Econometrics and Statistics 000 (2016) 1–11

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

On efficient Bayesian inference for models with stochastic volatility

D.K. Sakaria, J.E. Griffin*

School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, CT2 7NF, UK

ARTICLE INFO

Article history:
Received 19 March 2016
Revised 1 August 2016
Accepted 3 August 2016
Available online xxx

Keywords:
Stochastic volatility
Bayesian methods
Markov chain Monte Carlo
Mixture offset representation

ABSTRACT

An efficient method for Bayesian inference in stochastic volatility models uses a linear state space representation to define a Gibbs sampler in which the volatilities are jointly updated. This method involves the choice of an offset parameter and we illustrate how its choice can have an important effect on the posterior inference. A Metropolis–Hastings algorithm is developed to robustify this approach to choice of the offset parameter. The method is illustrated on simulated data with known parameters, the daily log returns of the Eurostoxx index and a Bayesian vector autoregressive model with stochastic volatility.

1. Introduction

It is known that the volatility of many economic variables vary over time. Initial work on time-varying volatility often considered asset price returns. Over the long term, the volatility of equity returns may appear to be stable but usually there are periods of high volatility and calm market periods when the volatility may be low (Enders, 2004). Several approaches have been developed to model this time-varying volatility. In ARCH and GARCH models (Engle, 1982; Bollerslev, 1986), the volatility is modeled as a function of the lagged values of the asset returns and the volatility. Alternatively, stochastic volatility models assume that the volatility follows a known stochastic process such as an AR process for the logarithm of volatility (see e.g., Harvey and Shephard, 1996).

In this paper, we will concentrate on the Bayesian estimation of stochastic volatility models (see e.g., Jacquier et al., 1994; Kim et al., 1998; Chib et al., 2002). The asset returns may be expressed as functions of past returns or other economic variables and the log volatility is modeled as a separate AR process. A simple stochastic volatility model assumes that y_t, the log return at time t, can be expressed as

$$\begin{aligned} y_t &= e^{h_t/2}\nu_t, \qquad t = 1, \ldots, T \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \sigma_\eta \eta_t. \end{aligned} \tag{1}$$

where ν_t and η_t are independent error terms for which $\nu_t \overset{i.i.d.}{\sim} \mathrm{N}(0, 1)$ and $\eta_t \overset{i.i.d.}{\sim} \mathrm{N}(0, 1)$, and h_t is the log volatility at time t. The model assumes that the log volatility h_t follows an AR(1) process with parameters μ, ϕ, and σ_η.

Bayesian inference is complicated since this is a non-linear state space model. Several Markov chain Monte Carlo (MCMC) methods have been developed to sample this class of models. Jacquier et al. (1994) used one-at-a-time updating of h_t

* Corresponding author.
E-mail address: J.E.Griffin-28@kent.ac.uk (J.E. Griffin).

http://dx.doi.org/10.1016/j.ecosta.2016.08.002

with a carefully chosen proposal distribution (one-at-a-time or single move updating is often criticised for highly correlated samples). Samplers which update a block of h_t's often lead to better mixing. For example, Jensen and Maheu (2014) propose to update a block of h_t in an asymmetric, nonparametric stochastic volatility model.

The model (1) can be expressed in linear state space form for h_t using transformed data $\log y_t^2 = h_t + \log \nu_t^2$. Kim et al. (1998) (KSC) approximate the distribution of $\log \nu_t^2$ using a normal mixture distribution leading to a Gaussian linear state space form for h_t conditional on the mixture states for each observation. This allows the volatilities $h_1, \ldots, h_T$ to be updated using Forward Filtering Backward Sampling (FFBS) techniques (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994). In order to make the approximation robust for small values of y_t, a small offset parameter c is used and $\log(y_t^2 + c)$ is used in place of $\log y_t^2$ as the transformed data. This leads to samples from an approximate posterior distribution for the parameters of the SV model and $h_1, \ldots, h_T$. KSC suggest an importance sampling scheme for estimating posterior quantities using the approximate posterior as the importance sampling distribution. However, as with any importance sampler, results can become biased if the importance sampling distribution (the approximation) is sufficiently different to the actual posterior distribution. This can be the case if c is poorly chosen. The approach has been developed in various directions. Chib et al. (2002) consider models with Student-t distributed innovation, exogeneous variables and jumps in observations. Omori and Watanabe (2008) consider an asymmetric stochastic volatility model and allow correlation between the returns and the volatility which allows the modelling for the leverage effect. A multivariate normal approximation is used to express the model in a linear state space form with Gaussian errors. Results show a better performance compared to a single move sampler. More recently, Kastner and Frühwirth-Schnatter (2014) developed centring methods.

More recently, the KSC sampler has been applied to more complicated models. For example, Belmonte et al. (2013) consider dynamic regression models with stochastic volatility

$$
\begin{aligned}
y_t &= X_t \beta_t + e^{h_t/2} \nu_t, \qquad t = 1, \ldots, T \\
h_t &= \mu + \phi(h_{t-1} - \mu) + \sigma_\eta \eta_t.
\end{aligned} \tag{2}
$$

Clark (2012) builds a vector autoregressive model with stochastic volatility, which will be further considered in this paper. Let y_t be a $(p \times 1)$-dimensional vector of economics variables and x_t be a $(q \times 1)$-dimensional vector of deterministic variables measured at time t. The data modelled as

$$
\Pi(L)(y_t - \Psi x_t) = \epsilon_t \tag{3}
$$

where Ψ is a $(p \times q)$-dimensional vector of coefficients, $\Pi(L) = I_p - \Psi_1 L - \Psi_2 L^2 \ldots \Psi_k L^k$ is a lag polynomial and ν_t are independent errors. The errors ϵ_t are modelled using a factor stochastic volatility model. Let A be a lower triangular matrix with 1's on the diagonal then

$$
\begin{aligned}
\epsilon_t &= A^{-1} \Lambda_t^{0.5} \nu_t, \qquad \nu_t \sim \mathrm{N}(0, I_p), \\
\Lambda_t &= \mathrm{diag}(e^{h_{1,t}}, e^{h_{2,t}}, \ldots, e^{h_{p,t}}), \\
h_{i,t} &= h_{i,t-1} + \sigma_{\eta,i} \eta_{i,t}, \qquad \eta_{i,t} \overset{iid}{\sim} \mathrm{N}(0,1) \qquad \forall i = 1, 2, \ldots, p.
\end{aligned}
$$

Bayesian inference is made using a Gibbs sampler and the volatilities are updated using the KSC method in the approximate model (i.e., using $r_t^\star = \log((y_t - X_t \beta_t)^2 + c)$ for the dynamic regression model or $r_t^\star = A\Psi(L)(y_t - \Psi x_t) + c$ for the vector autoregression model) but other parameters (such as β_t) are updated using the correct (rather than the approximate) stochastic volatility model. Although, this seems to have little effect on inference, the Gibbs sampler is not properly specified. In addition, in these models, the effect of c is harder to understand since the scale of $r_t^\star$ can change substantially between iterations.

This paper makes two main contributions. Firstly, we develop an MCMC framework for sampling from the posterior distribution of the SV model (rather than an approximation to the SV model) using the KSC method as a proposal in a Metropolis–Hastings step for updating the volatilities. Secondly, we introduce a method for specifying the offset parameter using standardisation that robustifies the MCMC algorithm to the scale of the data.

The paper initially considers the problem of sampling the time-varying volatilities in the stochastic volatility model in (1) and considers more complicated models in the examples. The remainder of the paper is organised as follows. Section 2 describes the Kim et al. (1998) method to linearise the log volatility model and the difficulty of using an appropriate value of the offset parameter c is highlighted. In Section 3, a standardisation method is introduced and a Metropolis–Hastings (M–H) step is described to propose volatility parameter $h_1, \ldots, h_T$ using Forward Filtering Backward Sampling (FFBS). Results using simulated data, Eurostoxx daily log returns and a vector autoregressive model with stochastic volatility are discussed in Section 4. Section 5 concludes.

2. Linear state space method

In this section, we review the sampling method of KSC and illustrate the effect of choosing c. KSC suggest transforming the observations in the SV model in (1) so that

$$
\log\, y_t^2 = h_t + \log\, \nu_t^2, \tag{4}
$$