The Journal of Open Source Software

stochvolTMB: An R-package for likelihood estimation of stochastic volatility models

Norwegian Computing Center

DOI: XX.XXXXX/joss.XXXXX

Software

- Review
- Repository
- Archive

Submitted:
Published:

Summary

Stochastic volatility (SV) models are often used to model financial returns that exhibit time-varying and autocorrelated variance. The first SV model was introduced by Taylor (1982) and models the logarithm of the variance as a latent autoregressive process of order one. Parameter estimation of stochastic volatility models can be challenging and a variety of methods have been proposed, such as simulated likelihood (Liesenfeld 2006), quasi-maximum likelihood (Harvey, Ruiz, and Shephard 1994) and Markov Chain Monte Carlo methods (MCMC) (Pitt and Shephard 1999; Kastner 2016). `stochvolTMB` takes a frequentist approach and estimates the parameters using maximum likelihood, similar to Skaug and Yu (2014). The latent variables are integrated out using the Laplace approximation. The models are implemented in `C++` using the `R`-package (R Core Team 2019) `TMB` (Kristensen et al. 2016) for fast and efficient estimation. `TMB` utilizes the `Eigen` (Guennebaud, Jacob, and others 2010) library for numerical linear algebra and `CppAD` (Bell 2005) for automatic differentiation of the negative log-likelihood. This can lead to substantial speed-up compared to MCMC methods.

Implementation

`stochvolTMB` implements stochastic volatility models of the form

$$
\begin{aligned}
y_t &= \sigma_y e^{h_t/2}\epsilon_t, \quad t = 1, \dots, T, \\
h_{t+1} &= \phi h_t + \sigma_h \eta_t, \quad t = 1, \dots, T-1, \\
\eta_t &\overset{\text{iid}}{\sim} \mathcal{N}(0,1), \\
\epsilon_t &\overset{\text{iid}}{\sim} F, \\
h_1 &\sim \mathcal{N}\left(0, \frac{\sigma_h}{\sqrt{(1-\phi^2)}}\right)
\end{aligned} \tag{1}
$$

where y_t is the observed log return for day t, h_t is the logarithm of the conditional variance of day t and $\boldsymbol{\theta} = (\phi, \sigma_y, \sigma_h)$ are the fixed parameters. Four distributions are implemented for ϵ_t: (1) The standard normal distribution; (2) The t-distribution with ν degrees of freedom; (3) The skew-normal distribution with skewness parameter α; and (4) The leverage model where (ϵ_t, η_t) are multivariate normal with zero mean and correlation coefficient ρ. The last three distributions add an additional fixed parameter to $\boldsymbol{\theta}$. `stochvolTMB` also supports generic functions such as `plot`, `summary`, `predict` and `AIC`. The plotting is implemented using `ggplot2` (Wickham (2016)) and data processing utilizes the `R`-package `data.table` (Dowle and Srinivasan 2019).

The parameter estimation is done in an iterative two-step procedure: (1) Optimize the joint negative log-likelihood with respect to the latent log-volatility $\boldsymbol{h} = (h_1, \dots, h_T)$ holding $\boldsymbol{\theta}$ fixed, and (2) Optimizing the Laplace approximation of the joint negative log-likelihood w.r.t $\boldsymbol{\theta}$. This procedure is iterated until convergence. Standard deviations for the log-volatility and the fixed parameters are obtained by the delta-method (Kristensen et al. 2016).

`stochvolTMB` different from `R`-package `stochvol` (Kastner 2016) as `stochvol` performs Bayesian inference using MCMC. By using optimization instead of simulation we are able to obtain a 5-10 times speed up, dependent on the data, model and number of observations.

References

Bell, Brad. 2005. "CppAD: A Package for C++ Algorithmic Differentiation." http://www.coin-or.org/CppAD.

Dowle, Matt, and Arun Srinivasan. 2019. *data.table: Extension of 'Data.frame'*. https://CRAN.R-project.org/package=data.table.

Guennebaud, Gaël, Benoît Jacob, and others. 2010. "Eigen V3." http://eigen.tuxfamily.org.

Harvey, Andrew, Esther Ruiz, and Neil Shephard. 1994. "Multivariate Stochastic Variance Models." *Review of Economic Studies* 61 (2): 247–64. http://EconPapers.repec.org/RePEc:oup:restud:v:61:y:1994:i:2:p:247-264.

Kastner, Gregor. 2016. "Dealing with Stochastic Volatility in Time Series Using the R Package stochvol." *Journal of Statistical Software* 69 (5): 1–30. https://doi.org/10.18637/jss.v069.i05.

Kristensen, Kasper, Anders Nielsen, Casper Berg, Hans Skaug, and Bradley Bell. 2016. "TMB: Automatic Differentiation and Laplace Approximation." *Journal of Statistical Software, Articles* 70 (5): 1–21. https://doi.org/10.18637/jss.v070.i05.

Liesenfeld, Richard. 2006. "Classical and Bayesian Analysis of Univariate and Multivariate Stochastic Volatility Models." http://www.tandfonline.com/doi/abs/10.1080/07474930600713424.

Pitt, Michael K., and Neil Shephard. 1999. "Time-Varying Covariances: A Factor Stochastic Volatility Approach." In *Bayesian Statistics, 6 (Alcoceber, 1998)*, 547–70. Oxford Univ. Press, New York.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R foundation for statistical computing. https://www.R-project.org/.

Skaug, Hans J., and Jun Yu. 2014. "A Flexible and Automated Likelihood Based Framework for Inference in Stochastic Volatility Models." *"Computational Statistics & Data Analysis"* "76": 642–54. https://doi.org/https://doi.org/10.1016/j.csda.2013.10.005.

Taylor, S. J. 1982. "Financial Returns Modelled by the Product of Two Stochastic Processes-a Study of the Daily Sugar Prices 1961-75." *Time Series Analysis : Theory and Practice* 1: 203–26. https://ci.nii.ac.jp/naid/10018822959/en/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.