bmgarch: An R-Package for Bayesian Multivariate GARCH models

Philippe Rast[*1] **and Stephen R. Martin**[2]

1 University of California, Davis **2** Comscore, Inc.

DOI: 10.21105/joss.03452

Software

- Review
- Repository
- Archive

Editor: Mehmet Hakan Satman

Reviewers:

- @ChristopherLucas
- @mwt

Submitted: 19 June 2021
Published: 08 August 2021

Introduction

Generalized autoregressive conditional heteroskedasticity models (GARCH) and their multivariate extension (MGARCH) are part of the economists' toolbox ever since their introduction in the early 1980s (Bollerslev, 1986; Engle, 1982; Tsay, 2013). Typically, the goal is to generate forecasts of volatility and covolatility for the next day or the near future in time series of assets or market indices. While GARCH models are primarily used in the econometric context, they can be used to capture and forecast heteroskedasticity in any time series. In fact, Rast et al. (2020) presented a parameterization for predicting and capturing within-person variability in human behavior in intensive longitudinal designs.

The main focus of MGARCH models is the estimation of the $p \times p$ conditional covariance matrix $\mathbf{H}_t$ that varies over $t = 1, ..., N$ time points and defines the (co-)volatility of p time series as $\boldsymbol{\epsilon}_t = \mathbf{H}_t^{1/2}\boldsymbol{\eta}_t$, where $\boldsymbol{\epsilon}_t$ is a $p \times 1$ vector of returns or residuals, assuming a process with a conditional mean of zero. $\boldsymbol{\eta}_t$ is a serially independent multivariate white noise process with covariance matrix $\mathbf{I}_p$. While conceptually straightforward, the crux is to define a model that maintains stationarity of the variances and covariances among the time series for all time points as well as positive definiteness for all $\mathbf{H}_t$. As such, a substantial amount of research of the past decades revolved around parameterizations of the conditional covariance matrix that fulfill all those desiderata (for a comparison of some of the most common parameterizations see Almeida et al., 2018).

Statement of need

While there are a number of readily available packages for univariate GARCH models in R, multivariate implementations are scarcely available. Currently, we are aware of only two packages in the Comprehensive R Archive Network (CRAN). One is `mgarchBEKK` (Schmidbauer et al., 2016) which implements BEKK as well as a bivariate asymmetric GARCH model. The other is `rmgarch` (Ghalanos, 2019), which includes DCC, GO-GARCH and Copula-GARCH models. Both packages are based on maximum likelihood methods. Moreover, some MGARCH models are implemented in proprietary software (such as Stata), but their focus is mostly limited to CCC, DCC, and VCC models without the option of estimating the location (e.g., time-varying mean returns) and scale (e.g., time-varying volatility) at the same time (Carnero & Eratalay, 2014). Again, these propietary solutions are based on maximum likelihood methods, while in contrast, `bmgarch` implements Bayesian methods for parameter estimation and inference.

At this current time, `bmgarch` implements a CCC (Bollerslev, 1990), DCC (Engle, 2002), pdBEKK (Rast et al., 2020), and a BEKK (Engle & Kroner, 1995) model based on either Gaussian or Student-t distributed innovations. All parameterizations allow arbitrary ARCH and GARCH orders. The Bayesian framework is ideally suited to handle the necessary constraints imposed on the MGARCH specifications to both keep the solution stationary and all

*corresponding author

The Journal of Open Source Software

t conditional covariance matrices $\mathbf{H}_t$ positive definite. Moreover, the model allows one to examine the posterior distribution in order to obtain the full range of Bayesian inference.

The model parameters are estimated using the probabilistic programming language Stan (Stan Development Team, 2016), which also allows one to use stan related diagnostics implemented in `rstan` (Stan Development Team, 2020).

bmgarch

The package is designed to take multivariate time series only. Hence, in terms of data, the minimum requirement is a data frame or matrix with at least two columns, representing the time series. Note that `bmgarch` currently does not support missing values in time series, nor in the predictor variables.

The default behavior in `bmgarch` is to estimate a CCC(1,1) parameterized model of order 1 for the ARCH and GARCH components assuming Student-t distributed innovations and a constant means (zero, typically) structure. The model is fit using 4 chains with 2000 iterations each (half of those are warm-up and the other half are sampling). As such, the minimum call will be `bmgarch( data = X )` with `X` being a matrix or data frame with at least 2 columns.

In order to include the effect of an external predictor variable on the unconditional and constant covariance term C, one can include a time-varying predictor `Y` of the same dimension as `X` by including the preditor with `xC = Y`, resulting in`bmgarch( data = X, xC = Y )`. Further, `bmgarch` takes a number of model related arguments governing the type of parameterization (`parameterization = {"CCC", "DCC", "BEKK", "pdBEKK"}`), the order of the GARCH (`P = {1, ..., t-1}`) and ARCH (`Q = {1, ..., t-1}`) processes, the mean structure (`meanstructure = {"constant", "arma"}`) and the type of distribution (`distribution = {"Student_t", "Gaussian"}`). The function also takes arguments with respect to sampling (number of `iterations` and number of `chains`) as well as whether data should be z-standardized before sampling (`standardize_data = {FALSE, TRUE}`) – this last argument can facilitate the estimation process.

Objects of the `bmgarch` family can be passed to the `plot()` function, `print()`, and `summary()` for a summary table.

Moreover, each estimated model can be passed to the `forecasting()` function to generate m-ahead forecasts of variances, correlations, and, if the meanstructure is included (such as with ARMA(1,1)), forecasts of means. Moreover, `bmgarch` integrates model ensemble techniques to generate model weighted forecasts (and estimates) based on Bayesian model averaging and stacking techniques described by Yao et al. (2018) and Bürkner et al. (2020). The forecasted objects can again be plotted and summarized via the corresponding generic functions.

Summary

The `bmgarch` package implements four MGARCH parameterizations and estimates all parameters via a Bayesian framework. Sampling from the posterior distribution is based on HMC and is implemented via stan. The `bmgarch` objects can be printed and plotted as well as passed on to a forecasting function that runs either simple or ensemble forecasts.

Acknowledgements

The work reported in this publication was supported by the National Institute On Aging of the National Institutes of Health under Award Number R01AG050720 to PR. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Almeida, D. de, Hotta, L. K., & Ruiz, E. (2018). MGARCH models: Trade-off between feasibility and flexibility. *International Journal of Forecasting*, *34*(1), 45–63. https://doi.org/10.1016/J.IJFORECAST.2017.08.003

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model. *The Review of Economics and Statistics*, *72*(3), 498——505. https://doi.org/10.2307/2109358

Bürkner, P.-C., Gabry, J., & Vehtari, A. (2020). Approximate leave-future-out cross-validation for Bayesian time series models. *Journal of Statistical Computation and Simulation*, 1–25. https://doi.org/10.1080/00949655.2020.1783262

Carnero, M. A., & Eratalay, M. H. (2014). Estimating VAR-MGARCH models in multiple steps. *Studies in Nonlinear Dynamics & Econometrics*, *18*(3), 339–365. https://doi.org/10.1515/snde-2012-0065

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

Engle, R. F. (2002). Dynamic Conditional Correlation. *Journal of Business & Economic Statistics*, *20*(3), 339–350. https://doi.org/10.1198/073500102288618487

Engle, R. F., & Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric Theory*, *11*(1), 122–150. https://doi.org/10.1017/S0266466600009063

Ghalanos, A. (2019). *Rmgarch: Multivariate GARCH models*.

Rast, P., Martin, S. R., Liu, S., & Williams, D. R. (2020). A New Frontier for Studying Within-Person Variability: Bayesian Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Psychological Methods*, 1939–1463. https://doi.org/10.1037/met0000357

Schmidbauer, H., Roesch, A., & Tunalioglu, V. S. (2016). *mgarchBEKK: Simulating, estimating and diagnosing MGARCH (BEKK and mGJR) processes*.

Stan Development Team. (2016). *The Stan C++ Library* (2.14.0 ed.). http://mc-stan.org

Stan Development Team. (2020). *RStan: The R interface to Stan*. http://mc-stan.org/

Tsay, R. S. (2013). *An Introduction to Analysis of Financial Data with R* (p. 416). John Wiley & Sons. ISBN: 978-0-470-89081-3

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using Stacking to Average Bayesian Predictive Distributions. *Bayesian Analysis*, *13*(3), 917–1007. https://doi.org/10.1214/17-BA1091