Inference Methods for Stochastic Volatility Models

Maddalena Cavicchioli

Ca’ Foscari University of Venice
Advanced School of Economics
Cannaregio 231, Venice, Italy
maddalena.cavicchioli@unive.it

Abstract

In the present paper we consider estimation procedures for stationary Stochastic Volatility models, making inferences about the latent volatility of the process. We show that a sequence of generalized least squares regressions enables us to determine the estimates. Finally, we make inferences iteratively by using the Kalman Filter algorithm.

Mathematics Subject Classification: 62M10, 62M20, 91B84, 93E03

Keywords: Stochastic Volatility, Generalized Least Squares, Kalman Filter

1 Introduction

Stochastic Volatility (SV) models have been received a growing interest in time series analysis since they find many financial applications as, for example, option pricing, asset allocation and risk management. For a comprehensive discussion on SV models see, for example, Taylor [3] and Tsyplakov [4]. Let us consider the basic stochastic volatility model given by

\begin{align}
y_t &= \exp \left\{ \frac{1}{2} h_t(\theta) \right\} u_t \\
h_t(\theta) &= \mu + \rho h_{t-1}(\theta) + v_t
\end{align}

where the error terms \( u_t \sim \mathcal{N}(0,1) \) and \( v_t \sim \mathcal{N}(0, \sigma_v^2) \) are assumed to be independent of one other, and the parameter vector \( \theta = (\mu, \rho, \sigma_v^2) \) is in a compact parameter set \( \Theta \subset (0, +\infty)^3 \). To ensure stationarity, we always set \( |\rho| < 1 \). In this case we have

\begin{equation}
h_t(\theta) = \mu(1-\rho)^{-1} + \sum_{i=0}^{\infty} \rho^i v_{t-i}
\end{equation}
hence $E(h_t) = \mu(1 - \rho)^{-1}$ and $\text{Var}(h_t) = \sigma_v^2(1 - \rho^2)^{-1}$. Squaring (1.1) and taking logs, we get the state-space representation

$$
(1.3) \quad x_t = \alpha + h_t(\theta) + \epsilon_t \\
\quad h_t(\theta) = \mu + \rho h_{t-1}(\theta) + \epsilon_t
$$

where $x_t = \log y_t^2$, $\alpha = E[\log u_t^2]$ is a real constant ($\simeq -1.27$) and $\epsilon_t = \log u_t^2 - \alpha$ is a martingale difference but not normal, i.e., $\epsilon_t \sim \text{IID}(0, \sigma^2_\epsilon)$, where $\sigma^2_\epsilon = \pi^2/2$ (here log denotes the natural logarithm). See Breidt and Carriquiry [1]. Following Kim-Nelson [2], Chp.3, we present two alternative ways of making inferences about $h_t$ (volatility) conditional on information available up to time $t$ (given $\theta$). In Section 2 we show that a sequence of generalized least squares (GLS) regressions enables us to determine $h_{t|t} = E(h_t|\Psi_t)$, where $\Psi_t$ denotes the information set up to time $t$. In Section 3 we make inferences about $h_t$ by employing the Kalman Filter algorithm.

## 2 Generalized Least Squares Estimation

As usual, we approximate the SV model in (1.3) by a Gaussian state-space model. From (1.3) we get

$$
(2.1) \quad h_t = \mu(1 - \rho)^{-1}(1 - \rho^{t-1}) + \rho^{t-1}h_1 + \sum_{i=0}^{t-2} \rho^i v_{t-i}
$$

for $t \geq 2$. Thus

$$
\begin{bmatrix}
    h_1 \\
    h_2 \\
    \vdots \\
    h_{t-1} \\
    h_t
\end{bmatrix}
= \begin{bmatrix}
    \rho^{-t+1} \\
    \rho^{-t+2} \\
    \vdots \\
    \rho^{-1} \\
    1
\end{bmatrix}
\begin{bmatrix}
    h_t \\
    \vdots \\
    \rho^{-1} \\
    1
\end{bmatrix}
- \mu \begin{bmatrix}
    \rho^{-t+1} \\
    \rho^{-t+2} \\
    \vdots \\
    \rho^{-1} \\
    0
\end{bmatrix}

\text{Define:}

$$
\mathbf{a}_t = (\rho^{-t+1} \quad \rho^{-t+2} \ldots \rho^{-1} \quad 1)'
$$

and

$$
\mathbf{B}_t = \begin{bmatrix}
    \rho^{-1} & \rho^{-2} & \ldots & \rho^{-t+2} & \rho^{-t+1} \\
    0 & \rho^{-1} & \ldots & \rho^{-t+3} & \rho^{-t+2} \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \ldots & 0 & \rho^{-1} \\
    0 & 0 & \ldots & 0 & 0
\end{bmatrix}
\quad \quad
\mathbf{C}_t = \begin{bmatrix}
    \alpha - \mu(1 - \rho)^{-1}(\rho^{-t+1} - 1) \\
    \alpha - \mu(1 - \rho)^{-1}(\rho^{-t+2} - 1) \\
    \vdots \\
    \alpha - \mu \rho^{-1} \\
    \alpha
\end{bmatrix}
$$
Inference methods for stochastic volatility models

Using the measurement equation in (1.3) and the above matrix relation, we have

\[(2.2) \quad x_t = C_t + a_t h_t + \epsilon_t\]

where \(x_t = (x_1 \ldots x_t)\), \(\epsilon_t = -B_t(v_2 \ldots v_t) + e_t\) and \(e_t = (e_1 e_2 \ldots e_t)'\). Then we have

\[(2.3) \quad E(\epsilon_t \epsilon_t') = \sigma_e^2 I_t + \sigma_v^2 B_t B_t' = \Omega_t.\]

One could apply GLS to model (2.2) for \(t = 2, \ldots, T\). Then we get

\[(2.4) \quad h_{t|t} = (a_t' \Omega_t^{-1} a_t)^{-1} a_t' \Omega_t^{-1} (x_t - C_t)\]

hence

\[(2.5) \quad h_{t|t} = h_t + (a_t' \Omega_t^{-1} a_t)^{-1} a_t' \Omega_t^{-1} \epsilon_t.\]

Then we have

\[E(h_{t|t}) = E(h_t) = \mu(1 - \rho)^{-1}\]

and

\[P_{t|t} = E[h_t - h_{t|t}]^2 = (a_t' \Omega_t^{-1} a_t)^{-1}.\]

Define

\[b_t = B_t' a_t ||a_t||^{-2} = \frac{1}{1 - \rho^{-2t}} \begin{bmatrix} \rho^{-t}(1 - \rho^{-2}) \\ \rho^{-t+1}(1 - \rho^{-4}) \\ \vdots \\ \rho^{-2t}(1 - \rho^{-2t+2}) \end{bmatrix}.\]

**Theorem 2.1** With the above notation, we have

\[P_{t|t} = \sigma_e^2 ||a_t||^{-2} + \sigma_v^2 ||b_t||^2\]

where

\[||a_t||^2 = \frac{1 - \rho^{-2t}}{1 - \rho^{-2}}, \quad ||b_t||^2 = \frac{\rho^{-4t+2} - 1}{\rho^2(1 - \rho^2)(1 - \rho^{-2t})^2} - \frac{2t - 1}{\rho^{2t+2}(1 - \rho^{-2t})^2}.\]

**Proof.** We apply the Sherman-Morrison-Woodbury (SMW) formula, i.e., if \(A\) and \(C\) are invertible matrices, then

\[(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}.\]

Setting \(A = \sigma_e^2 I_t, B = B_t, C = \sigma_v^2 I_{t-1}\) and \(D = B_t',\) from (2.3) we get

\[\Omega_t^{-1} = \sigma_e^{-2} I_t - \sigma_e^{-2} B_t (\sigma_v^{-2} I_{t-1} + \sigma_e^{-2} B_t B_t')^{-1} \sigma_e^{-2} B_t'.\]
hence
\[ a_i′Ω^{-1}a_i = \sigma^{-2}_e a_i′a_i - \sigma^{-2}_e a_i′B_t(\sigma^{-2}I_{t-1} + \sigma^{-2}B_t')^{-1}\sigma^{-2}B'_t a_i. \]

Apply again the SMW formula with \( A = \sigma^{-2}_e a_i′a_i, B = -\sigma^{-2}_e a_i′B_t, C = (\sigma^{-2}I_{t-1} + \sigma^{-2}B_t)′^{-1} \) and \( D = \sigma^{-2}B'_t a_i \). Then we have
\[ P_{t|t} = (a_i′Ω^{-1}a_i)^{-1} = \sigma^2_e (a_i′a_i)^{-1} - (a_i′a_i)^{-1}a_i′B_t \]
\[ = \sigma^2_e \|a_t\|^{-2} - (a_i′a_i)^{-1}a_i′B_t(\sigma^{-2}I_{t-1})^{-1}B′_t a_i(a_i′a_i)^{-1} \]
\[ = \sigma^2_e \|a_t\|^{-2} + \sigma^2_e \|b_t\|^2 \]

which gives the result of the statement.

**Theorem 2.2**
\[ P_\infty = \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} P_{t|t} = \sigma^2_e (1 - \rho^2)^{-1} = \text{var}(h_t). \]

**Proof.** To compute the partial sums of the series we took advantage of the software "Mathematica". We have (recall that \( |\rho| < 1 \))
\[ \sum_{t=1}^{T} \|a_t\|^{-2} = \sum_{t=1}^{T} \frac{1 - \rho^{-2}}{1 - \rho^{-2t}} = \frac{(\rho^2 - 1)\psi^{(0)}(T + 1) + (1 - \rho^2)\psi^{(0)}(1)}{\rho^2 \log(\rho^2)} \]
where \( \psi^{(0)}(z) = \psi(z) = \partial \log \Gamma_q(z)/\partial z \) denotes the \( q \)-digamma function. Since
\[ \lim_{T \to +\infty} \frac{1}{T} \psi^{(0)}(T + 1) = 0 \]
we get
\[ \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \|a_t\|^{-2} = 0. \]

Furthermore, we have
\[ \sum_{t=1}^{T} \frac{\rho^{-4t+2} - 1}{\rho^2(1 - \rho^2)(1 - \rho^{-2t})^2} = \frac{\psi^{(1)}(T + 1)}{\rho^2 \log^2(\rho^2)} + \frac{(\rho^2 + 1)\psi^{(0)}(T + 1)}{\rho^2(\rho^2 - 1) \log(\rho^2)} + \frac{T}{1 - \rho^2} + c \]
where \( c \) is the numerical constant
\[ c = -\frac{\psi^{(1)}(1)}{\rho^2 \log^2(\rho^2)} - \frac{(\rho^2 + 1)\psi^{(0)}(1)}{\rho^2(\rho^2 - 1) \log(\rho^2)} \]
and $\psi_q^{(1)}(z)$ denotes the first derivative of the $q$-digamma function. Now (2.6) and

\[
\lim_{T \to +\infty} \frac{1}{T} \psi_{\rho^2}^{(1)}(T + 1) = 0 \tag{2.8}
\]

imply

\[
\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{\rho^{-4t^2} - 1}{\rho^2(1 - \rho^2)(1 - \rho^{-2t})^2} = \frac{1}{1 - \rho^2}. \tag{2.9}
\]

It remains to consider the series

\[
\sum_{t=1}^{T} \frac{2t - 1}{\rho^2(1 - \rho^{-2t})^2} = \sum_{t=1}^{T} \frac{2t - 1}{\rho^2(\rho^t - \rho^{-t})^2}.
\]

We take the first Taylor expansion around $\rho_0 \in (0, 1)$ of the function $(\rho^t - \rho^{-t})^2$ and use the following sequence of inequalities

\[
0 < \frac{2t - 1}{(\rho^t - \rho^{-t})^2} \sim \frac{2t - 1}{(\rho_0 - \rho_0^{-t})^2 + 2t(\rho_0 - \rho_0^{-t})(\rho_0^{-t-1} + \rho_0^{-t-1})(\rho - \rho_0)} \leq \frac{1}{(\rho_0 - \rho_0^{-t})(\rho_0^{-t-1} + \rho_0^{-t-1})(\rho - \rho_0)}
\]

for $0 < \rho < \rho_0$. But we have

\[
\sum_{t=1}^{T} \frac{1}{(\rho^t - \rho^{-t})(\rho^{t-1} + \rho^{-t-1})} = \frac{\rho\psi_0^{(0)}(T + 1) - \rho\psi_0^{(0)}(T - \frac{i\pi}{2\log(\rho)} + 1)}{4\log(\rho)}
\]

\[
+ \frac{\rho\psi_0^{(0)}(T + \frac{i\pi}{2\log(\rho)} + 1) + \rho\psi_0^{(0)}(T - \frac{i\pi}{2\log(\rho)} + 1)}{4\log(\rho)} + d
\]

where $d$ is the numerical constant

\[
d = \frac{-\rho\psi_0^{(0)}(1) + \rho\psi_0^{(0)}(1 - \frac{i\pi}{2\log(\rho)}) + \rho\psi_0^{(0)}(1 + \frac{i\pi}{2\log(\rho)}) - \rho\psi_0^{(0)}(1 - \frac{i\pi}{2\log(\rho)})}{4\log(\rho)}.
\]

Then we have

\[
\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{1}{(\rho^t - \rho^{-t})(\rho^{t-1} + \rho^{-t-1})} = 0
\]

as

\[
\lim_{T \to +\infty} \frac{1}{T} \psi_0^{(0)}(T + 1) = \lim_{T \to +\infty} \frac{1}{T} \psi_0^{(0)}(T + \frac{i\pi}{2\log(\rho)} + 1)
\]

\[
= \lim_{T \to +\infty} \frac{1}{T} \psi_0^{(0)}(T - \frac{i\pi}{\log(\rho)} + 1) = 0.
\]
This implies
\[
\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \frac{2t - 1}{\rho^2 (\rho^t - \rho^{-t})^2} = 0
\]
hence
\[
(2.10) \quad \lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} ||b_t||^2 = \frac{1}{1 - \rho^2}.
\]
Finally, from (2.7) and (2.10), we get the result of the statement.

3 Estimation by the Kalman Filter

Since Model (1.3) is in linear state-space form, predicted filtered and smoothed values of \( h_t \) can be computed recursively via the Kalman Filter algorithm. See Kim-Nelson [2], Sec.3.1.2. Define \( h_{t|\tau} = E[h_t|\Psi_\tau] \), \( P_{t|\tau} = E[h_t - h_{t|\tau}]^2 \), \( x_{t|\tau} = E[x_t|\Psi_\tau] \), \( \eta_{t|\tau} = x_t - x_{t|\tau} \) and \( f_{t|\tau} = E[\eta_{t|\tau}]^2 \) for \( 1 \leq \tau \leq T \). For the one-step-ahead prediction, we have
\[
\begin{align*}
    h_{t|t-1} &= \mu + \rho h_{t-1|t-1} \quad P_{t|t-1} = \rho^2 P_{t-1|t-1} + \sigma_v^2 \\
    \eta_{t|t-1} &= x_t - x_{t|t-1} = x_t - \alpha - h_{t|t-1} = (h_t - h_{t|t-1}) + e_t \\
    f_{t|t-1} &= P_{t|t-1} + \sigma_e^2.
\end{align*}
\]
The initial states of the recursion are
\[
    h_{0|0} = \mu(1 - \rho)^{-1} \quad \text{and} \quad P_{0|0} = \sigma_v^2(1 - \rho^2)^{-1}.
\]
The updating is given by
\[
\begin{align*}
    h_{t|t} &= h_{t|t-1} + K_t \eta_{t|t-1} = h_{t|t-1} + K_t(x_t - \alpha - h_{t|t-1}) \\
    P_{t|t} &= P_{t|t-1} - K_t P_{t|t-1} = P_{t|t-1} - P_{t|t-1} f_{t|t-1}^{-1} f_{t|t-1}^{-1}
\end{align*}
\]
where
\[
    K_t = P_{t|t-1} f_{t|t-1}^{-1} = P_{t|t-1}(P_{t|t-1} + \sigma_e^2)^{-1}
\]
is the Kalman gain. From these recursions, one can construct the (quasi) Gaussian log-likelihood
\[
\ell(\theta|\Psi_T) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \log f_{t|t-1} - \frac{1}{2} \sum_{t=1}^{T} \eta_{t|t-1} f_{t|t-1}^{-1}.
\]
The smoothed estimates and their variances are given by
\[
\begin{align*}
    h_{t|T} &= h_{t|t} + P_{t|t} (h_{t+1|T} - \mu - \rho h_{t|t}) \\
    P_{t|T} &= P_{t|t} + [P_{t|t}]^2 (P_{t+1|T} - P_{t+1|t})
\end{align*}
\]
where

\[ P_t^* = \rho P_{t+1|t'} P_{t+1|t}. \]

Solving the difference equations in \( h_{t|t} \) and \( P_{t|t} \) gives the estimates. More precisely, we have

\[
P_{t|t} = \sigma_v^2 \rho^{2t} (1 - \rho^2)^{-1} \prod_{i=0}^{t-1} (1 - K_{t-i}) + \sigma_v^2 \sum_{k=0}^{t-1} \rho^{2k} \prod_{j=0}^{k} (1 - K_{t-j})
\]

\[
h_{t|t} = \mu (1 - \rho)^{-1} + \sum_{i=0}^{\infty} \rho^i K_{t-i} \eta_{t-i|t-i-1}
\]

\[
\eta_{t|t-1} = \alpha_t(L) v_t + \beta_t(L) e_t
\]

where \( \alpha_t(L) = 1 + \sum_{r=1}^{t} \rho^r \prod_{s=1}^{r} (1 - K_{t-s}) L^r \) and \( \beta_t(L) = \alpha_t(L)(1 - \rho L) \) (here \( L \) denotes the lag operator).

**References**


Received: November, 2012