Assignment Week 5 Time Series Econometrics

Rose Barzilai | 2596415 | r.n.barzilai@student.vu.nl
Zuzana Leova | 2591033 | z.leova@student.vu.nl
Charlotte Taman | 2100584 | c.e.j.taman@student.vu.nl
Femke Vedder | 2555497 | f.m.l.vedder@student.vu.nl

Group 1

March 12 2018

1 Assignment 5.1: SV model (continued)

We have the SV model given by:

$$
\begin{aligned}
y_t &= \mu + \sigma_t \epsilon_t, & \epsilon_t &\sim \mathcal{N}(0,1) \\
\log \sigma_t^2 &= \xi + H_t, & & \\
H_{t+1} &= \phi H_t + \sigma_\theta \theta_t, & \eta_t &\sim \mathcal{N}(0,1).
\end{aligned}
$$

(f) Compute the smoothed mean of H_t based on the the SV model for y_t by using the importance sampling method.

Firstly, we apply SPDK that we had already implemented in the assignment of Week 5. In the previous section, we have computed the smoothed mode h_t based on the approximate model by using Kalman filter and smoother. We will now use this estimate to obtain approximating model with normal density $g(Y_n)$

The smoothed estimate of x is denoted by $\overline{x}$.

$$
\overline{x} = \mathbb{E}[x(\theta)|Y_n] = \int x(\theta) p(\theta|Y_n) d\theta \tag{1}
$$

The conditional density $p(\alpha|Y_n)$ depends on an unknown parameter vector ψ (in our case estimated in the Assignment of week 3). In practice, the explicit expressions are not available for the density $p(\theta|Y_n)$, we will use importance sampling to find an approximate density for $p(\theta|Y_n)$ and for which random draws are available, so we can sample from it, making an appropriate adjustment to the integral 1 In order to do this we need to introduce the Gaussian importance density $g(\theta|Y_n)$, which will resemble $p(\theta|Y_n)$ as closely as possible. Random draws are available from this density and we can sample from this, when making an appropriate adjustment to the integral in (1).
Following the equations derived in section **11.2** of the Time Series Analysis by State Space Methods book (shortly DK book), we know that

$$
\overline{x} = \frac{\mathbb{E}[x(\alpha) w^*(\theta, Y_n)]}{\mathbb{E}[w^*(\theta, Y_n)]}, \qquad \text{where} \qquad w^*(\theta, Y_n) = \frac{p(Y_n|\theta)}{g(Y_n|\theta)}. \tag{2}
$$

The Monte Carlo estimate $\hat{x}$ of the mean is then given as

$$
\hat{x} = \frac{\sum_{i=1}^N x_i w_i}{\sum_{i=1}^N w_i}, \qquad \text{where} \qquad x_i = x(\theta^{(i)}) \text{ and } \qquad w_i = w(\theta^{(i)}, Y_n). \tag{3}
$$

The first step of the SPDK approach is to construct the approximate model $g(\cdot)$. We obtain $\hat{\theta} = \mathbb{E}[\theta|Y_n]$ from the original model $p(.)$ through mode estimation, as we did in the previous section. We use this $\hat{\theta}$ to obtain A (which is equal to H) and y^* (which is equal to z of the formula for smoothed mode from week 4) once more. Both A and y^* depend on Y_n, since we use $\hat{\theta}$ to obtain them. As mentioned in section **14.5.2** of **DK** book, if an observation is very close to zero or equal to zero, it should be replaced by a small constant value to avoid numerical problems. Therefore if y is very small we replace it by 0.001 before mode computation by KFS

Now the approximating model of $g(Y_n, \theta) \equiv g(y^*, \theta)$ is $y^* = \theta + u$ where $u \sim N(0, A)$. This model is linear Gaussian.

In order to compute weights, we need to perform draw simulation smoothing of $\theta^{(i)} \sim g(\theta|Y_n)$, which we used in week 3 of the course. Because the importance density $g(\theta|Y_n) \equiv g(\theta|y^*)$ now comes from a linear Gaussian model we can apply the following formula to draw $\tilde{\theta}^{(i)}$ from conditional density by simulation:

$$\tilde{\theta}^{(i)} = \hat{\theta} + \theta^{+,(i)} - \hat{\theta}^{+,(i)} \tag{4}$$

Here we used the KFS method to obtain the $\hat{\theta}$ and the $\hat{\theta}^{+}$, where the first is the smoothed mean from the approximating Gaussian model ($\hat{\theta} = \mathbb{E}[\theta|y^*]$) and the second the smoothed mean from the simulated linear Gaussian model ($\hat{\theta}^{+} = \mathbb{E}[\theta|y^{+}]$). In order to draw $\theta^{+,(i)}$ we needed to simulate the innovations for $\eta_t^{+} \sim N(0, Q_t)$ and $u_t^{+} \sim N(0, A_t)$ for $t = 1, \ldots, n$. a_1 is initialized as $a_1 = 0$ and P_1 is initialized as $P_1 = \frac{\sigma_\eta^2}{1-\phi^2}$. We have recursively computed α_{t+1}^{+}, θ_t^{+} and y_t^{+} and then repeat the whole process to simulate N=100 draws of $\tilde{\theta}^{(i)}$.
For the numerical stability, we worked with $m_i = \log p^{(i)} - \log g^{(i)}$ for weights computation, therefore the formula for the mode computation $\hat{x}$ (3) changes to

$$\hat{x} = \frac{\sum_{i=1}^{N} x_i \exp\left(m_i - \overline{m}\right)}{\sum_{i=1}^{N} \exp\left(m_i - \overline{m}\right)}, \tag{5}$$

where $\overline{m}$ is mean $\frac{\sum_{i=1}^{N} m_i}{N}$. For this we need to evaluate likelihood of densities $\log p(Y_n|\theta^{(i)})$ and $\log g(Y_n|\theta^{(i)})$. We have implemented it using the following formula, stated in section **10.6.5** of **DK**

$$\log p\left(y_t|\tilde{\theta}_t{}^{(i)}\right) = -\frac{1}{2}[\log 2\pi\sigma^2 + \tilde{\theta}_t{}^{(i)} + (y_t - c)^2 \exp(-\tilde{\theta}_t{}^{(i)})], \tag{6}$$

where σ is defined as in section $\exp(\frac{1}{2}\omega/(1-\phi))$ and $c = 0$. Because the density $\log g(Y_n|\theta^{(i)})$ is a Gaussian approximation of original density, to compute it we used a standard formula for likelihood function of normal distribution with mean $\theta^{(i)}$ and variance A_t to compute it. So the final formula looks like

$$\log g\left(y_t^*|\tilde{\theta}_t{}^{(i)}\right) = -\frac{1}{2}\log 2\pi A_t - \frac{1}{2}\frac{\left(y_t^* - c - \tilde{\theta}_t{}^{(i)}\right)^2}{A_t}. \tag{7}$$

The final densities $\log p(Y_n|\tilde{\theta}^{(i)})$ and $\log g(Y_n|\tilde{\theta}^{(i)})$ are then computed as $\log p(Y_n|\tilde{\theta}^{(i)}) = \sum_{t=1}^{T} \log p(y_t|\tilde{\theta}_t{}^{(i)})$ and $\log g(Y_n|\tilde{\theta}^{(i)}) \equiv \log g(y^*|\tilde{\theta}^{(i)}) = \sum_{t=1}^{T} \log g(y_t^*|\tilde{\theta}_t{}^{(i)})$.

Because $p^{(i)} = \exp\left(\sum_{t=1}^{T} \log p\left(y_t|\tilde{\theta}_t{}^{(i)}\right)\right)$ and $g^{(i)} = \exp\left(\sum_{t=1}^{T} \log g\left(y_t^*|\tilde{\theta}_t{}^{(i)}\right)\right)$, we can compute the estimated mean as in (5). The implementation of the mentioned algorithm is available in code *llik_fun_IS.m* and *WK5_Assignment_5_1*. The figure 1 displays resulting smoothed mode compared with the estimates from previous weeks.

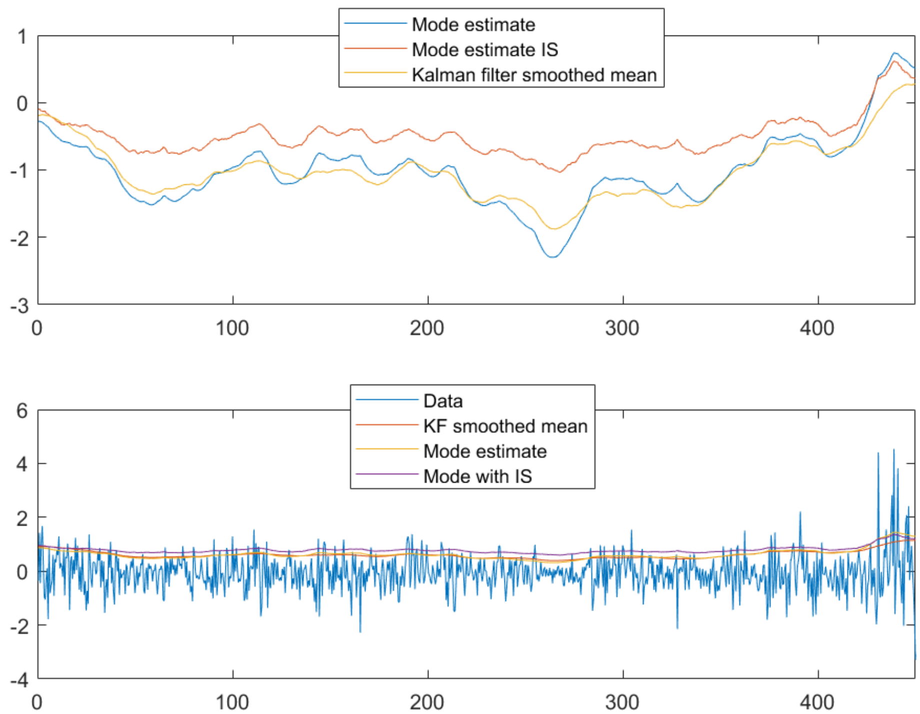

Figure 1: Plot of the transformed returns with filtered mean, mode and mode estimated using importance sampling.

Next consider the SV model with Student's t disturbances with $\epsilon_t \sim t(v)$, where $t(v)$ refers to the standardized Student's t distribution with v degrees of freedom.

(g) Estimate the parameters by maximizing the simulated likelihood function for the Student's t SV model using the importance sampling method. Note: use the random numbers for each likelihood evaluation when maximizing the likelihood function.

In this section we followed similar steps as in the previous section. The first step is to compute smoothed mode via KFS routine until the convergence criterion is reached. We used the same criterion as in the case of normally distributed disturbances. We used formulas in section **10.6.5** of **DK** book as

$$A_t = 2(\nu+1)^{-1}(\tilde{q}_t - 1)^{-1}\tilde{q}_t^2, \qquad x_t = \tilde{\theta}_t + \tilde{q}_t - \frac{1}{2}A_t, \tag{8}$$

where $\tilde{q}_t$ is defined as

$$\tilde{q}_t = 1 + \exp(-\tilde{\theta}_t)\frac{z_t^2}{\nu - 2}, \tag{9}$$

and z_t^2 are standardized observations computed as $z_t = (y_t - c)/\sigma$. Because $q_t > 1$, all A_t's are positive with probability one and we can proceed with the estimation of the mode. Resulting A_t

and mode were used to evaluate $\log g(Y_n)$ of the approximating model

$$y_t^* = \theta_t + u_t, \qquad u_t \sim \mathcal{N}(0, A_t) \tag{10}$$

$$\theta_t = Z_t \alpha_t \tag{11}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \qquad \eta_t \sim \mathcal{N}(0, Q_t). \tag{12}$$

Because the model is linear Gaussian with fixed A_t, the likelihood $\log g(Y_n) \equiv \log g(y^*)$ we obtained with standard KF methods that can be found in **7.2.1** of **DK** book.

After that importance sampling of $\theta_t^{(i)}$ was performed, exactly as in the previous section, where smoothed mode $\tilde{\theta}$ and variance A_t were used as inputs. Only the logarithm of conditional density $p(y_t|\theta_t)$ was computed differently, more specifically as

$$\log p(y_t|\tilde{\theta}_t^{(i)}) = \log \left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma(\frac{\nu}{2})} \right) - \frac{1}{2} \log((\nu-2)\sigma^2) - \frac{1}{2} \left(\tilde{\theta}_t^{(i)} + (\nu+1) \log \tilde{q}_t^{(i)} \right), \tag{13}$$

where $\tilde{q}_t^{(i)}$ is computed as in (9).

The resulting likelihood is computed as

$$\hat{\mathcal{L}}(\psi) = \log g(Y_n) + \bar{m} - \log N + \log \left(\sum_{i=1}^{N} \exp(m_i - \bar{m}) \right), \tag{14}$$

where m is defined as in f) and N is the number of importance sampling of $\tilde{\theta}$, in our case $N = 100$. We maximized average log-likelihood function $\hat{\mathcal{L}}(\psi)$using the vector of parameters $\hat{\psi} = (0.007, 0.9, 0.95, 5)$ as the initialization vector, with use of same set of random numbers for each iteration $i = 1, \ldots, N$ to avoid any errors that might steam from their randomness. Resulting parameters are displayed in the table 1.

$\hat{\sigma_\eta}^2$	$\hat{\sigma^2}$	$\hat{\phi}$	$\hat{\nu}$
0.0189	0.6480	0.9808	18.7109

Table 1: Estimates of vector *psi* for model with $\epsilon_t \sim \mathrm{t}(\nu)$.

References

[1] Durbin, J., Koopman, S.J. (2012). *Time Series Analysis by State Space Methods, Second Edition.* Oxford University Press