Journal of Statistical Software

November 2024, Volume 111, Issue 4. doi: 10.18637/jss.v111.i04

BEKKs: An R Package for Estimation of Conditional Volatility of Multivariate Time Series

Markus J. Fülle
University of Göttingen

Alexander Lange
University of Göttingen

Christian M. Hafner
UCLouvain

Helmut Herwartz
University of Göttingen

Abstract

We describe the R package **BEKKs**, which implements the estimation and diagnostic analysis of a prominent family of multivariate generalized autoregressive conditionally heteroskedastic (MGARCH) processes, the so-called BEKK models. Unlike existing software packages, we make use of analytical derivatives implemented in efficient C++ code for nonlinear log-likelihood optimization. This allows fast parameter estimation even in higher model dimensions $N > 3$. The baseline BEKK model is complemented with an asymmetric parameterization that allows for a flexible modeling of conditional (co)variances. Furthermore, we provide the user with the simplified scalar and diagonal BEKK models to deal with high dimensionality of heteroskedastic time series. The package is designed in an object-oriented way featuring a comprehensive toolbox of methods to investigate and interpret, for instance, volatility impulse response functions, risk estimation and forecasting (VaR) and a backtesting algorithm to compare the forecasting performance of alternative BEKK models. For illustrative purposes, we analyze a bivariate ETF return series (S&P, US treasury bonds) and a four-dimensional system comprising, in addition, a gold ETF and changes of a log oil price by means of the suggested package. We find that the **BEKKs** package is more than 100 times faster for time series systems of dimension $N > 3$ than other existing packages.

Keywords: BEKK model, multivariate GARCH, leverage effect, value-at-risk, impulse response functions, R.

1. Introduction

Models for the volatility of time series have played an enormous role in empirical work, primarily in economics and finance. The association of volatility with the notion of risk often shifts it to the focus of attention in many financial applications. Univariate volatility models in discrete time have been developed at the end of the last century, the most prominent being the stochastic volatility (SV) model of Taylor (1982), and the generalized autoregressive conditional heteroskedasticity (GARCH) model of Engle (1982) and Bollerslev (1986). SV models treat volatility as a latent factor, whereas in GARCH models it is observable when conditioning on past information. While SV models offer a high degree of flexibility, estimation and inference are challenging, see e.g., Shephard (2004) for selected readings on SV models.

GARCH models, at least in the univariate case, are by now thoroughly investigated and understood, both from a theoretical and empirical point of view, see e.g., Hafner (2008) for a review. There is a consensus that maximum likelihood provides a straightforward and efficient method for estimation and inference. Many variants of the classical GARCH model have been proposed to take into account empirical phenomena such as the prominent leverage effect of stock return volatility. On the software side, there are many packages, and even with a number of thousands of observations, estimation and inference are usually only a matter of fractions of a second.

In the multivariate case, modelling volatility has been challenging from the beginning, due to the fact that for N series, the number of volatilities and co-volatilities (i.e., covariances or correlations) increases as $\mathcal{O}(N^2)$, and the number of parameters often increases at a similar rate. Balancing flexibility with feasibility has been the main challenge for modeling and estimation. Multivariate GARCH (MGARCH) models have been introduced soon after the univariate version, and the first MGARCH model is the so-called vec-model of Bollerslev, Engle, and Wooldridge (1988), which allows a flexible modeling of conditional volatilities and covariances, but does not guarantee a positive definite covariance matrix.

A restricted version of the vec-model is the so-called BEKK model of Engle and Kroner (1995), based on an early working paper by Baba, Engle, Kraft, and Kroner (1983). It is slightly less general than the vec-model but has the important advantage of ensuring positivity of the conditional covariance matrix. Extensions of the classical BEKK model have been proposed, for example to take into account the leverage effect as in Kroner and Ng (1998). Estimation theory for the BEKK model has been developed over recent years and can be considered established by now, with important references being Comte and Lieberman (2003) and Hafner and Preminger (2009).

Another type of MGARCH models is the dynamic conditional correlation (DCC) model of Engle (2002) generalizing the constant conditional correlation (CCC) model of Bollerslev (1990), (see also Tse and Tsui 2002 for a similar specification). These models, and their extensions, separate the problem of modeling volatilities and correlations into two steps: a first, with estimation of the volatilities, and a second for the correlations, based on standardized observations using the volatilities of the first step. This has practical and computational advantages, although some inconsistency arguments have been raised by Aielli (2013), who suggested an extension of the model to circumvent this problem. A major drawback, to date, of the DCC class of models is that the estimation theory is very complicated, and to best of our knowledge has not been established as of today.

Because of the theoretical drawbacks of the DCC class of models, we concentrate in this paper on the BEKK models, for which a sound and complete estimation theory is available. Caporin and McAleer (2012) provide a comparison of the theoretical properties of the BEKK and DCC model classes, while their empirical properties are compared in Caporin and McAleer (2008). The main conclusion is that BEKK can consistently estimate conditional correlations under testable assumptions, which to date has not been shown for the DCC model, while empirically the BEKK model has not been significantly outperformed by the DCC model. We therefore refrain from covering both models here and focus on the BEKK model and its different versions.

In the range of R packages, the widely-applied **rugarch** package of Ghalanos (2024) includes the evaluation of univariate GARCH models and a broad toolbox to predict volatilities and risks by means of estimated models. Moreover, the **rmgarch** package (Ghalanos 2022) provides a similar functionality for multivariate time series using the CCC and DCC MGARCH models but does not cover the estimation of flexible MGARCH models in BEKK form. The **RATS** software (Estima 2024) comprises estimation tools for MGARCH models (including BEKK) but is not provided as an open source tool for a broad user-base. Currently, there are three R packages offering BEKK estimation. However, they have some limitations which we may be addressed by the newly designed package. The **MTS** package (Tsay and Wood 2022) allows only estimation up to three-dimensional processes, and relies on purely numerical optimization. While the second competing package **mgarchBEKK** (Schmidbauer, Roesch, and Tunalioglu 2022) can be applied to higher dimensional time series, it is also computationally less efficient than **BEKKs**, since it relies on numerical optimization techniques. The third competing package **bmgarch** (Rast and Martin 2023) utilizes Bayesian estimation algorithms making it substantially slower and computationally costly, in particular, if the interest turns towards systems that imply rich cross equation dependencies of conditional (co)variances. Moreover, the competing packages do not offer risk evaluation by a backtesting procedure such as our new **BEKKs** (Fülle, Lange, Hafner, and Herwartz 2024) package. Stepping outside the R framework, full BEKK model implementations are missing for EViews (IHS Markit 2021) (allows only diagonal BEKK), Stata (StataCorp 2021) and Python (Van Rossum *et al.* 2011). Stata and EViews, however, provide the usership with DCC and CCC MGARCH models. A DCC implementation is available in Python via the module **mgarch** (Srivastavas 2022).

Our objective in this paper is to first recall some theoretical properties of the BEKK model, including estimation theory and analytical derivatives needed for iterative model evaluation and the assessment of the asymptotic covariance matrix. Hafner and Herwartz (2008) emphasize as a core conclusion from their simulation studies that utilizing analytical derivatives is extremely beneficial for both the estimation of MGARCH models and the provision of reliable inferential outcomes as, for instance, t ratios. Using instead numerical derivatives tends to become unreliable in higher dimensions, especially for the evaluation of the Hessian matrix. Accordingly, model estimation and inferential analysis in the **BEKKs** package builds upon the analytical first and second order derivatives as provided in Hafner and Herwartz (2008). While **BEKKs** provides analytical derivatives in R, core optimization procedures have been implemented in C++ to enable marked performance improvements in terms of computation time and/or speed. To enhance the functionalities of **BEKKs** and going beyond alternative packages (e.g., **bmgarch**, **mgarchBEKK** or **MTS**), we furthermore provide various diagnostic checks for the estimated model such as portmanteau statistics of the squares and cross-products of standardized residuals. Most notably, we provide a unique, flexible asymmetric parametrization of volatility (spillovers) in the vein of Grier, Henry, Olekalns, and Shields (2004). In empirical applications, volatility impulse response functions (VIRFs), as defined by Hafner and Herwartz (2006), have become increasingly popular, so that we implement functions to directly obtain VIRFs based on an estimated BEKK model.

According to Engle (2001), MGARCH models are of particular interest when it comes to estimate volatilities and risks of financial assets and portfolios. To this end, we provide the usership with a forecasting tool for volatilities and risks assuming different distributions (for instance, Gaussian, standardized Student's t distribution and the empirical distribution of model-implied residuals). In order to compare the risk forecasting by means of the value-at-risk (VaR) (Jorion 1996, 2007), we provide a backtesting procedure and coverage tests of Kupiec (1995) and Christoffersen (1998). The test statistics are imported from the **GAS** package (Ardia, Boudt, and Catania 2019). We finish the paper with two examples of empirical applications to demonstrate the functionality of the package and computation time in comparison with competing packages.

2. MGARCH

Let r_t define an N-dimensional vector of returns of speculative assets with $t \in \{1, \ldots, T\}$. The information available at time t is $\Omega_t = \{r_1, \ldots, r_t\}$. To describe conditional first and second order properties of r_t, MGARCH models align with the following representation:

$$\begin{aligned} r_t &= \mu_t + e_t, \\ e_t &= H_t(\theta)^{1/2}\xi_t, \end{aligned} \tag{1}$$

where $\mu_t = \mathsf{E}[r_t|\Omega_{t-1}]$ and $H_t(\theta) = \mathsf{COV}[r_t|\Omega_{t-1}]$. Model parameters that formalize the relation between H_t and the history of the process are collected in the parameter vector $\theta \in \Theta$, with Θ indicating the respective parameter space. For notational convenience we henceforth set $H_t = H_t(\theta)$. Moreover, $H_t^{1/2}$ in Equation 1 denotes a decomposition matrix of H_t such that $H_t = H_t^{1/2}\left(H_t^{1/2}\right)^\top$.[1] The shocks ξ_t are assumed i.i.d. and $\xi_t \sim (0, I_N)$, where I_N is the N-dimensional identity matrix. While empirical return processes typically show marked time variation in second order moments, the mean of the elements in r_t is often a-priori assumed to be time invariant such that centered vector residuals are subjected to model estimation, i.e., $e_t = r_t - \bar{r}$, where $\bar{r}$ denotes the unconditional mean of r_t. Throughout, we consider the process e_t to be weakly (or covariance) stationary.

Maximum likelihood (ML) estimation of MGARCH models requires a distributional assumption. As the true distribution of model innovations ξ_t is unknown, it has become a widespread approach to perform quasi ML (QML) estimation to maximize the Gaussian log-likelihood for three reasons: First, the ad-hoc assumption of Gaussian model innovations deliberates the analyst from the need to elicit a particular alternative and potentially false distribution. Second, as we argue below under the notion of QML-estimation, the consequences of log-likelihood misspecification are yet well understood. Third, the maximization of the Gaussian log-likelihood is numerically convenient in the present case as it only depends on the MGARCH implied (co)variance profile and does not involve the estimation of further

[1] Throughout this work we use the spectral decomposition of H_t.

distributional parameters. The Gaussian log-likelihood reads as

$$\begin{aligned}\mathcal{L}(\theta) &= \sum_{t=1}^{T} l_t(\theta) \\ &= \sum_{t=1}^{T} \left(-\frac{N}{2}\log(2\pi) - \frac{1}{2}\log(\det H_t) - \frac{1}{2} e_t^\top H_t^{-1} e_t \right),\end{aligned}$$

conditional on an available sample $(e_t)_{t=1,2,...T}$.

The score function and the Hessian have, respectively, the following representation

$$\frac{\partial l(\theta)}{\partial \theta_i} = -0.5 \sum_{t=1}^{T} \text{trace}\left(\frac{\partial H_t}{\partial \theta_i} H_t^{-1} - e_t e_t^\top H_t^{-1} \frac{\partial H_t}{\partial \theta_i} H_t^{-1} \right)$$

and

$$\begin{aligned}\frac{\partial^2 l(\theta)}{\partial \theta_i \partial \theta_j} = 0.5 \sum_{t=1}^{T} \text{trace}\Bigg(& \frac{\partial^2 H_t}{\partial \theta_i \partial \theta_j} H_t^{-1} - e_t e_t^\top H_t^{-1} \frac{\partial^2 H_t}{\partial \theta_i \partial \theta_j} H_t^{-1} - \frac{\partial H_t}{\partial \theta_i} H_t^{-1} \frac{\partial H_t}{\partial \theta_j} H_t^{-1} \\ & + e_t e_t^\top H_t^{-1} \frac{\partial H_t}{\partial \theta_j} H_t^{-1} \frac{\partial H_t}{\partial \theta_i} H_t^{-1} + e_t e_t^\top H_t^{-1} \frac{\partial H_t}{\partial \theta_i} H_t^{-1} \frac{\partial H_t}{\partial \theta_j} H_t^{-1} \Bigg).\end{aligned}$$

Under conditions listed in Comte and Lieberman (2003) and Hafner and Preminger (2009), the QML estimator $\widehat{\theta} = \arg\max_{\Theta} \mathcal{L}(\theta)$ is consistent and asymptotically normally distributed, even in the case of non-normally distributed innovations, i.e., under misspecification of the Gaussian likelihood. The asymptotic distribution of the QML estimator is given by

$$\sqrt{T}(\widehat{\theta} - \theta_0) \xrightarrow{\mathcal{D}} N\left(0, \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}\right), \tag{2}$$

with

$$\mathcal{I} = \mathsf{E}\left[\left.\frac{\partial l_t(\theta)}{\partial \theta}\right|_{\theta_0} \left.\frac{\partial l_t(\theta)}{\partial \theta^\top}\right|_{\theta_0} \right], \quad \mathcal{J} = -\mathsf{E}\left[\left.\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta^\top}\right|_{\theta_0} \right]$$

and where the expectation is taken with respect to the true model parameters. The matrix $\mathcal{I}$ is the expectation of the outer product of the score vector evaluated at the true parameter vector θ_0, whereas $\mathcal{J}$ is the negative expectation of the Hessian evaluated at θ_0. If the residual process ξ_t is Gaussian, $\mathcal{I} = \mathcal{J}$ and the asymptotic covariance matrix reduces to $\mathcal{I}^{-1}$ (Bollerslev and Wooldridge 1992).

With respect to the functional form relating the dynamic covariance to the history Ω_{t-1} the so-called vec specification of MGARCH models encompasses all linear specifications. Due to its generality the vec specification provides an encompassing tool to compare dynamic features, as e.g., impulse response functions, implied by alternatively restricted models (Hafner and Herwartz 2006). It has practical drawbacks, however, due to the large number of parameters and the difficulties of keeping H_t positive definite. For the sake of notational simplicity we consider the vec model of order $(1,1)$ as given by

$$\text{vech}(H_t) = \omega + \mathcal{A}\text{vech}(e_{t-1}e_{t-1}^\top) + \mathcal{G}\text{vech}(H_{t-1}), \tag{3}$$

where ω is a $N(N+1)/2 \times 1$ parameter vector and $\mathcal{A}$ and $\mathcal{G}$ are square parameter matrices of dimension $N(N+1)/2$.

Alternative MGARCH models now differ in the functional form relating the elements of H_t to the information set Ω_{t-1}. In this work we focus on model representations that fit into the class of so-called BEKK models. Henceforth, let K_{NN} be the commutation matrix, L_N the elimination matrix, D_N the duplication matrix, and its generalized inverse D_N^+.

2.1. BEKK

In its most general form the so-called BEKK model (Engle and Kroner 1995) of order p, q, K (BEKK(p, q, K)) reads as[2]

$$H_t = CC^\top + \sum_{k=1}^{K}\sum_{i=1}^{q} A_{ik}^\top e_{t-i}e_{t-i}^\top A_{ik} + \sum_{k=1}^{K}\sum_{i=1}^{p} G_{ik}^\top H_{t-i} G_{ik},$$

with C, A_{ik}, and G_{ik} being $N \times N$ parameter matrices of which C is lower triangular. The summation limit K determines the generality of the BEKK specification within the class of positive definite covariance processes. A necessary and sufficient condition for covariance stationarity is given by

$$\rho\left(\sum_{i=1}^{q}\sum_{k=1}^{K} (A_{ik}\otimes A_{ik})^\top + \sum_{i=1}^{p}\sum_{k=1}^{K} (G_{ik}\otimes G_{ik})^\top\right) < 1,$$

with $\rho(A)$ being the spectral radius of a matrix A.[3] Model implied covariance paths are positive definite under suitable initialization and the additional condition that

$$\operatorname{diag}(C) \geq 0 \tag{4}$$

must hold. The most popular is the BEKK$(1, 1, 1)$ representation. It is formulated using one-period lags only and reads as

$$H_t = CC^\top + A^\top e_{t-1}e_{t-1}^\top A + G^\top H_{t-1} G,$$

where $\theta = (\operatorname{vech}(C)^\top, \operatorname{vec}(A)^\top, \operatorname{vec}(G)^\top)^\top$.

Partial derivatives of the log-likelihood function obtain as

$$\begin{aligned}
\frac{\partial \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top} &= 2D_N D_N^+(C\otimes I_N)L_N^\top + (G\otimes G)^\top \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vech}(C)^\top},\\
\frac{\partial \operatorname{vec}(H_t)}{\partial \operatorname{vec}(A)^\top} &= 2D_N D_N^+(I_N\otimes A^\top e_{t-1}e_{t-1}^\top) + (G\otimes G)^\top \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(A)^\top},\\
\frac{\partial \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top} &= 2D_N D_N^+(I_N\otimes G^\top \operatorname{vec}(H_{t-1})) + (G\otimes G)^\top \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(G)^\top}.
\end{aligned}$$

Explicit representations of second order derivatives have been provided in Hafner and Herwartz (2008). For convenience these derivatives are encountered in Appendix A.

[2]The acronym BEKK is named after an initial working paper version (Baba *et al.* 1983) of Engle and Kroner (1995).

[3]Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a (complex) matrix $A \in \mathbb{C}^{n\times n}$. Then, the spectral radius of A is defined as $\rho(A) = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$.

2.2. Asymmetric BEKK

Similar to the asymmetric DCC model of Cappiello, Engle, and Sheppard (2006), the asymmetric BEKK(1, 1, 1) model proposed by Kroner and Ng (1998) is given by

$$H_t = CC^\top + A^\top e_{t-1}e_{t-1}^\top A + B^\top \eta_{t-1}\eta_{t-1}B + G^\top H_{t-1}G, \tag{5}$$

with $\eta_t = \mathbb{I}(e_t < 0) \odot e_t$, where $\mathbb{I}(\cdot)$ is the elementwise indicator function. Alternative specifications have been proposed by Grier *et al.* (2004), where the indicator function operates jointly on all components of e_t. A general specification would be given by model in Equation 5 with $\eta_t = \left\{\prod_{j=1}^N \mathbb{I}(e_{t,j} \in I_j)\right\} e_t$, where $I_j \subset \mathbb{R}$, for example $I_j = (-\infty, 0]$. Here, the asymmetry term becomes active only if all N individual conditions hold.

We develop a condition for covariance stationarity for the asymmetric BEKK model that holds for both the Kroner and Ng (1998) and Grier *et al.* (2004) versions. It is convenient to observe that $\mathsf{E}[\eta_t\eta_t^\top] = \mathsf{E}[e_te_t^\top] \odot W = \Sigma \odot W$ for some symmetric positive definite matrix W, where Σ is the unconditional variance-covariance matrix of e_t. Then, note that

$$\mathsf{E}[\text{vec}(\eta_t\eta_t^\top)] = \text{diag}(\text{vec}(W))\text{vec}(\Sigma),$$

where diag(a) is a diagonal matrix with vector a on the diagonal. Therefore, the vectorized unconditional covariance of the asymmetric BEKK model is

$$\text{vec}(\Sigma) = \left\{I_{N^2} - (A \otimes A)^\top - (G \otimes G)^\top - (B \otimes B)^\top \text{diag}(\text{vec}(W))\right\}^{-1} \text{vec}(CC^\top),$$

provided that all eigenvalues of the matrix

$$(A \otimes A)^\top + (G \otimes G)^\top + (B \otimes B)^\top \text{diag}(\text{vec}(W))$$

are smaller than one in modulus. To estimate the matrix W, we replace expectations by sample averages,

$$S_1 := \frac{1}{T}\sum_{t=1}^T \eta_t\eta_t^\top = \frac{1}{T}\sum_{t=1}^T e_te_t^\top \odot \widehat{W} =: S_2 \odot \widehat{W}$$

and hence

$$\widehat{W} = S_1 \div S_2$$

where $\div$ is elementwise division.

The score function of the model comprises an extension of the score function of the BEKK(1, 1, 1) process, i.e,

$$\frac{\partial \, \text{vec}\,(H_t)}{\partial \, \text{vec}(B)^\top} = 2D_ND_N^+(I_N \otimes B^\top \eta_{t-1}\eta_{t-1}^\top) + (G \otimes G)^\top \frac{\partial H_{t-1}}{\partial \, \text{vec}(B)^\top}.$$

second order derivatives for the asymmetric BEKK are in Appendix A.

2.3. Restricted BEKK specifications

The estimation of MGARCH models relies on nonlinear optimization routines. Since the number of BEKK model parameters increases for medium and higher dimensional return vectors at rate $\mathcal{O}(N^2)$ the literature yet provides a set of model representations developed to respect principles of model parsimony and avoid a curse of dimensionality. BEKK variants that deserve attention in this regard are, for instance, the diagonal BEKK (dBEKK, Baba *et al.* 1983) and the scalar BEKK (sBEKK, Ding and Engle 2001). Whereas the (asymmetric) diagonal BEKK process consists of parameter matrices A, B and G restricted to diagonal matrices, the scalar BEKK is obtained by substituting scalar parameters for parameter matrices (i.e., introducing scalars a, b and g to replace A, B and G). The dBEKK and its first and second order derivatives are therefore similar to the respective equations for the BEKK model discussed above. However, the zeros obtained by setting the off-diagonal parameters to zero need to be sliced out. The asymmetric sBEKK reads as

$$H_t = CC^\top + a e_{t-1} e_{t-1}^\top + b \eta_{t-1} \eta_{t-1}^\top + g H_{t-1},$$

with $a, b, g \in \mathbb{R}$. To guarantee positive definiteness of the process, additionally to Equation 4, the parameters must be positive, i.e., $a, b, g \geq 0$. For covariance stationarity of the sBEKK, respectively, asymmetric sBEKK process, all eigenvalues of

$$a \cdot I_N + b \cdot W + g \cdot I_N$$

must be smaller than one in modulus.

The score functions read as

$$\begin{aligned}
\frac{\partial \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top} &= 2 \cdot D_N D_N^{+} (C \otimes I_N) L_N^\top + g \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vech}(C)^\top} \\
\frac{\partial \operatorname{vec}(H_t)}{\partial a} &= \operatorname{vec}(e_{t-1} e_{t-1}^\top) + g \frac{\partial \operatorname{vec}(H_{t-1})}{\partial a} \\
\frac{\partial \operatorname{vec}(H_t)}{\partial g} &= \operatorname{vec}(H_{t-1}) + g \frac{\partial \operatorname{vec}(H_{t-1})}{\partial g} \\
\frac{\partial \operatorname{vec}(H_t)}{\partial b} &= \operatorname{vec}(\eta_{t-1} \eta_{t-1}^\top) + g \frac{\partial \operatorname{vec}(H_{t-1})}{\partial b}.
\end{aligned}$$

The symmetric sBEKK obtains by setting $b = 0$. Second order derivatives for the scalar BEKK are given in Appendix A.

2.4. Estimation

With the analytical derivatives at hand, we are able to use an estimation algorithm which does not rely on numerical derivatives. For the package we use the so-called BHHH algorithm (Berndt, Hall, Hall, and Hausman 1974), which does not require to evaluate the Hessian matrix making the estimation process even faster and computationally less costly. The update of the parameter estimate at the i-th iteration, conditional on starting values θ_0, is given by

$$\theta_{i+1} = \theta_i - \lambda_i \left[\frac{\partial l(\theta_i)}{\partial \theta_i} \frac{\partial l(\theta_i)}{\partial \theta_i^\top} \right]^{-1} \frac{\partial l(\theta_i)}{\partial \theta_i},$$

where λ_i is a flexible step size used for iteration i.

3. Methods and tools

In order to further interpret and use the BEKK models, we provide the following methods within the package.

3.1. Portmanteau test for remaining second order correlation

Let $\zeta_t = \text{vech}(\xi_t \xi_t^\top)$ denote the $N(N+1)/2$ dimensional vector of cross products of MGARCH implied model residuals. In case of a well specified MGARCH model, the random variables in ζ_t should be free of serial correlation. We consider a Portmanteau test of the null hypothesis of no remaining correlation as a general model diagnostic to assess the convenience of a specific empirical model, i.e.,

$$H_0 : \mathsf{E}(\zeta_t \zeta_{t-i}^\top) = 0, \quad i = 1, \dots, h.$$

The respective Portmanteau statistic joint with its approximate asymptotic distribution is

$$Q_h = T \sum_{j=1}^{h} \text{trace}(\hat{C}_j^\top \hat{C}_0^{-1} \hat{C}_j^\top \hat{C}_0^{-1}) \approx \chi^2((h-2)N^2),$$

where $\hat{C}_i = T^{-1} \sum_{t=i+1}^{T} \hat{\zeta}_t \hat{\zeta}_{t-i}^\top$ see Lütkepohl (2007, p. 577) and Li and McLeod (1981) for VARMA models.

3.2. VIRFs

As it is typical in multivariate dynamic systems, the parametric model is too complex to directly assess the marginal effects of particular variables on the covariance path H_t. Impulse response functions illustrate the dynamic effects of isolated shocks occurring in one variable at a particular point in time on a system of variables (see e.g., Lütkepohl 2007, Chap. 9). Building on the concept of conditional moment profiles (Gallant, Rossi, and Tauchen 1993), Hafner and Herwartz (2006) have defined so-called variance impulse response functions to assess the marginal effects of shocks on the second order properties of observables e_t at horizon ν as follows:

$$\mathcal{V}_{t+\nu}(\xi^*, \Omega_{t-1}) = \mathsf{E}[\text{vech}(H_{t+\nu})|\Omega_{t-1}, \xi_t = \xi^*] - \mathsf{E}[\text{vech}(H_{t+\nu})|\Omega_{t-1}]. \tag{6}$$

In Equation 6, ξ^* is a shock vector of interest such that $\mathcal{V}_{t+\nu}(\xi^*, \Omega_{t-1})$ provides a direct comparison of two (future) covariance patterns in the spirit of conditional moment profiles of Gallant *et al.* (1993). One pattern is determined under the shock scenario and the other one describes a counterfactual "steady state". The shock of interest ξ^* might consist of an isolated unit impulse with other elements being zero or it could be obtained from an (estimated) empirical model $\xi^* = H_t^{-1/2} e_t$, where $H_t^{1/2}$ is a given covariance decomposition. It is convenient to derive VIRFs from the vec form, see Equation 3, of the BEKK model,

$$\text{vech}(H_t) = \mathcal{C} + \mathcal{A}\text{vech}(e_{t-1} e_{t-1}^\top) + \mathcal{G}\text{vech}(H_{t-1}), \tag{7}$$

where $\mathcal{C} = \text{vech}(CC^\top)$, $\mathcal{A} = D_N^+(A \otimes A)^\top D_N$, and $\mathcal{G} = D_N^+(G \otimes G)^\top D_N$. Referring to model parameters in Equation 7 and conditional covariance decompositions $H_t^{1/2}$, VIRFs are determined recursively as

$$\begin{aligned} \mathcal{V}_{t+1}(\xi^*, \Omega_{t-1}) &= \mathcal{A}\text{vech}(H_t^{1/2} \xi^* \xi^{*\top} H_t^{1/2} - H_t), \\ \mathcal{V}_{t+\nu}(\xi^*, \Omega_{t-1}) &= \{\mathcal{A} + \mathcal{G}\} \mathcal{V}_{t+\nu-1}(\xi^*, \Omega_{t-1}), \quad \nu \geq 2. \end{aligned}$$

To determine confidence bands of the VIRFs, we use the Delta method. Accordingly, the asymptotic distribution of the VIRF function is given by

$$\sqrt{T} \left[\hat{\mathcal{V}}_{t+1}(\xi^*, \Omega_{t-1}) - \mathcal{V}_{t+1}(\xi^*, \Omega_{t-1}) \right] \xrightarrow{D} N\left(0, \nabla \mathcal{V}_{t+1}(\xi^*, \Omega_{t-1})^\top \Sigma \nabla \mathcal{V}_{t+1}(\xi^*, \Omega_{t-1})\right), \tag{8}$$

where $\Sigma = \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}$, see Equation 2, and $\hat{\mathcal{V}}_{t+1}(\xi^*, \Omega_{t-1})$ denotes the VIRF using estimated BEKK$(1,1,1)$ parameters and the subsequent H_t process. In case of Gaussian error terms, Σ reduces to $\mathcal{I}^{-1}$. The ∇ symbol stands for the first derivative operator.

3.3. VaR

In order to estimate risks of (speculative) portfolio returns, we support risk assessment within the package. For this purpose, we use the common VaR as measure to estimate the risks. It quantifies the loss of an asset or portfolio Z which is not exceeded with probability $1-\alpha$, i.e.,

$$\mathrm{VaR}_\alpha(Z) = F_Z^{-1}(\alpha) = \inf\{z \in \mathbb{R} : \; F_Z(z) > \alpha\},$$

where F_Z is the (estimated) CDF of Z. The most common distributional assumption for F is to impose normality. However, Patton (2004) and Chen, Fan, and Patton (2004), for instance, provide evidence against (conditional) normality of financial return series. We henceforth provide two alternative distributions in the package. First, we allow the user to use a centered, Student's t distribution, where the degrees of freedom are estimated by means of a moment estimator using the estimated BEKK residuals (Student 1908). *Given the consistency of the QML parameter estimates, this provides a consistent, albeit inefficient, estimate of the degrees of freedom parameter.* The second alternative is to use the empirical distribution function (eCDF) of the BEKK residuals.

To determine the most effective forecasting model and respective distributional assumptions, we provide a backtesting algorithm for VaR using a rolling window approach. In terms of backtesting, we include the **GAS** package (Ardia *et al.* 2019) and provide the user with common tests of unconditional and conditional coverage, such as given by Kupiec (1995) and Christoffersen (1998).

4. Package overview

The starting point to use the **BEKKs** package is the function `bekk_spec()`, where the user can specify the BEKK model type, the selection of initial values for log-likelihood optimization and the asymmetric patterns to account for. The default is set to the full symmetric BEKK model with fixed initial values for optimization. We will give a detailed description of the main function below.

4.1. Function for specification

The `bekk_spec()` function returns a so-called S3 class object '`bekkSpec`' which further can be passed to the methods for estimation, forecasting and simulation. The function takes the following arguments:

- `model`: A list with the following elements must be provided. First, `type` specifies the BEKK$(1,1,1)$ model type, i.e., one character string of full BEKK (`"bekk"`), diagonal BEKK (`"dbekk"`) or scalar BEKK (`"sbekk"`). The second list element is entitled `asymmetric` and must be a logical specifying whether a symmetric (default) (`asymmetric = FALSE`) or an asymmetric model (`asymmetric = TRUE`) shall be estimated, i.e., all model types can be either fitted in their symmetric or asymmetric form.

- `init_values`: Either specifying the method to obtain initial values for the BEKK optimization by method `bekk_fit()` or a vector of BEKK parameters when the '`bekkFit`' object is directly passed to `simulate()` for simulating a multivariate time series following a BEKK(1, 1, 1) process. Initial values for '`bekkFit`' during BHHH algorithm are set by default, i.e, by setting `init_values = NULL`, to a recursive grid search to obtain start values. Moreover, it can be either a numerical vector of suitable dimension such that the user may provide an own initial guess, or a character vector i.e., "random" to use a random starting value generator (set a seed in advance for reproducible results), or "simple" for relying on a simple initial values generator based on typical values for BEKK parameters found in the literature.[4] If the object from this function is passed to `simulate`, `"init_values"` are used as parameters for the data generating process.

- `signs`: A vector stacking 1 or -1 to identify the patterns for asymmetric volatility effects. Using `"1"` indicates that positive values have an additional effect on volatility, while `"-1"` states that negative values are considered to have an additional effect on volatility. The positioning of the 1, respective, -1 indicates the time series for which the asymmetric pattern is set. For example, a vector `c(-1, 1, -1)` in a three-dimensional model indicates that the asymmetric term is added to the volatility if negative returns of the first and third time series occur jointly with positive returns of the second time series. The default case is to assign a vector of -1 to `signs` such that jointly negative returns are assumed to exert an additional influence on conditional volatility by default.

- `N`: An integer specifying the dimension of the time series object. This is only needed when a '`bekkSpec`' object is used for simulating a MGARCH process by means of `bekk_spec()`.

It is worthwhile to mention that we have implemented the asymmetric model according to Grier *et al.* (2004), where the asymmetric volatility term only enters the conditional volatility if the condition imposed through `signs` is jointly fulfilled by the return vector e_t.

4.2. Function for estimation

Having set up '`bekkSpec`' object, the function `bekk_fit()` is used to estimate the parameters of the specified BEKK model for a given time series. The function returns a S3 class object '`bekkFit`'. The following variables are the input for the estimation procedure:

- `spec`: An object of class '`bekkSpec`'.

- `data`: A multivariate data object. Can be a numeric matrix or '`ts`'/'`xts`'/'`zoo`' object.

- `QML_t_ratios`: A logical variable. If `QML_t_ratios = TRUE`, the t ratios and standard errors of the BEKK parameter matrices are exactly calculated by means of second order derivatives. If `QML_t_ratios = FALSE`, the t ratios and standard errors are calculated conditionally on a correctly specified distribution of the BEKK residuals. For more details, see Section 2.

- `max_iter`: An integer specifying the maximum number of BHHH algorithm iterations.

[4] The C matrix is assumed to be diagonal, where the diagonal is set to $0.05\sum_{t=1}^{T}\text{diag}(e_t e_t^\top)/T$ and $\text{diag}(A) = 0.32$ and $\text{diag}(G) = 0.9$ with the off-diagonal elements set to zero.

- `crit`: Determines the precision of the BHHH algorithm.

4.3. Forecasting and simulation

Having specified and estimated a BEKK$(1,1,1)$ model, the resulting object of class '`bekkFit`' may be used for further analysis of the underlying time series data. First, the package allows to simulate a time series by means of the method `predict()`:

- `object`: A fitted BEKK$(1,1,1)$ model of class '`bekkFit`'.
- `n.ahead`: Number of periods to forecast conditional volatility. Default is a one-step ahead forecast.
- `ci`: Numeric point in $[0,1]$ defining the level for confidence bands of the conditional volatility forecast. Default is a 95% level confidence band.

For the simulation, the function `simulate()` can be used. The input `object` is either a '`bekkFit`' or a '`bekkSpec`' object. The length of the resulting series is set by `nsim`. The function then returns a simulated time series. For replication, a seed should be set in advance.

4.4. Estimating and backtesting portfolio risks

In order to estimate and forecast risks of a given time series by means of the BEKK$(1,1,1)$ model, we supply the user with a toolbox of methods for risks assessment by means of the VaR under different distributions. To estimate and forecast the VaR, the `VaR` function takes the following arguments:

- `x`: An object of class '`bekkFit`' from the function `bekk_fit()` or an object of class '`bekkForecast`' from the function `predict()` for which the VaR shall be calculated.
- `p`: A numerical value that determines the confidence level for the estimated risk (i.e., 1 minus the so-called VaR coverage level). The default value is set at 0.99 in accordance with the Basel Regulation (The Basel Committee on Banking Supervision 2011).
- `portfolio_weights`: A vector determining the portfolio weights to calculate the portfolio VaR. If set to `NULL`, univariate VaRs are calculated for each series.
- `distribution`: A character string determining the assumed distribution of the residuals. Implemented are `"normal"`, `"empirical"` and `"t"`. The default is set to the empirical distribution of the residuals. When using `"t"`, a Student t distribution is used for the marginal residuals' distribution. The degrees of freedom are estimated by a moment estimation using the empirical kurtosis of the marginal residuals.

To compare the risk forecasting performance of different models, the method `backtesting()` applies a rolling window approach for an estimated BEKK$(1,1,1)$ model. For this purpose, the data is divided into an in-sample and an out-of-sample period. Then, the method determines the VaR for each out-of-sample time instance. The user can specify the length, the forecasting horizon, distribution and define the number of cores for parallel computation. More specifically, the function takes the following input additionally to the inputs of the `VaR` method listed above:

- `window_length`: An integer specifying the length of the rolling window to which the model is fitted.
- `n.ahead`: Number of periods to forecast conditional volatility. Default is a one-period ahead forecast.
- `nc`: Number of cores to be used for parallel computation.

The function returns a series of estimated VaR values, the number of hits (i.e., time instances where the true loss was larger than the estimated VaR) and gives a series of test statistics for conditional and unconditional coverage.

4.5. Estimating volatility impulse response functions

The package allows for the estimation of volatility impulse response functions. For instance, we allow the user to specify the time instance, quantile of shock, and series where the shock occurs for which the VIRFs are calculated. The corresponding function `virf` takes the following input.

- `x`: An object of class '`bekkFit`'.
- `time`: Time instance for which the VIRF is determined.
- `q`: A number specifying the quantile to be considered for a shock on which basis VIRFs are generated.
- `n.ahead`: Number of periods to forecast conditional volatility. Default is a 10-period ahead forecast.
- `index_series`: An integer defining the number of series for which a shock is assumed.
- `ci`: A number defining the confidence level for the confidence bands.
- `time_shock`: Logical indicating if the estimated residuals at time instance specified by `time` shall be used as a shock.

5. Empirical illustration

To illustrate the functions and methods of the **BEKKs** package, we first illustrate the estimation of a dynamic covariance model. Subsequently, we describe the estimation of the VaR of a portfolio putting some more weight (60%) on the high yield assets (i.e., US equities) while less risky long-term US Treasury bonds account for 40% of the portfolio value. For these purposes we consider a symmetric BEKK model. Second, we add gold and crude oil to the system of considered assets. Apart from describing the second order moment structure by means of an asymmetric BEKK model, we estimate the VaR of an portfolio, with is composed of 60% stocks, 20% bonds, 10% gold and 10% crude oil.

5.1. Bivariate symmetric BEKK

The first step of the analysis is to load the **BEKKs** package into the workspace. Furthermore, the **quantmod** package (Ryan and Ulrich 2024) enables to download financial market

Method	class	Methods for class	Description
`bekk_spec`	‘`bekkSpec`’	`print`	Defines type of BEKK model, how to estimate initial values and asymmetric patterns to account for.
`bekk_fit`	‘`bekkFit`’	`plot`, `summary`, `print`, `logLik`, `residuals`	Estimates a BEKK model of type defined by `bekk_spec()` for a given time series.
`predict`	‘`bekkForecast`’		Forecasting volatility of time series using estimated parameters of class ‘`bekkFit`’.
`simulate`	‘`bekkSim`’		Simulating time series using estimates from `bekk_fit()` or predefined parameters of estimates from ‘`bekkSpec`’ objects.
`VaR`	‘`VaR`’	`plot`, `summary`, `print`	Estimates (forecasts) the VaR for given ‘`bekkFit`’ (‘`bekkForecast`’) objects.
`backtest`	‘`backtest`’	`plot`, `summary`, `print`	A function for backtesting a model of class ‘`bekkFit`’ in terms of VaR-forcasting accuracy by means of rolling windows.
`virf`	‘`virf`’	`plot`	Estimates volatility impulse responses and confidence bands of a ‘`bekkFit`’ object for a specified shock.
`portmanteau.test`	‘`portmanteau.test`’	`print`	Performing a Portmanteau test on remaining ARCH effects in the residuals of an object of class ‘`bekkFit`’.

Table 1: Package design of **BEKKs**.

data directly from Yahoo! Finance and **ggfortify** (Tang, Horikoshi, and Li 2016) allows for straightforward time series plotting.

```
R> library("BEKKs")
R> library("quantmod")
R> library("ggfortify")
```

Our hypothetical portfolio in this section consists of 60% of the SPDR S&P 500 Trust ETF, which trades on the NYSE under the symbol SPY. This fund is the largest exchange-traded

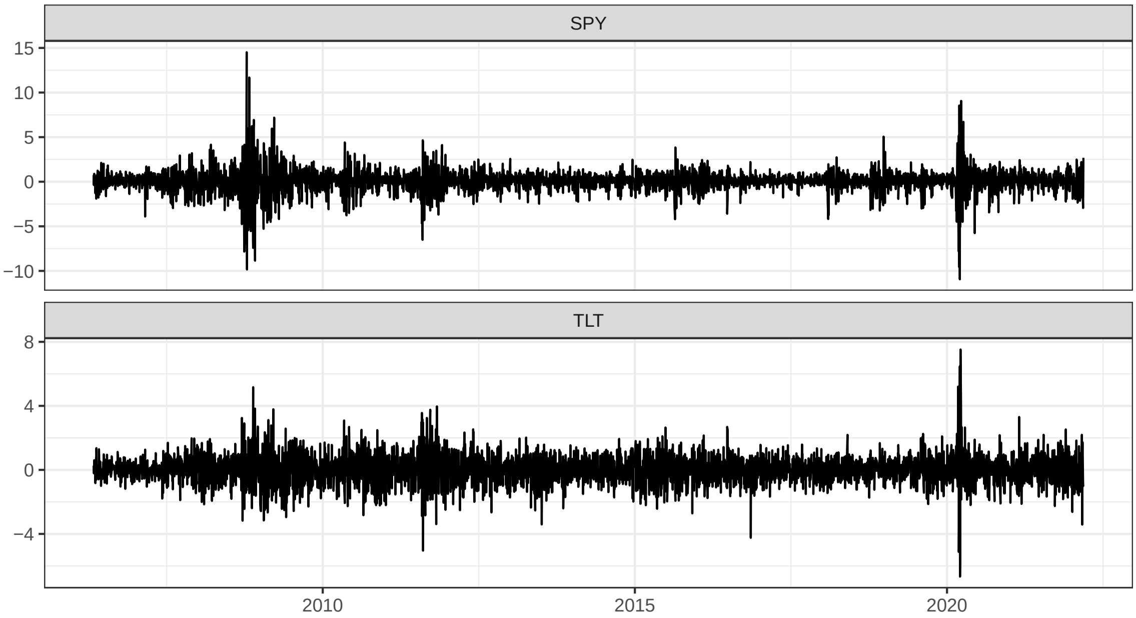

Figure 1: Daily percentage changes in S&P 500 ETF (SPY) and iShares 20+ Year Treasury Bond ETF (TLT).

fund in the world and has been designed to track the S&P 500 stock market index. The remaining 40% of our portfolio consists of the iShares 20+ Year Treasury Bond ETF (symbol: TLT). The fund invests in US dollar-denominated fixed-income US government bonds with a remaining maturity of at least twenty years and thus tracks the Barclays US 20+ Year Treasury Bond Index.

The following code downloads the daily prices of the two assets from Yahoo! Finance for the period May 1, 2006 until March 10, 2022 and converts the adjusted[5] closing prices into percentage changes.

```
R> assets <- c("SPY", "TLT")
R> getSymbols(assets, from = "2006-05-01", to = "2022-03-10",
+    src = "yahoo")
R> s_data <- na.omit(ROC(merge(Ad(SPY), Ad(TLT)), type = "discrete")) * 100
R> colnames(s_data) <- assets
R> autoplot(s_data) + theme_bw()
```

Figure 1 shows the daily returns in SPY and TLT. The next step is to use the **BEKKs** package to estimate the conditional standard deviations and correlations of the two assets. To do so, we specify a symmetric BEKK model by generating an object of class ‘`bekkSpec`’ with the default input of `bekk_spec()` and pass the object jointly with the data to the `bekk_fit()` method.

```
R> objBEKK1 <- bekk_spec()
```

[5]Adjusted closing prices are the closing prices after adjusting for all relevant splits and dividend distributions.

```
R> m1 <- bekk_fit(objBEKK1, s_data, QML_t_ratios = FALSE, max_iter = 50)
R> summary(m1)

BEKK estimation results
-----------------------
Log-likelihood: -9905.655
BEKK model stationary: TRUE
Number of BHHH iterations: 16
AIC: 19830.31
BIC: 19840.19
Estimated parameter matrices:

C
            [,1]      [,2]
[1,]  0.17060406 0.0000000
[2,] -0.04284237 0.0931289

A
            [,1]        [,2]
[1,]  0.36084828 -0.03193617
[2,] -0.02705291  0.24183921

G
            [,1]       [,2]
[1,] 0.918128943 0.01215393
[2,] 0.001007041 0.96584175

Standard errors of parameter matrices:

C
            [,1]        [,2]
[1,] 0.007205855 0.000000000
[2,] 0.011133739 0.009159382

A
           [,1]        [,2]
[1,] 0.01141225 0.007473649
[2,] 0.01218979 0.009893556

G
            [,1]        [,2]
[1,] 0.005048616 0.003120268
[2,] 0.004851338 0.003243111
```

To test the residuals for normality, we use the R package **tseries** for Jarque-Bera tests (Trapletti and Hornik 2024).

```
R> library("tseries")
R> jarque.bera.test(residuals(m1)[,1])
R> jarque.bera.test(residuals(m1)[,2])

Jarque Bera Test

data:  residuals(m1)[, 1]
X-squared = 1246, df = 2, p-value < 2.2e-16

Jarque Bera Test

data:  residuals(m1)[, 2]
X-squared = 246.97, df = 2, p-value < 2.2e-16

R> portmanteau.test(m1, lags = 5)
R> portmanteau.test(m1, lags = 15)
R> portmanteau.test(m1, lags = 30)

Portmanteau Test (Lags = 5)

data:  Residuals of estimated BEKK process
statistic = 0.71517, df = 12, p-value = 1

Portmanteau Test (Lags = 15)

data:  Residuals of estimated BEKK process
statistic = 1.9165, df = 52, p-value = 1

Portmanteau Test (Lags = 30)

data:  Residuals of estimated BEKK process
statistic = 3.6395, df = 112, p-value = 1
```

For notational convenience with existing packages, the model-implied residuals ξ_t are defined as `e_t` inside the package. Testing the model-implied residuals ξ_t for normality obtains a strong rejection of the null hypothesis. We have therefore chosen the option `QML_t_ratios = TRUE` (Bollerslev and Wooldridge 1992). In case that model residuals align with the normal distribution it is recommended to use ML t ratios by setting `QML_t_ratios = FALSE`. To indicate that the employed MGARCH model captures the dynamic second order properties of the data, a Portmanteau test for residual correlation of ξ_t as given in Section 3.1 obtains for $h = 5, 15$ and $h = 30$ test statistics (p values) of about 0.715 (1), 1.917 (1) and 3.640 (1), respectively. Accordingly, we do not find evidence for remaining ARCH effects in the residuals. Hence, there is no need for a BEKK model of higher order.

```
R> objBEKK1.1 <- bekk_spec()
R> m1.1 <- bekk_fit(objBEKK1.1, s_data, QML_t_ratios = TRUE, max_iter = 50)
R> summary(m1.1)
```

```
BEKK estimation results
-----------------------
Log-likelihood: -9905.655
BEKK model stationary: TRUE
Number of BHHH iterations: 16
AIC: 19830.31
BIC: 19840.19
Estimated parameter matrices:

C
            [,1]      [,2]
[1,]  0.17060406 0.0000000
[2,] -0.04284237 0.0931289

A
            [,1]        [,2]
[1,]  0.36084828 -0.03193617
[2,] -0.02705291  0.24183921

G
            [,1]       [,2]
[1,] 0.918128943 0.01215393
[2,] 0.001007041 0.96584175

Standard errors of parameter matrices:

C
           [,1]     [,2]
[1,] 0.01827122 0.000000
[2,] 0.02841909 0.015247

A
           [,1]       [,2]
[1,] 0.02715270 0.02058446
[2,] 0.04202871 0.02459968

G
           [,1]        [,2]
[1,] 0.01094254 0.008801889
[2,] 0.01465061 0.007976558
```

The following command will generate the plot in Figure 2, which shows that the two ETFs are negatively correlated for the majority of time instances. Moreover, the S&P 500 ETF exhibits a larger standard deviation compared with the bond ETF throughout.

```
R> plot(m1.1)
```

The next step is to calculate the VaR of the two assets combined according to a predetermined

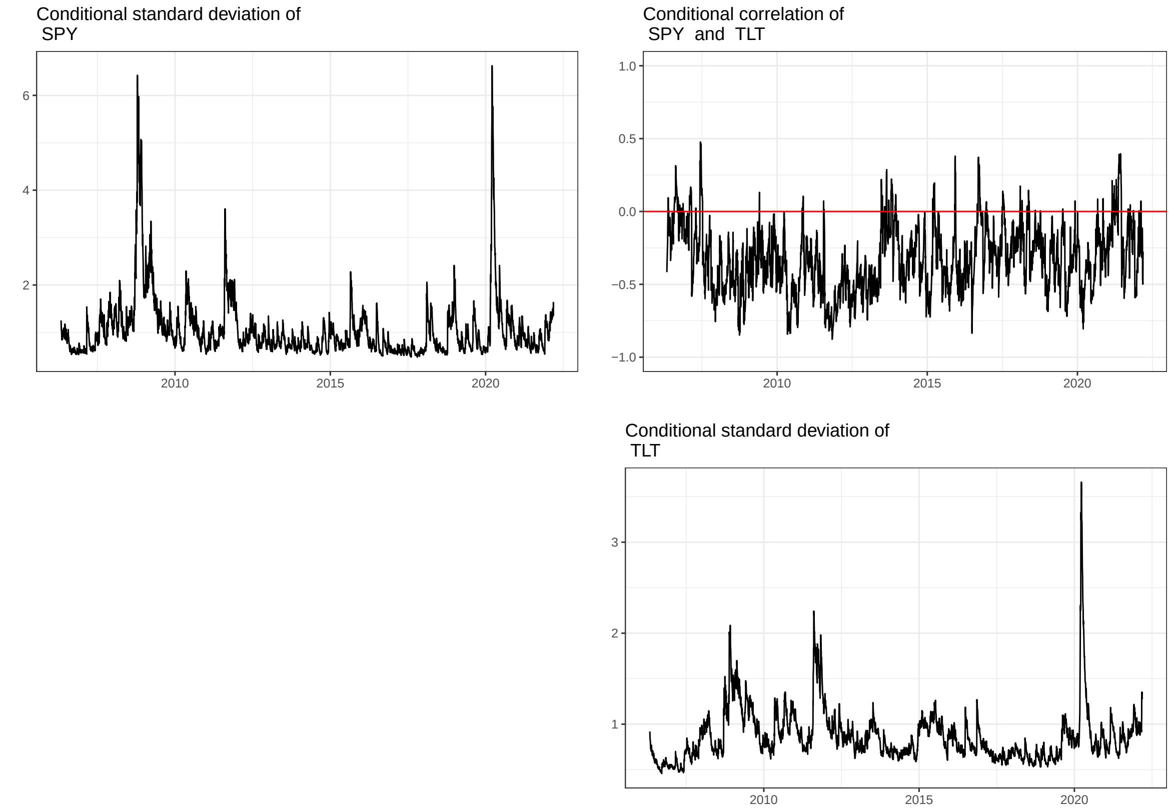

Figure 2: Estimated standard deviation and correlation of two-dimensional symmetric BEKK model.

60/40 weighting, which is a widely-used portfolio setup for stocks and bonds. The method `VaR()` is designed to either process an object of class '`bekkFit`' or '`bekkForecast`' and computes the VaR which is not exceeded with confidence of 99% by default, which corresponds to a 1% VaR coverage.

```
R> m1.1_var <- VaR(m1.1, p = 0.99, portfolio_weights = c(0.6, 0.4),
+    distribution = "empirical")
R> plot(m1.1_var)
```

The resulting plot in Figure 3 depicts the estimated VaR using the empirical distribution of the model-implied residuals of the considered portfolio with two strongly outlying estimates, i.e., (i) a VaR up to about −7% on the occasion of the financial crisis in 2008/09 and (ii) a VaR of about −6% occurring at the date of the COVID outbreak in March 2020.[6]

5.2. Large four-dimensional asymmetric model

In the second part of the empirical application, we augment our portfolio with two additional assets, i.e., (i) the SPDR Gold Shares ETF, the largest fund backed by physical bullion

[6]In the present illustration the VaR describes a critical negative return. In practice, VaR is often communicated in the form of a minimum positive loss for a given portfolio as implied by the negative return considered here. Hence, both measures are related in a one-to-one manner.

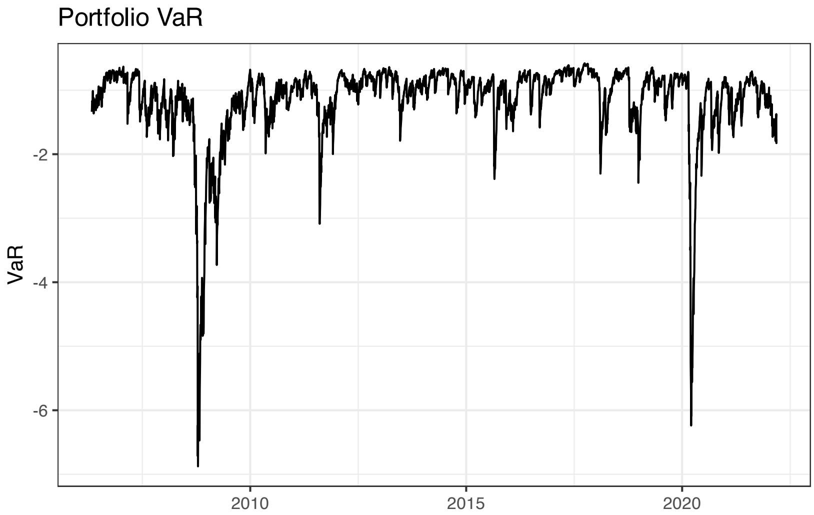

Figure 3: Estimated VaR of two-dimensional portfolio consisting of 40% treasury bond ETF and 60% S&P 500 ETF.

tracking the price of a tenth of an ounce of gold (symbol: GLD), and (ii) the US Oil fund tracking the price of West Texas Intermediate (WTI) crude oil by investing in oil future contracts that are traded on regulated futures exchanges (symbol: USO). In this section, we consider a portfolio which is weighting the S&P 500 ETF by 60%, 20% iShares 20+ Year Treasury Bond ETF and 10% for each gold and crude oil.

In analogy to the illustration above, we first download the data from Yahoo! Finance and convert the adjusted closing prices into percentage changes.

```
R> assets <- c("SPY", "TLT", "GLD", "USO")
R> getSymbols(assets, from = "2006-05-01", to = "2022-03-10",
+    src = "yahoo")
R> s_data_4dim <- na.omit(ROC(merge(Ad(SPY), Ad(TLT), Ad(GLD), Ad(USO)),
+    type = "discrete")) * 100
R> colnames(s_data_4dim) <- assets
```

In order to assess the performance of the new **BEKKs** package in higher dimensions (here $N = 4$), we compare the estimation time with competing packages.[7] As competing packages we take **mgarchBEKK** and **bmgarch**, since they support estimation of dimension $N > 3$ unlike **MTS**. The estimation time of **BEKKs** is 11.06 seconds, where **mgarchBEKK** takes 1118 and **bmgarch** 20372 seconds. Thus, we find that the new package is more than 100 times faster than the **mgarchBEKK** and more than 1000 times faster than the estimation using the **bmgarch** package.

```
R> library("mgarchBEKK")
R> library("bmgarch")
R> spec <- bekk_spec()
R> system.time(bekk_fit(spec, s_data_4dim, max_iter = 150))
```

[7] The details of the used hardware may be found in the section *Computational details* at the end of this paper.

```
       User      System     elapsed
      10.88        0.07       11.06
```

```
R> system.time(BEKK(as.matrix(s_data_4dim)))
```

```
H IS SINGULAR!...
H IS SINGULAR!...
H IS SINGULAR!...
       User      System     elapsed
    1117.61        0.22     1118.74
```

```
R> system.time(bmgarch(s_data_4dim, parameterization = "BEKK"))
```

```
       User      System       elapsed
      20371        0.07       20372
```

In contrast to the illustration above, we take an agnostic approach to model selection and fit all possible model specifications from the **BEKKs** package to the data. To do so, we explicitly define the alternative model type (BEKK, diagonal BEKK, or scalar BEKK) and asymmetry in the `bekk_spec()` function and fit for each specification a separate model. More specifically, we set up a pattern of negative stock and oil returns along with positive bond and gold returns to have an additional effect on volatility, i.e., we set `signs = c(-1, 1, 1, -1)`. This pattern is often seen in times of crisis, when investors seek for safe-havens and oil is depreciating due to weaker economic prospects. We compare the results with the widely-used pattern of jointly negative returns having an additional effect on volatility. We then compare the AIC and BIC of each model.

```
R> objBEKK2 <- bekk_spec(model = list(type = "bekk", asymmetric = TRUE))
R> m2 <- bekk_fit(objBEKK2, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter =  150)
R> objBEKK2.1 <- bekk_spec(model = list(type = "bekk", asymmetric = TRUE),
+    signs = c(-1, 1, 1, -1))
R> m2.1 <- bekk_fit(objBEKK2.1, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> summary(m2.1)
```

```
Asymmetric BEKK estimation results
----------------------------------
Log-likelihood: -23270.14
BEKK model stationary: TRUE
Number of BHHH iterations: 46
AIC: 46646.29
BIC: 46671.21
Estimated parameter matrices:

C
```

```
             [,1]       [,2]       [,3]      [,4]
[1,]  0.143220212 0.00000000 0.00000000 0.0000000
[2,] -0.018282995 0.08976385 0.00000000 0.0000000
[3,] -0.007197599 0.05198730 0.08337374 0.0000000
[4,]  0.036104166 0.06193355 0.03522780 0.2420883

A
            [,1]         [,2]         [,3]         [,4]
[1,] 0.305433822 -0.058274085 -0.023726806 -0.001634063
[2,] 0.020688588  0.169650389 -0.009010672 -0.036481854
[3,] 0.032359126 -0.010582467  0.189038919  0.024637101
[4,] 0.004239984 -0.007096682  0.001770778  0.242476169

B
             [,1]       [,2]       [,3]        [,4]
[1,]  0.199025005 0.05561223 0.03376652 -0.38415482
[2,] -0.082643329 0.14513009 0.15694756 -0.36076297
[3,] -0.150412606 0.09004088 0.01537866 -0.02505958
[4,]  0.008738239 0.02696692 0.02805908  0.30453009

G
              [,1]         [,2]         [,3]         [,4]
[1,]  0.9320597218  0.016462890  0.004344744 0.0007670556
[2,] -0.0103557485  0.977540050 -0.005435353 0.0025012306
[3,]  0.0009720747  0.001546335  0.978516092 0.0022435923
[4,] -0.0015203745 -0.001164771 -0.001046889 0.9567943453


Standard errors of parameter matrices:

C
           [,1]       [,2]       [,3]       [,4]
[1,] 0.02125217 0.00000000 0.00000000 0.00000000
[2,] 0.02977004 0.01669054 0.00000000 0.00000000
[3,] 0.02794263 0.03327243 0.02129155 0.00000000
[4,] 0.06154286 0.06220006 0.12320291 0.05171927

A
           [,1]        [,2]       [,3]       [,4]
[1,] 0.04670133 0.016563520 0.03391756 0.09305551
[2,] 0.06636472 0.024081310 0.06658100 0.07387436
[3,] 0.03009048 0.014687145 0.03563707 0.02537361
[4,] 0.01343847 0.006927696 0.01001872 0.04471055

B
           [,1]       [,2]       [,3]       [,4]
[1,] 0.15638471 0.08469091 0.23136129 0.26563482
```

```
[2,] 0.30550451 0.17982124 0.20675802 0.23479504
[3,] 0.08976606 0.12952582 0.38409450 0.23360952
[4,] 0.10278793 0.03145999 0.06577739 0.07582492

G
            [,1]        [,2]        [,3]        [,4]
[1,] 0.016161058 0.006232534 0.007524096 0.035420382
[2,] 0.015241503 0.006522420 0.010491911 0.021164555
[3,] 0.009629733 0.002977816 0.006423934 0.007477714
[4,] 0.005173787 0.001977315 0.002675551 0.013824837
```

```
R> objBEKK2.2 <- bekk_spec(model = list(type = "bekk", asymmetric = FALSE))
R> m2.2 <- bekk_fit(objBEKK2.2, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> objBEKK2.3 <- bekk_spec(model = list(type = "dbekk", asymmetric = FALSE))
R> m2.3 <- bekk_fit(objBEKK2.3, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> objBEKK2.4 <- bekk_spec(model = list(type = "dbekk", asymmetric = TRUE))
R> m2.4 <- bekk_fit(objBEKK2.4, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> objBEKK2.5 <- bekk_spec(model = list(type = "sbekk", asymmetric = FALSE))
R> m2.5 <- bekk_fit(objBEKK2.5, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> objBEKK2.6 <- bekk_spec(model = list(type = "sbekk", asymmetric = TRUE))
R> m2.6 <- bekk_fit(objBEKK2.6, s_data_4dim, QML_t_ratios = TRUE,
+    max_iter = 150)
R> logLik(m2, m2.1, m2.2, m2.3, m2.4, m2.5, m2.6)
```

```
  df       LLV      AIC      BIC
1 58 -23344.76 46795.51 46820.43
2 58 -23270.14 46646.29 46671.21
3 42 -23383.73 46841.46 46866.39
4 18 -23430.92 46875.84 46952.76
5 22 -23413.24 46842.48 46921.40
6 12 -23483.40 46980.80 47051.72
7 13 -23483.80 46983.60 47053.52
```

AIC as well as BIC show support for the asymmetric BEKK model with `signs = c(-1, 1, 1, -1)` as optimal choice given the data. This implies that higher volatility is seen in times of jointly negative stock and oil returns along with positive gold and bond returns. Accordingly, we find evidence for flight-to-safety (movements from stocks and oil towards gold and bonds) in times of higher uncertainty in financial markets. This underlines the practical importance of accounting for different asymmetric patterns as outlined in Grier *et al.* (2004). The code below provides a brief summary of the chosen model including a graphical display of model-implied second order moments in Figure 4.

```
R> plot(m2.1)
```

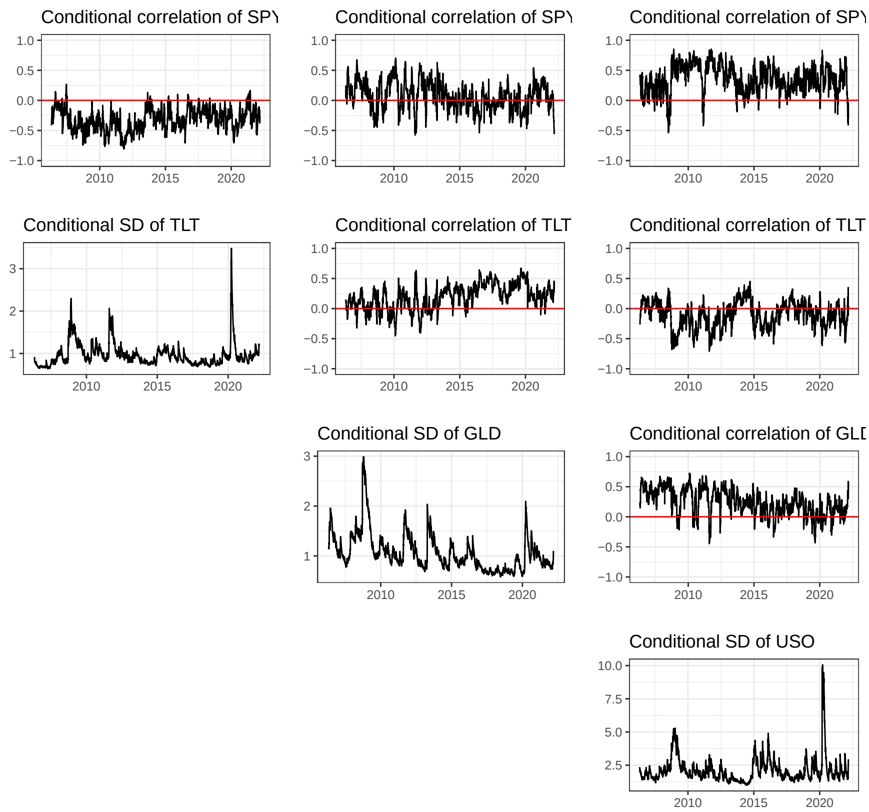

Figure 4: Estimated standard deviations and correlations of best-fitting four-dimensional asymmetric BEKK model with `signs = c(-1, 1, 1, -1)`.

Based on this model we can calculate the VaR for a weighted portfolio by:

```
R> m2.1_var <- VaR(m2.1, p = 0.99,
+    portfolio_weights = c(0.6, 0.2, 0.1, 0.1),
+    distribution = "t")
```

and graphically display the VaR by (see Figure 5):

```
R> plot(m2.1_var)
```

To compare the one-day-ahead VaR-forecasting performance of those models, we apply a backtest procedure for the 99% coverage level. For the in-sample, we choose a window length of 10 years which equals roughly 2600 trading days. We assume a Student's t distribution for the marginal distributions of the BEKK-implied residuals.

```
R> m2.1_backtest <- backtest(m2.1, window_length = 2600, p = 0.99,
+    portfolio_weights = c(0.6, 0.2, 0.1, 0.1), n.ahead = 1, nc = 8,
+    distribution = "t")
R> summary(m2.1_backtest)

Asymmetric BEKK backtesting results
-----------------------------------
```

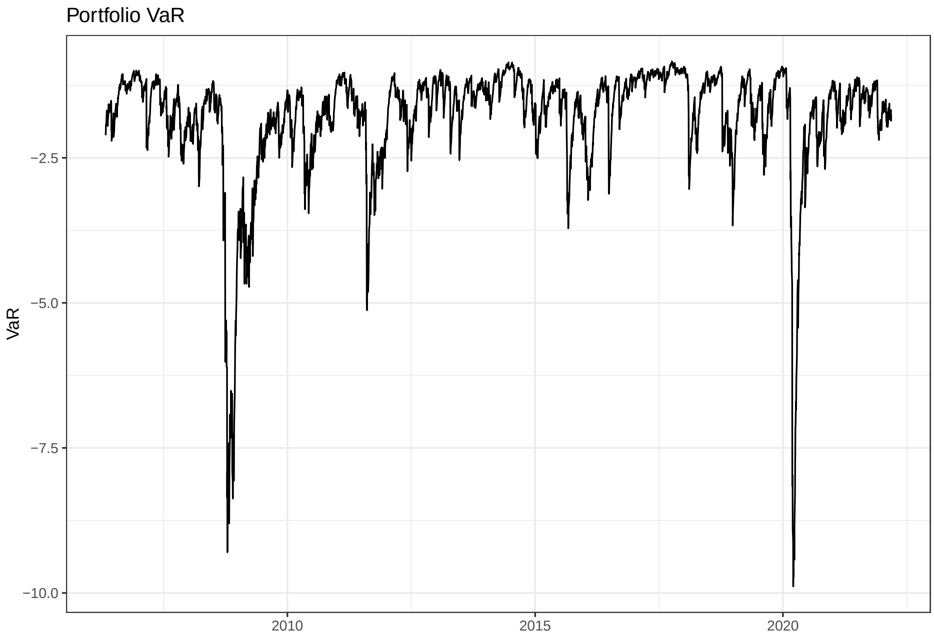

Figure 5: Estimated daily 1% VaR implied by the best fitting BEKK model of the four-dimensional portfolio with weights 60/20/10/10 % assuming Student's t distribution for the residuals.

```
Value-at-risk confidence level: 0.99
Window length: 2600
Portfolio weights: 0.6 0.2 0.1 0.1
-----------------------------------
Hit rate: 0.017

Unconditional coverage test of Kupiec:

Test:    4.9996666
p-value: 0.0253522

conditional coverage test of Christoffersen:

Test:    12.789464881
p-value:  0.001670333

R> m2.6_backtest <- backtest(m2.6, window_length = 2600, p = 0.99,
+    portfolio_weights = c(0.6, 0.2, 0.1, 0.1), n.ahead = 1, nc = 8,
+    distribution = "t")
R> summary(m2.6_backtest)

Asymmetric scalar BEKK backtesting results
------------------------------------------
Value-at-risk confidence level: 0.99
```

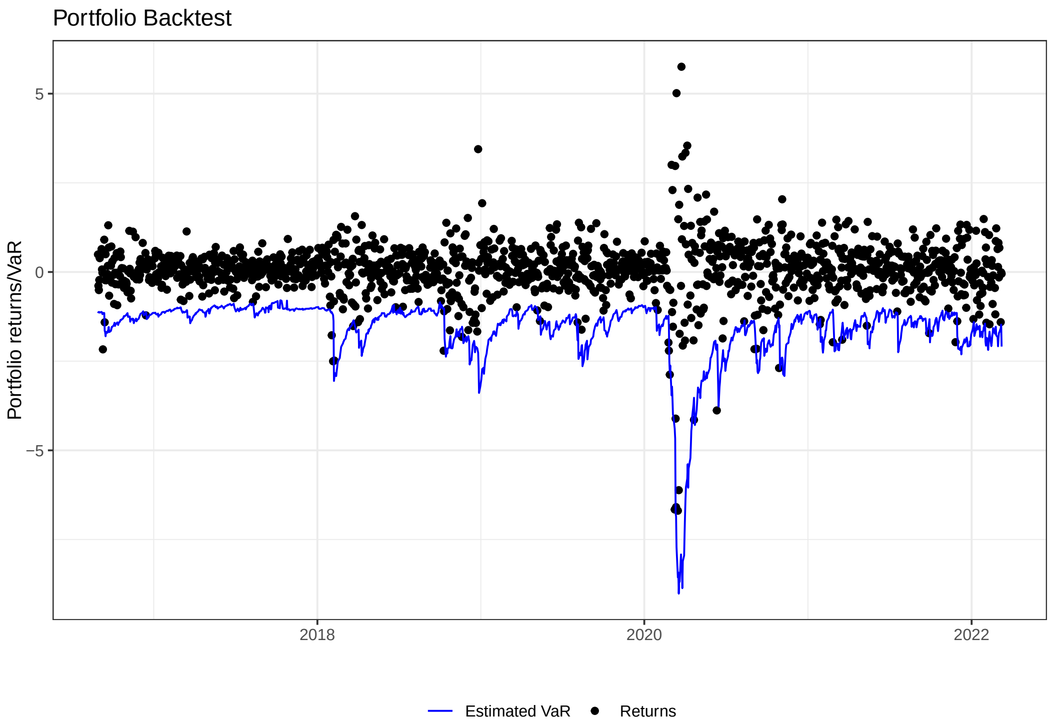

Figure 6: Backtesting results of the four-dimensional portfolio with weights 60/20/10/10 % for the best-fitting asymmetric BEKK model.

```
Window length: 2600
Portfolio weights: 0.6 0.2 0.1 0.1
------------------------------------------
Hit rate: 0.019

Unconditional coverage test of Kupiec:

Test:     8.434235788
p-value: 0.003682216

conditional coverage test of Christoffersen:

Test:     11.250187561
p-value:  0.003606225
```

We find that the asymmetric BEKK model leads to better one-period-ahead forecasts than the asymmetric scalar BEKK model. For the 1% conditional VaR, we find 17 violations of the estimated risk level out of 1000 forecasts. This is also underlined by the p values of the test for unconditional and conditional coverage which are somewhat larger for the full asymmetric BEKK model. In Figure 6, we display the estimated one-day-ahead VaR forecasts joint with realized portfolio returns by:

```
R> plot(m2.1_backtest)
```

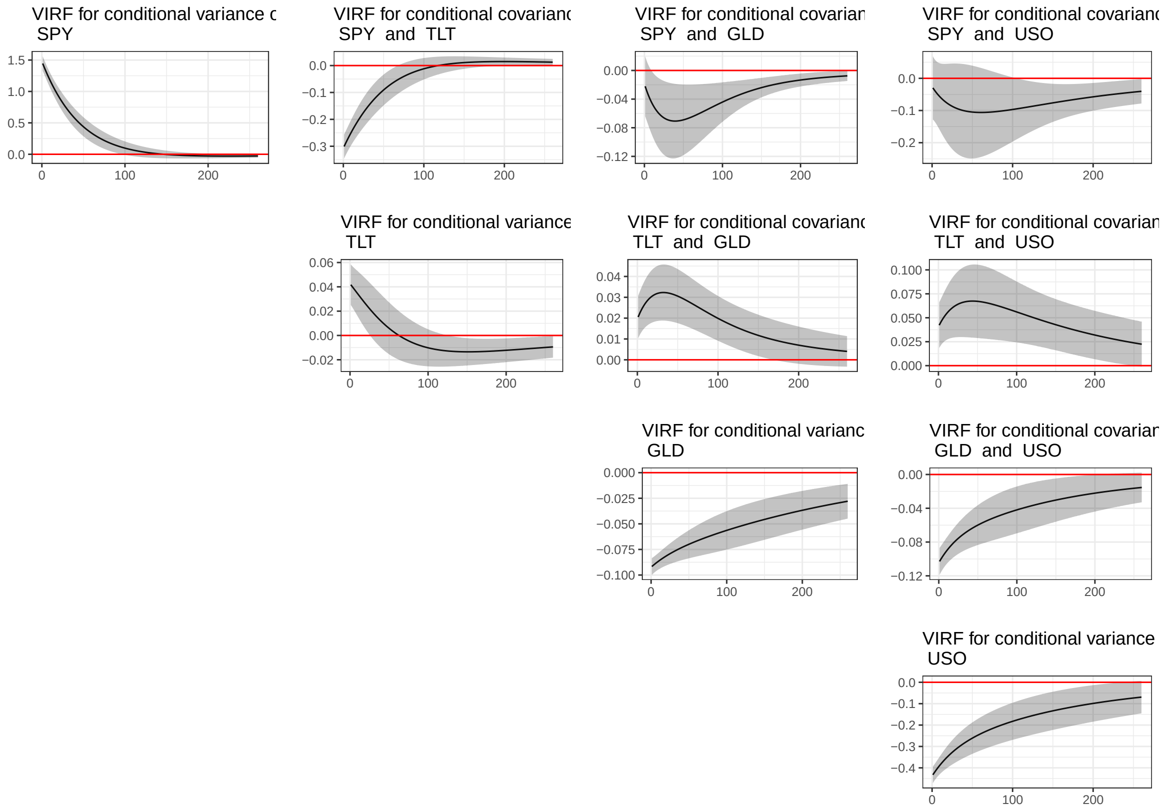

Figure 7: VIRFs for a shock in the S&P 500 index. 90% Confidence bands are displayed as grey shaded area.

The asymmetric BEKK model with t distribution provides a solid risk forecasting performance. In particular, it performs accurately during the COVID crisis in March 2020. To demonstrate the estimation of VIRFs within the **BEKKs** package, we estimate the VIRFs for the day of the Lehman default (September 15, 2008) assuming a shock in the S&P 500 index which is equal to the lowest 1% of the estimated residuals of the S&P 500 series.

```
R> m2.2_virf <- virf(m2.2, time = "2008-09-15", q = 0.01, index_series = 1,
+    n.ahead = 260, ci = 0.9, time_shock = FALSE)
```

The influences on second order moments are displayed for a time horizon of roughly one year, i.e., 260 trading days. Along with the VIRFs, we plot 90% confidence intervals as implied by the Delta method (see Equation 8) in Figure 7 by:

```
R> plot(m2.2_virf)
```

The shock results in an increased volatility of the S&P 500 index. Whereas the volatility of so-called safe-havens (here: gold and US Treasury bonds) is less affected by the shock originating in the stock market. This underlines the stability associated with these assets. Moreover, the covariance between the safe-havens and stocks is initially reduced by the shock which might end up in a negative correlation being in line with the safe-haven characteristic in times of turmoil.

6. Conclusion

We have introduced the R package **BEKKs** that allows for fast estimation, simulation and forecasting of conditional volatilities and risks of multivariate time series of speculative returns. We explained the main functions of the package by performing an investigation of a multivariate time series consisting of stocks, commodities and bonds. An evaluation of that four dimensional time series using competing packages demonstrates that the **BEKKs** package is substantially faster. Yet, we have concentrated on BEKK models of order $(1, 1)$, since these models are by far the most applied. Moreover, potential likelihood gains of higher order BEKK models, say for instance BEKK$(2, 2)$, are in conflict with an even more increased curse of dimensionality. Accordingly, we leave such extensions as a topic for future work. Moreover, including the package into existing tools such as **PerformanceAnalytics** (Peterson and Carl 2020) could help to offer manifold portfolio analysis using the flexibility of full BEKK models to an even broader usership.

Computational details

The results in this paper were obtained using R version 4.1.1 (R Core Team 2024) with the packages **BEKKs** version 1.4.5 (Fülle *et al.* 2024), **bmgarch** version 1.1.0 (Rast and Martin 2023) and **mgarchBEKK** version 0.0.2 (Schmidbauer *et al.* 2022).

Computations were performed on Windows 10 x86 64-w64-mingw32/x64 (64-bit) with processor Intel Core i7-8850H CPU @ 2.60GHz, 2592 MHz, 6 core(s), 12 logical(s).

References

Aielli GP (2013). "Dynamic Conditional Correlation: On Properties and Estimation." *Journal of Business & Economic Statistics*, **31**(3), 282–299. doi:10.1080/07350015.2013.771027.

Ardia D, Boudt K, Catania L (2019). "Generalized Autoregressive Score Models in R: The **GAS** Package." *Journal of Statistical Software*, **88**(6), 1–28. doi:10.18637/jss.v088.i06.

Baba Y, Engle RF, Kraft DF, Kroner KF (1983). "Multivariate Simultaneous Generalized ARCH Models." *Technical report*, Department of Economics, University of California, San Diego. URL http://www.jstor.org/stable/3532933.

Berndt ER, Hall BH, Hall RE, Hausman JA (1974). "Estimation and Inference in Nonlinear Structural Models." *Annals of Economic and Social Measurement*, **3**, 653–666. doi:10.1016/j.enconman.2023.117831.

Bollerslev T (1986). "Generalized Autoregressive Conditional Heteroscedasticity." *Journal of Econometrics*, **31**, 307–327. doi: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. doi:10.2307/2109358.

Bollerslev T, Engle RF, Wooldridge JM (1988). "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy*, **96**, 116–131. doi:10.1086/261527.

Bollerslev T, Wooldridge J (1992). "Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances." *Econometric Reviews*, **11**, 143–172. doi:10.1080/07474939208800229.

Caporin M, McAleer M (2008). "Scalar BEKK and Indirect DCC." *Journal of Forecasting*, **27**(4), 537–549. doi:10.1002/for.1074.

Caporin M, McAleer M (2012). "Do We Really Need Both BEKK and DCC? A Tale of Two Multivariate GARCH Models." *Journal of Economic Surveys*, **26**(4), 736–751. doi:10.1111/j.1467-6419.2011.00683.x.

Cappiello L, Engle RF, Sheppard K (2006). "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns." *Journal of Financial Econometrics*, **4**(4), 537–572. doi:10.1093/jjfinec/nbl005.

Chen X, Fan Y, Patton AJ (2004). "Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to US Equity Returns and Exchange Rates." *Working Paper 483*, London Economics Financial Markets Group. doi:10.2139/ssrn.513024.

Christoffersen PF (1998). "Evaluating Interval Forecasts." *International Economic Review*, **39**(4), 841–862. doi:10.2307/2527341.

Comte F, Lieberman O (2003). "Asymptotic Theory for Multivariate GARCH Processes." *Journal of Multivariate Analysis*, **84**, 61–84. doi:10.1016/s0047-259x(02)00009-x.

Ding Z, Engle RF (2001). "Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing." *NYU Working Paper Fin-01-029*, New York University. doi:10.2139/ssrn.4727500.

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

Engle RF (2001). "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives*, **15**(4), 157–168. doi:10.1257/jep.15.4.157.

Engle RF (2002). "Dynamical Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroscedastic Models." *Journal of Business and Economic Statistics*, **20**(3), 339–350. doi:10.1198/073500102288618487.

Engle RF, Kroner KF (1995). "Multivariate Simultaneous Generalized ARCH." *Econometric Theory*, **11**, 122–150. doi:10.1017/s0266466600009063.

Estima (2024). ***RATS**: Regression Analysis of Time Series*. URL https://www.estima.com/.

Fülle MJ, Lange A, Hafner CM, Herwartz H (2024). ***BEKKs**: Multivariate Conditional Volatility Modelling and Forecasting*. doi:10.32614/CRAN.package.BEKKs. R package version 1.4.5.

Gallant AR, Rossi PE, Tauchen G (1993). "Nonlinear Dynamic Structures." *Econometrica*, **61**(4), 871–907. doi:10.2307/2951766.

Ghalanos A (2022). ***rmgarch**: Multivariate GARCH Models*. doi:10.32614/CRAN.package.rmgarch. R package version 1.3-9.

Ghalanos A (2024). ***rugarch**: Univariate GARCH Models*. doi:10.32614/CRAN.package.rugarch. R package version 1.5-3.

Grier KB, Henry OT, Olekalns N, Shields K (2004). "The Asymmetric Effects of Uncertainty on Inflation and Output Growth." *Journal of Applied Econometrics*, **19**(5), 551–565. doi:10.1002/jae.763.

Hafner CM (2008). "GARCH Modeling." In RA Meyers (ed.), *Encyclopedia of Complexity and Systems Science*. Springer-Verlag.

Hafner CM, Herwartz H (2006). "Volatility Impulse Responses for Multivariate GARCH Models: An Exchange Rate Illustration." *Journal of International Money and Finance*, **25**, 719–740. doi:10.1016/j.jimonfin.2006.04.006.

Hafner CM, Herwartz H (2008). "Analytical Quasi Maximum Likelihood Inference in Multivariate Volatility Models." *Metrika*, **67**, 219–239. doi:10.1007/s00184-007-0130-y.

Hafner CM, Preminger A (2009). "On Asymptotic Theory for Multivariate GARCH Models." *Journal of Multivariate Analysis*, **100**(9), 2044–2054. doi:10.1016/j.jmva.2009.03.011.

IHS Markit (2021). *EViews 12 for Windws*. Irvine, CA. URL http://www.EViews.com/.

Jorion P (1996). "Risk2: Measuring the Risk in Value at Risk." *Financial Analysts Journal*, **52**(6), 47–56. doi:10.2469/faj.v52.n6.2039.

Jorion P (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw Hill, New York.

Kroner KF, Ng VK (1998). "Modeling Asymmetric Comovements of Asset Returns." *The Review of Financial Studies*, **11**, 817–844. doi:10.1093/rfs/11.4.817.

Kupiec PH (1995). "Techniques for Verifying the Accuracy of Risk Measurement Models." *The Journal of Derivatives*, **3**(2), 73–84. doi:10.3905/jod.1995.407942.

Li WK, McLeod AI (1981). "Distribution of the Residual Autocorrelations in Multivariate ARMA Time Series Models." *Journal of the Royal Statistical Society B*, **43**(2), 231–239. doi:10.1111/j.2517-6161.1981.tb01175.x.

Lütkepohl H (2007). *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin. Corrected 2nd printing.

Patton AJ (2004). "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation." *Journal of Financial Econometrics*, **2**(1), 130–168. doi:10.1093/jjfinec/nbh006.

Peterson BG, Carl P (2020). ***PerformanceAnalytics**: Econometric Tools for Performance and Risk Analysis*. doi:10.32614/CRAN.package.PerformanceAnalytics. R package version 2.0.4.

Rast P, Martin S (2023). ***bmgarch**: Bayesian Multivariate GARCH Models*. doi:10.32614/CRAN.package.bmgarch. R Package Version 2.0.0.

R Core Team (2024). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ryan JA, Ulrich JM (2024). ***quantmod**: Quantitative Financial Modelling Framework*. doi:10.32614/CRAN.package.quantmod. R package version 0.4.26.

Schmidbauer H, Roesch A, Tunalioglu VS (2022). ***mgarchBEKK**: Simulating, Estimating and Diagnosing MGARCH (BEKK and mGJR) Processes*. doi:10.32614/CRAN.package.mgarchBEKK. R package version 0.0.5.

Shephard N (ed.) (2004). *Stochastic Volatility: Selected Readings*. Oxford University Press.

Srivastavas P (2022). ***mgarch***. R package version 0.3.0, URL https://github.com/srivastavaprashant/mgarch.

StataCorp (2021). *Stata Statistical Software: Release 17*. StataCorp LLC, College Station, TX. URL https://www.stata.com/.

Student (1908). "The Probable Error of a Mean." *Biometrika*, **6**(1), 1–25. doi:10.1093/biomet/6.1.1.

Tang Y, Horikoshi M, Li W (2016). "**ggfortify**: Unified Interface to Visualize Statistical Result of Popular R Packages." *The R Journal*, **8**(2), 474–485. doi:10.32614/rj-2016-060.

Taylor SJ (1982). "Financial Returns Modelled by the Product of Two Stochastic Processes – A Study of Daily Sugar Prices, 1961–79." In OD Anderson (ed.), *Time Series Analysis: Theory and Practice 1*, pp. 203–226. North-Holland Publishing Company. doi:10.1093/oso/9780199257195.003.0003.

The Basel Committee on Banking Supervision (2011). "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems."

Trapletti A, Hornik K (2024). ***tseries**: Time Series Analysis and Computational Finance*. doi:10.32614/CRAN.package.tseries. R package version 0.10-58.

Tsay RS, Wood D (2022). ***MTS**: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*. doi:10.32614/CRAN.package.MTS. R package version 1.2.1.

Tse Y, Tsui A (2002). "A Multivariate Generalized Autoregressive Conditional Heteroscedastic Model with Time-Varying Correlations." *Journal of Business and Economic Statistics*, **20**(3), 351–362. doi:10.1198/073500102288618496.

Van Rossum G, *et al.* (2011). *Python Programming Language*. URL https://www.python.org/.

A. Second order derivatives

A.1. Symmetric BEKK

$$
\begin{aligned}
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(A)^\top \partial \operatorname{vec}(A)^\top} =& C_3 \left[\operatorname{vec}(I_N) \otimes \left(e_{t-1}e_{t-1}^\top \otimes I_N\right) K_{NN}\right] \\
&+ \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(A)^\top \partial \operatorname{vec}(A)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(A)^\top \partial \operatorname{vec}(G)^\top} =& \left(\left(\frac{H_{t-1}}{\partial \operatorname{vec}(A)^\top}\right)^\top \otimes I_{N^2}\right) C_1 \left[K_{NN} \otimes \operatorname{vec}\left(G^\top\right) + \operatorname{vec}\left(G^\top\right) \otimes K_{NN}\right] \\
&+ \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 H_{t-1}}{\partial \operatorname{vec}(A)^\top \partial \operatorname{vec}(G)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(A)^\top \partial \operatorname{vech}(C)^\top} =& 0 \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(A)^\top} =& C_3 \left[\operatorname{vec}(I_N) \otimes \left(I_N \otimes G^\top\right) \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(A)^\top}\right] \\
&+ \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(A)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(G)^\top} =& C_3 \left[\operatorname{vec}(I_N) \otimes \left\{(H_{t-1} \otimes I_N) K_{NN} + \left(I_N \otimes B^\top\right) \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(G)^\top}\right\}\right] \\
&+ \left(\left(\frac{\partial vec(H_{t-1})}{\partial \operatorname{vec}(G)^\top}\right)^\top \otimes I_{N^2}\right) C_1 \left[K_{NN} \otimes \operatorname{vec}\left(G^\top\right) + \operatorname{vec}\left(G^\top\right) \otimes K_{NN}\right] \\
&+ \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(G)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vech}(C)^\top} =& C_3 \left[\operatorname{vec}(I_N) \otimes \left\{\left(I_N \otimes G^\top\right) \frac{\partial vec(H_{t-1})}{\partial \operatorname{vech}(C)^\top}\right\}\right] \\
&+ \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vech}(C)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vec}(A)^\top} =& 0 \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vec}(G)^\top} =& \left(\left(\frac{vec(H_{t-1})}{\partial \operatorname{vech}(C)^\top}\right)^\top \otimes I_{N^2}\right) C_1 \left[K_{NN} \otimes \operatorname{vec}\left(G^\top\right) + \operatorname{vec}\left(G^\top\right) \otimes K_{NN}\right] \\
&+ \left[I_{N^*} \otimes (G \otimes G)^\top\right] \frac{\partial^2 H_{t-1}}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vec}(G)^\top} \\
\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vech}(C)^\top} =& 2\left(L_N \otimes D_N D_N^+\right) C_1 \left[I_{N^2} \otimes \operatorname{vec}(I_N)\right] L_N^\top \\
&+ \left[I_{N^*} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vech}(C)^\top}
\end{aligned}
$$

with

$$C_1 = I_N \otimes K_{NN} \otimes I_N$$
$$C_2 = 2\left(I_{N^2} \otimes D_N D_N^+\right)$$
$$C_3 = C_2 C_1$$

A.2. Additional derivatives for the asymmetric BEKK

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(B)^\top \partial \operatorname{vec}(B)^\top} = C_3\left[\operatorname{vec}(I_N) \otimes \left(\eta_{t-1}\eta_{t-1}^\top \otimes I_N\right) K_{NN}\right] + \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_t)}{\partial^2 \operatorname{vec}(B)^\top}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(B)^\top \partial \operatorname{vec}(G)^\top} = \left(\left(\frac{\partial H_{t-1}}{\partial \operatorname{vec}(B)^\top}\right)^\top \otimes I_{N^2}\right) C_1 \left[K_{NN} \otimes \operatorname{vec}\left(G^\top\right) + \operatorname{vec}\left(G^\top\right) \otimes K_{NN}\right] + \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 H_{t-1}}{\partial \operatorname{vec}(B)^\top \partial \operatorname{vec}(G)^\top}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(B)^\top \partial \operatorname{vech}(C)^\top} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(B)^\top} = C_3\left[\operatorname{vec}(I_N) \otimes \left(I_N \otimes G^\top\right) \frac{\partial H_{t-1}}{\partial \operatorname{vec}(B)^\top}\right] + \left[I_{N^2} \otimes (G \otimes G)^\top\right] \frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vec}(G)^\top \partial \operatorname{vec}(B)^\top}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vec}(B)^\top} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(B)^\top \partial \operatorname{vec}(A)^\top} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(A)^\top \partial \operatorname{vec}(B)^\top} = 0$$

A.3. Symmetric Scalar BEKK

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial a^2} = g \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial a^2}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial g^2} = 2 \frac{\partial \operatorname{vec}(H_{t-1})}{\partial g} + g \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial g^2}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial a \partial g} = g \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial a \partial g} + \frac{\partial \operatorname{vec}(H_{t-1})}{\partial a}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial a} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial g} = g \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vech}(C)^\top \partial g} + \frac{\partial \operatorname{vec}(H_{t-1})}{\partial \operatorname{vech}(C)^\top}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vech}(C)^\top} = 2\left(L_N \otimes D_N D_N^+\right) C_1 \left[I_{N^2} \otimes \operatorname{vec}(I_N)\right] L_N^\top + g \frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial \operatorname{vech}(C)^\top \partial \operatorname{vech}(C)^\top}.$$

A.4. Additional derivatives for the asymmetric scalar BEKK

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial b^2} = g\frac{\partial \operatorname{vec}(H_{t-1})}{\partial b^2}$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial \operatorname{vech}(C)^\top \partial b} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial a \partial b} = 0$$

$$\frac{\partial^2 \operatorname{vec}(H_t)}{\partial g \partial b} = \frac{\partial \operatorname{vec}(H_{t-1})}{\partial b} + g\frac{\partial^2 \operatorname{vec}(H_{t-1})}{\partial g \partial b}$$

Affiliation:

Markus J. Fülle, Alexander Lange, Helmut Herwartz
Chair of Econometrics
Faculty of Economics
Universität Göttingen
Humboldtallee 3, 37073 Göttingen, Germany
E-mail: fuelle@uni-goettingen.de

Christian M. Hafner
Louvain Institute of Data Analysis and Modelling
Université Catholique de Louvain
Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium
E-mail: christian.hafner@uclouvain.be

Journal of Statistical Software https://www.jstatsoft.org/
published by the Foundation for Open Access Statistics https://www.foastat.org/
November 2024, Volume 111, Issue 4 *Submitted:* 2023-04-15
doi:10.18637/jss.v111.i04 *Accepted:* 2024-04-20