_{t-1}, \boldsymbol{\theta})),$$
$$x_1|\boldsymbol{\theta} \sim \mathcal{N}(\cdot|\mu_1(\boldsymbol{\theta}), \sigma_1^2(\boldsymbol{\theta})).$$

Pseudo-marginal methods target the marginal posteriors of the parameters directly, by Monte Carlo-integrating out the latent variables.
This approach has the potential to produce efficient exploration of the parameters, but relies on the ability to produce an unbiased, low-variance Monte Carlo estimate of the said posterior.

Method

Our approach is to combine **pseudo-marginal Hamiltonian Monte Carlo (HMC)** (Lindsten and Doucet, 2016) with **Efficient Importance Sampling (EIS)** (Richard and Zhang, 2007).

HMC offers the possibility of producing close to iid samples by using the dynamics of a synthetic Hamiltonian (i.e. energy preserving) dynamical system as the proposal mechanism.

Pseudo-marginal HMC substitutes $p(\mathbf{y}|\boldsymbol{\theta})$ for an unbiased Monte Carlo estimate: $E_{\mathbf{u}}(\hat{p}(\mathbf{y}|\boldsymbol{\theta}, \mathbf{u})) = p(\mathbf{y}|\boldsymbol{\theta})\ \forall\ \boldsymbol{\theta}$, where $\mathbf{u}$ is a set of random generated numbers. Here calculated using EIS.

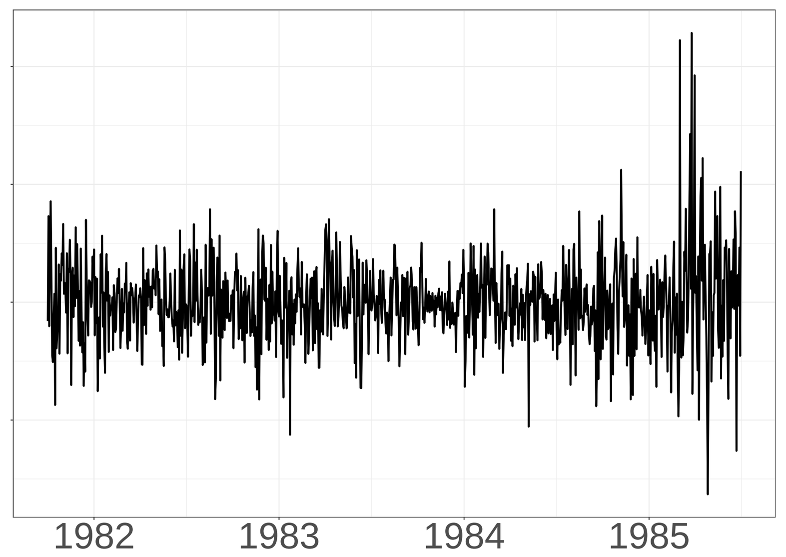
To correct for Monte Carlo variation in $\hat{p}(\mathbf{y}|\boldsymbol{\theta}, \mathbf{u})$, the target distribution must be augmented with $\mathbf{u}$, resulting in the following Hamiltonian:

$$H = -\log p(\boldsymbol{\theta}) - \log \hat{p}(\mathbf{y}|\boldsymbol{\theta}, \mathbf{u}) + \frac{1}{2}\mathbf{u}^T\mathbf{u} + \frac{1}{2}\mathbf{p}_{\boldsymbol{\theta}}^T \mathcal{M}^{-1}\mathbf{p}_{\boldsymbol{\theta}} + \frac{1}{2}\mathbf{p}_{\mathbf{u}}^T\mathbf{p}_{\mathbf{u}}$$

Stan, a Bayesian modeling software based on HMC, is used as a benchmark for the performance of our method.

Models and observations

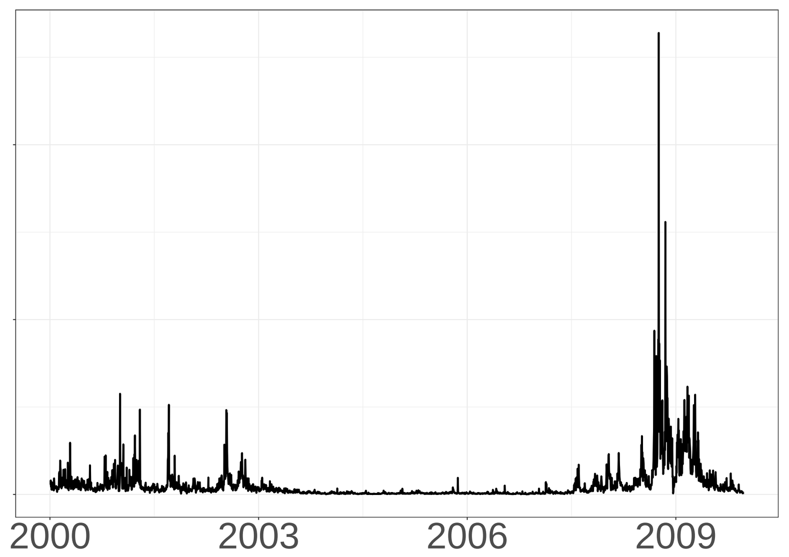
Dollar/Pound exchange rate (log-returns)



Stochastic volatility model

$$g_t = \mathcal{N}(0, \exp(x_t))$$
$$\mu_t = \gamma + \delta x_{t-1}$$
$$\sigma_t = v$$

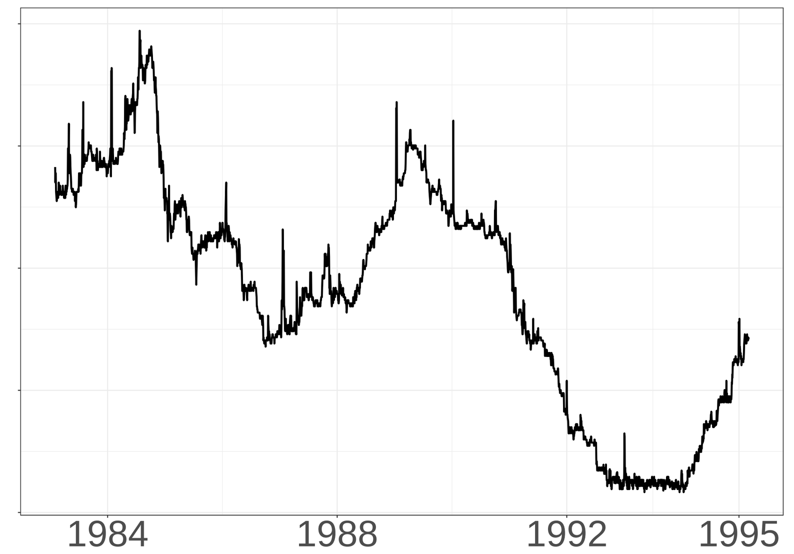
Variance of American Express stock price



Gamma model

$$g_t = \Gamma(1/\tau, \beta\tau\exp(x_t))$$
$$\mu_t = \phi x_{t-1}$$
$$\sigma_t = \sigma$$

Eurodollar deposit spot rates



Constant elasticity of variance diffusion model

$$g_t = \mathcal{N}(x_t, \sigma_y^2)$$
$$\mu_t = x_{t-1} + \Delta(\alpha - \beta x_{t-1})$$
$$\sigma_t = \sigma_x x_{t-1}^{\gamma}\sqrt{\Delta}$$

Results

The following table shows the Effective Sample Size per second for each model, as an average of each parameter in $\boldsymbol{\theta}$:

	SV model	Gamma model	CEV model
HMC-EIS	1.8	1.3	1.9
Stan	30	1.8	0

Conclusion

Hamiltonian Monte Carlo with Efficient Importance sampling produces stable and accurate results. The algorithm is able to produce near perfect samples, and the computational cost is competitive for the more advanced models.