Macroeconometrics

Lecture 21 Bayesian estimation of SV models using auxiliary mixture

Tomasz Woźniak

Department of Economics
University of Melbourne

A simple Stochastic Volatility model

Auxiliary mixture

Bayesian estimation of SV models

Bayesian estimation of SV–AR models

Introduction to heteroskedastic models

Compulsory readings:

Woźniak (2021) Bayesian estimation of simple SV models using auxiliary mixture, Lecture notes

Objectives.

- To present Gibbs sampling for Stochastic Volatility models
- To introduce auxiliary mixture technique
- To use an inverse probability transform sampling method

Learning outcomes.

- Applying log-linearisation for feasible computations
- Transforming a non-linear model to a Gaussian state-space specification
- Applying a normal mixture approximation of a distribution for a real-valued random variable

A simple Stochastic Volatility model

A simple Stochastic Volatility model

A model with a conditional mean specification.

Let $\mu_t(\alpha)$ denote a conditional mean of y_t that is a function of a parameter (vector) α.

$$y_t = \mu_t(\alpha) + \exp\left\{\frac{1}{2}h_t\right\}\epsilon_t$$

$$y_t - \mu_t(\alpha) = \exp\left\{\frac{1}{2}h_t\right\}\epsilon_t$$

$$\downarrow$$

$$y_{\mu.t} = \exp\left\{\frac{1}{2}h_t\right\}\epsilon_t$$

$$h_t = h_{t-1} + \sigma_v v_t$$

$$\epsilon_t \sim \mathcal{N}(0, 1)$$

$$v_t \sim \mathcal{N}(0, 1)$$

h_0 – estimated parameter of the model

Log-linearization of the measurement equation

Perform the log-linearization of the measurement equation by:

taking the square of both sides of the equation

taking the logarithm of both sides of the equation

$$y_{\mu.t} = \exp\left\{\frac{1}{2}h_t\right\}\epsilon_t$$

$$\log y_{\mu.t}^2 = h_t + \log \epsilon_t^2$$

$$\tilde{y}_t = h_t + \tilde{\epsilon}_t$$

$$\tilde{\epsilon}_t \sim \log \chi_1^2$$

Matrix notation

A simple Stochastic Volatility model

$$\tilde{y} = h + \tilde{\epsilon}$$

$$Hh = h_0 e_{1.T} + \sigma_v v$$

$$\tilde{\epsilon}_t \sim iid \log \chi^2_1$$

$$v \sim \mathcal{N}_T\left(\mathbf{0}_T, I_T\right)$$

Define the following $T \times 1$ matrices

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_T \end{bmatrix} \quad h = \begin{bmatrix} h_1 \\ \vdots \\ h_T \end{bmatrix} \quad \tilde{\epsilon} = \begin{bmatrix} \tilde{\epsilon}_1 \\ \vdots \\ \tilde{\epsilon}_T \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_T \end{bmatrix} \quad e_{1.T} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Auxiliary mixture

Auxiliary mixture

Approximate the $\log \chi_1^2$ distribution by a mixture of ten normal distributions given by:

$$\log \chi_1^2 \approx \sum_{m=1}^{10} Pr(s_t = m)\mathcal{N}\left(\mu_m, \sigma_m^2\right)$$

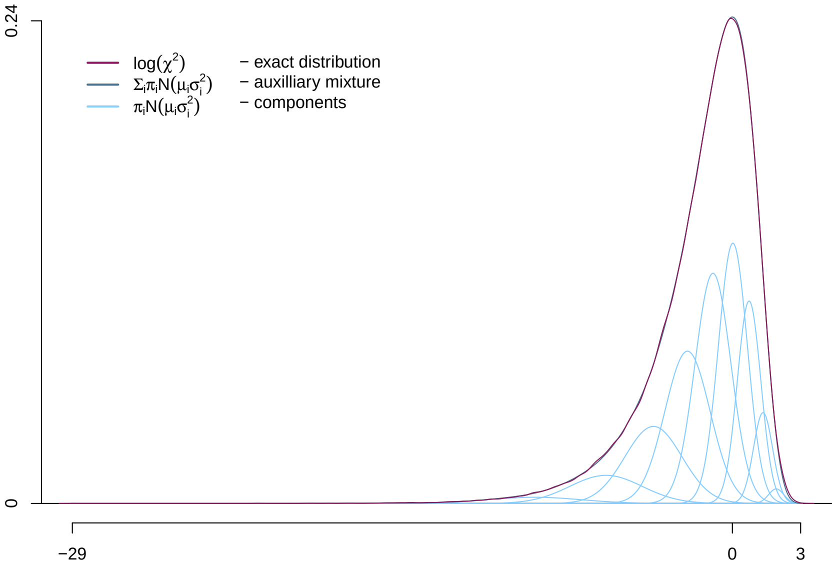
$s_t \in \{1, \dots, 10\}$ is a discrete-valued random indicator of the mixture component

$\mu_m, \sigma_m^2, Pr(s_t = m)$ are predetermined

Auxiliary mixture

m	$Pr(s_t = m)$	μ_m	σ_m^2
1	0.00609	1.92677	0.11265
2	0.04775	1.34744	0.17788
3	0.13057	0.73504	0.26768
4	0.20674	0.02266	0.40611
5	0.22715	-0.85173	0.62699
6	0.18842	-1.97278	0.98583
7	0.12047	-3.46788	1.57469
8	0.05591	-5.55246	2.54498
9	0.01575	-8.68384	4.16591
10	0.00115	-14.65000	7.33342

Auxiliary mixture



Auxiliary mixture

Conditional distribution of $\tilde{\epsilon}_t$

$$\tilde{\epsilon}_t | s_t = m \sim \mathcal{N}\left(\mu_m, \sigma_m^2\right)$$

A simple Stochastic Volatility model

$$\tilde{y} = h + \tilde{\epsilon} \tag{1}$$

$$Hh = h_0 e_{1.T} + \sigma_v v \tag{2}$$

$$\tilde{\epsilon} | s \sim \mathcal{N}_T\left(\mu_s, \mathrm{diag}\left(\sigma_s^2\right)\right) \tag{3}$$

$$v \sim \mathcal{N}_T\left(\mathbf{0}_T, I_T\right) \tag{4}$$

$$s = \begin{bmatrix} s_1 & \dots & s_T \end{bmatrix}' \quad \mu_s = \begin{bmatrix} \mu_{s_1} & \dots & \mu_{s_T} \end{bmatrix}' \quad \sigma_s^2 = \begin{bmatrix} \sigma_{s_1}^2 & \dots & \sigma_{s_T}^2 \end{bmatrix}'$$

Bayesian estimation of SV models

Prior distributions

Hierarchical prior structure is given by:

$$p\left(h, s, h_0, \sigma_v^2\right) = p\left(h|h_0, \sigma_v^2\right) p\left(h_0\right) p\left(\sigma_v^2\right) p\left(s\right)$$

Eqs. (2) and (4) determine a conditional prior distribution of h:

$$\begin{aligned} p\left(h|h_0, \sigma_v^2\right) &\sim \mathcal{N}_T\left(h_0 H^{-1} e_{1.T}, \sigma_v^2 (H'H)^{-1}\right) \\ &\propto \det\left(\sigma_v^2 I_T\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\frac{1}{\sigma_v^2}(Hh - h_0 e_{1.T})'(Hh - h_0 e_{1.T})\right\} \end{aligned}$$

whereas the marginal prior distributions for h_0, σ_v^2, and s are:

$$\begin{aligned} p(h_0) &= \mathcal{N}(0, \underline{\sigma}_h^2) \\ p\left(\sigma_v^2\right) &= \mathcal{IG}2(\underline{s}, \underline{\nu}) \\ p\left(s_t\right) &= \mathcal{M}ultinomial\left(\{m\}_{m=1}^{10}, \{Pr(s_t = m)\}_{m=1}^{10}\right) \end{aligned}$$

Full-conditional posterior distributions

Full-conditional posterior distribution for h

Equations (1) and (3) determine the conditional likelihood that is proportional to

$$\exp\left\{-\frac{1}{2}(h-(\tilde{y}-\mu_s))'\text{diag}\left(\sigma_s^2\right)^{-1}(h-(\tilde{y}-\mu_s))\right\}$$

which leads to:

$$h|y,s,h_0,\sigma_v^2 \sim \mathcal{N}_T\left(\overline{h},\overline{V}_h\right)$$

$$\overline{V}_h = \left[\text{diag}\left(\sigma_s^2\right)^{-1} + \sigma_v^{-2}H'H\right]^{-1}$$

$$\overline{h} = \overline{V}_h\left[\text{diag}\left(\sigma_s^2\right)^{-1}(\tilde{y}-\mu_s) + \sigma_v^{-2}h_0 e_{1.T}\right]$$

Full-conditional posterior distributions

Full-conditional posterior distribution for s

is a multinomial distribution with the probabilities proportional to

$$\omega_{m.t} = Pr[s_t = m] p\left(\tilde{y}_t | h_t, s_t = m\right)$$

for $m = 1, \dots, 10$, $p\left(\tilde{y}_t | h_t, s_t = m\right)$ is based on eqs (1) & (3)

For each t and m obtain $\omega_{m.t}$ using parallel computations and compute the probabilities of the multinomial full conditional posterior distribution by

$$Pr[s_t = m | \tilde{y}_t, h_t] = \frac{\omega_{m.t}}{\sum_{i=1}^{10} \omega_{i.t}}$$

Sampling s_t independently for each t is straightforward

Full-conditional posterior distributions

Full-conditional posterior distributions for σ_v^2 and h_0

It is straightforward to show that:

$$\begin{aligned}
\sigma_v^2|y, s, h, h_0 &\sim \mathcal{IG}2\left(\overline{s}, \overline{\nu}\right) \\
\overline{\nu} &= \underline{\nu} + T \\
\overline{s} &= \underline{s} + (Hh - h_0 e_{1.T})'(Hh - h_0 e_{1.T})
\end{aligned}$$

$$\begin{aligned}
h_0|y, s, h, \sigma_v^2 &\sim \mathcal{N}\left(\overline{h}_0, \overline{\sigma}_h^2\right) \\
\overline{\sigma}_h^2 &= \left(\underline{\sigma}_h^{-2} + \sigma_v^{-2}\right)^{-1} \\
\overline{h}_0 &= \overline{\sigma}_h^2\left(\sigma_v^{-2} e' Hh\right)
\end{aligned}$$

Gibbs sampler

Initialize $h^{(0)}$, $s^{(0)}$, and $\sigma_v^{2(0)}$

For $i = 1, \dots, S$

Draw $h_0^{(i)} \sim \mathcal{N}\left(\overline{h}_0, \overline{\sigma}_h^2\right)$

Draw $\sigma_v^{2(i)} \sim \mathcal{IG}2\left(\overline{s}, \overline{\nu}\right)$

Draw $s_t^{(i)} \sim \mathcal{M}ultinomial\left(\{m\}_{m=1}^{10}, \{Pr[s_t = m|\tilde{y}_t, h_t^{(i)}]\}_{m=1}^{10}\right)$
for t=1,. . . ,T using inverse transform method

Draw $h^{(i)} \sim \mathcal{N}_T\left(\overline{h}, \overline{V}_h\right)$ using precision sampler

Return a sample drawn from the posterior distribution:

$$\left\{h^{(i)}, s^{(i)}, h_0^{(i)}, \sigma_v^{2(i)}\right\}_{i=1}^{S}$$

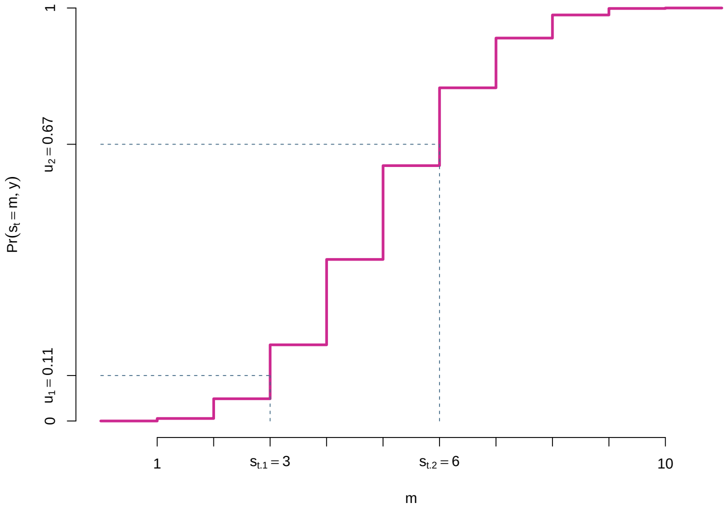
Simulation smoother

The precision matrix of the T-variate normal full-conditional posterior distribution $\overline{V}_h^{-1}$ is a tridiagonal matrix.

Draw random numbers from this distribution using the simulation smoother and computer routines for band or tridiagonal matrices.

Sampling random draws from Multinomial distribution

Inverse transform method



Bayesian estimation of the SV–AR models

Bayesian estimation of the SV–AR models

An Autoregressive Stochastic Volatility model

$$
\begin{aligned}
y_t &= \exp\left\{\frac{1}{2}h_t\right\}\epsilon_t \\
h_t &= \mu_0 + \alpha h_{t-1} + \sigma_v v_t \\
\epsilon_t &\sim \mathcal{N}(0,1) \\
v_t &\sim \mathcal{N}(0,1)
\end{aligned}
$$

$|\alpha| < 1$ – stationarity condition

Bayesian estimation of the SV–AR models

An Autoregressive Stochastic Volatility model

$$\tilde{y} = h + \tilde{\epsilon}$$

$$H_\alpha h = \mu_0 \imath_T + \alpha h_0 e_{1.T} + \sigma_v v$$

$$h = \mu_0 \imath_T + \alpha h_{-1} + \sigma_v v$$

$$\tilde{\epsilon} | s \sim \mathcal{N}_T\left(\mu_s, \operatorname{diag}\left(\sigma_s^2\right)\right)$$

$$v \sim \mathcal{N}_T\left(\mathbf{0}_T, I_T\right)$$

$$h_{-1} = \begin{bmatrix} h_0 & \dots & h_{T-1} \end{bmatrix}' \quad \imath_T = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}' \quad H_\alpha = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ -\alpha & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\alpha & 1 \end{bmatrix}$$

Prior distributions

Hierarchical prior structure is given by:

$$p\left(h, s, h_0, \mu_0, \alpha, \sigma_v^2\right) = p\left(h|\mu_0, \alpha, h_0, \sigma_v^2\right) p\left(\mu_0\right) p\left(\alpha\right) p\left(h_0\right) p\left(\sigma_v^2\right) p\left(s\right)$$

Conditional prior distribution of h:

$$p\left(h|\mu_0, \alpha, h_0, \sigma_v^2\right) \sim \mathcal{N}_T\left(\mu_0 H_\alpha^{-1}\imath_T + \alpha h_0 H_\alpha^{-1} e_{1.T}, \sigma_v^2 (H_\alpha' H_\alpha)^{-1}\right)$$

$$\propto \det\left(\sigma_v^2 I_T\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\frac{1}{\sigma_v^2}(Hh - \mu_0\imath_T - h_0 e_{1.T})'(Hh - \mu_0\imath_T - h_0 e_{1.T})\right\}$$

Marginal prior distributions for $\mu + 0, \alpha, h_0, \sigma_v^2$, and s:

$$\begin{aligned}
p(\mu_0) &= \mathcal{N}(\underline{\mu_0}, \underline{\sigma}_\mu^2) \\
p(\alpha) &= \mathcal{N}(\underline{\alpha}, \underline{\sigma}_\alpha^2)\mathcal{I}(|\alpha| < 1) \\
p(h_0) &= \mathcal{N}(0, \underline{\sigma}_h^2) \\
p\left(\sigma_v^2\right) &= \mathcal{IG}2(\underline{s}, \underline{\nu}) \\
p\left(s_t\right) &= \mathcal{M}ultinomial\left(\{m\}_{m=1}^{10}, \{Pr(s_t = m)\}_{m=1}^{10}\right)
\end{aligned}$$

Full-conditional posterior distributions

Full-conditional posterior distribution for h

Equations (1) and (3) determine the conditional likelihood that combined with the conditional prior distribution gives:

$$h|y, s, \mu_0, \alpha, h_0, \sigma_v^2 \sim \mathcal{N}_T\left(\overline{h}, \overline{V}_h\right)$$

$$\overline{V}_h = \left[\operatorname{diag}\left(\sigma_s^2\right)^{-1} + \sigma_v^{-2} H_\alpha' H_\alpha\right]^{-1}$$

$$\overline{h} = \overline{V}_h\left[\operatorname{diag}\left(\sigma_s^2\right)^{-1}(\tilde{y} - \mu_s) + \sigma_v^{-2} H_\alpha'\left(\mu_0 \imath_T + \alpha h_0 e_{1.T}\right)\right]$$

Full-conditional posterior distribution for s

Neither the form of the measurement equation or the multinomial prior distribution changes, therefore, the sampler stays the same.

Full-conditional posterior distributions

Full-conditional posterior distribution for μ_0

$$\mu_0|y, h, \alpha, h_0, \sigma_v^2 \sim \mathcal{N}\left(\overline{\mu}_0, \overline{\sigma}_\mu^2\right)$$

$$\overline{\sigma}_\mu^2 = \left[\underline{\sigma}_\mu^{-2} + T\sigma_v^{-2}\right]^{-1}$$

$$\overline{\mu}_0 = \overline{\sigma}_\mu^2\left[\underline{\mu}_0\underline{\sigma}_\mu^{-2} + \sigma_v^{-2}\left(H_\alpha h - \alpha h_0 e_{1.T}\right)' \imath_T\right]$$

Full-conditional posterior distribution for α

$$\alpha|y, h, \mu_0, h_0, \sigma_v^2 \sim \mathcal{N}\left(\overline{\alpha}, \overline{\sigma}_\alpha^2\right)\mathcal{I}(|\alpha| < 1)$$

$$\overline{\sigma}_\alpha^2 = \left[\underline{\sigma}_\alpha^{-2} + \sigma_v^{-2} h_{-1}' h_{-1}\right]^{-1}$$

$$\overline{\alpha} = \overline{\sigma}_\alpha^2\left[\underline{\alpha}\underline{\sigma}_\alpha^{-2} + \sigma_v^{-2} h_{-1}'\left(h - \mu_0 \imath_T\right)\right]$$

Full-conditional posterior distributions

Full-conditional posterior distribution for h_0

$$h_0|y, h, \alpha, \mu_0, \sigma_v^2 \sim \mathcal{N}\left(\overline{h}_0, \overline{\sigma}_h^2\right)$$

$$\overline{\sigma}_h^2 = \left[\underline{\sigma}_h^{-2} + \alpha^2 \sigma_v^{-2}\right]^{-1}$$

$$\overline{h}_0 = \overline{\sigma}_h^2 \left[\underline{h}_0 \underline{\sigma}_h^{-2} + \alpha h_1 \sigma_v^{-2}\right]$$

Full-conditional posterior distribution for σ_v^2

$$\sigma_v^2|y, h, \mu_0, \alpha, h_0 \sim \mathcal{IG}2\left(\overline{s}, \overline{\nu}\right)$$

$$\overline{s} = \underline{s} + \left(H_\alpha h - \mu_0 \imath_T - \alpha h_0 e_{1.t}\right)' \left(H_\alpha h - \mu_0 \imath_T - \alpha h_0 e_{1.t}\right)$$

$$\overline{\nu} = \overline{\nu} + T$$

Gibbs sampler

Initialize $h^{(0)}$, $s^{(0)}$, $\mu_0^{(0)}$, $\alpha^{(0)}$, and $\sigma_v^{2(0)}$

For $i = 1, \dots, S$

Draw $h_0^{(i)} \sim \mathcal{N}\left(\overline{h}_0, \overline{\sigma}_h^2\right)$

Draw $\sigma_v^{2(i)} \sim \mathcal{IG}2\left(\overline{s}, \overline{\nu}\right)$

Draw $\mu_0^{(i)} \sim \mathcal{N}\left(\overline{\mu}_0, \overline{\sigma}_\mu^2\right)$

Draw $\alpha^{(i)} \sim \mathcal{N}\left(\overline{\alpha}, \overline{\sigma}_\alpha^2\right)$

Draw $s_t^{(i)} \sim \mathcal{M}ultinomial\left(\{m\}_{m=1}^{10}, \{Pr[s_t = m|\tilde{y}_t, h_t^{(i)}]\}_{m=1}^{10}\right)$
for $t = 1, \dots, T$ using inverse transform method

Draw $h^{(i)} \sim \mathcal{N}_T\left(\overline{h}, \overline{V}_h\right)$ using precision sampler

Return a sample drawn from the posterior distribution:

$$\left\{h^{(i)}, s^{(i)}, \mu_0^{(i)}, \alpha^{(i)}, h_0^{(i)}, \sigma_v^{2(i)}\right\}_{i=1}^{S}$$

Introduction to heteroskedastic models

Introduction to heteroskedastic models

A linear regression with Stochastic Volatility

$$
\begin{aligned}
Y &= X\beta + E \\
E \mid X &\sim \mathcal{N}_T\left(\mathbf{0}_T, \operatorname{diag}\left(\sigma^2\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2 &= (\exp\{h_1\}, \dots, \exp\{h_T\}) \\
h_t &- \text{follows a Stochastic Volatility process}
\end{aligned}
$$

$$
\beta \sim \mathcal{N}\left(\underline{\beta}, \underline{V}_\beta\right)
$$

A model with conditional heteroskedasticity

- Improves the precision of the estimation of β
- Improves the in-sample fit of the model
- Greatly improves the forecasting performance of the model

Full conditional posterior distribution of β

$$\beta \mid Y, X, h \sim \mathcal{N}\left(\overline{\beta}, \overline{V}_{\beta}\right)$$

$$\overline{V}_{\beta} = \left[\underline{V}_{\beta}^{-1} + X'\text{diag}\left(\sigma^2\right)^{-1} X\right]^{-1}$$

$$\overline{\beta} = \overline{V}_{\beta}\left[\underline{V}_{\beta}^{-1}\underline{\beta} + X'\text{diag}\left(\sigma^2\right)^{-1} Y\right]$$

Conditionally on h the remaining model parameters can be sampled from their respective full conditional posterior distributions that take into account the conditional variances of the error term

Bayesian estimation of SV models

An efficient and computationally fast method of estimating SV models is Bayesian Gibbs sampler

Dedicated computational techniques include

auxiliary mixture
simulation smoother
operations on tridiagonal matrices
inverse transform method

The algorithm presented in this lecture is applicable to

univariate models with SV conditional heteroskedasticity
multivariate models with independent SV processes

Adapt your Gibbs sampler to use function `SV.Gibbs.iteration` from file `00 SV codes.R`