Stat Comput (2012) 22:849–855
DOI 10.1007/s11222-011-9251-2

Cholesky-GARCH models with applications to finance

Petros Dellaportas · Mohsen Pourahmadi

Received: 22 February 2010 / Accepted: 13 December 2010 / Published online: 5 April 2011

Abstract Instantaneous dependence among several asset returns is the main reason for the computational and statistical complexities in working with full multivariate GARCH models. Using the Cholesky decomposition of the covariance matrix of such returns, we introduce a broad class of multivariate models where univariate GARCH models are used for variances of individual assets and parsimonious models for the time-varying unit lower triangular matrices. This approach, while reducing the number of parameters and severity of the positive-definiteness constraint, has several advantages compared to the traditional orthogonal and related GARCH models. Its major drawback is the potential need for an a priori ordering or grouping of the stocks in a portfolio, which through a case study we show can be taken advantage of so far as reducing the forecast error of the volatilities and the dimension of the parameter space are concerned. Moreover, the Cholesky decomposition, unlike its competitors, decompose the normal likelihood function as a product of univariate normal likelihoods with independent parameters, resulting in fast estimation algorithms. Gaussian maximum likelihood methods of estimation of the parameters are developed. The methodology is implemented for a real financial dataset with seven assets, and its forecasting power is compared with other existing models.

P. Dellaportas (✉)
Department of Statistics, Athens University of Economics and Business, Athens, Greece
e-mail: petros@aueb.gr

M. Pourahmadi
Department of Statistics, Texas A&M University, College Station, TX, 77843, USA
e-mail: pourahm@stat.tamu.edu

Keywords Autoregressive conditional heteroscedastic models · Latent factor models · Time-varying ARMA coefficients · Cholesky decomposition · Principal components · Spectral decomposition · Stochastic volatility models · Maximum likelihood estimation

1 Introduction

Many tasks of modern financial management including portfolio selection, option pricing and risk assessment can be reduced to the prediction of a sequence of large $N \times N$ covariance matrices $\{\Sigma_t\}$ based on the (conditionally) independently $N(0, \Sigma_t)$-distributed data Y_t, $t = 1, 2, \ldots, n$, where Y_t is the shock (innovation) at time t of a multivariate time series of driftless returns of N assets in a portfolio (Tsay 2005). Since Σ_t is positive-definite and its number of entries grows quadratically in N, the problem of parsimonious modeling of $\{\Sigma_t\}$ is truly challenging and has been studied earnestly in the literature of finance in the last three decades (Engle 1982, 2002; Tsay 2005). The key idea is to write difference equations for $\{\Sigma_t\}$ similar to the univariate autoregressive and moving average (ARMA) models. More precisely, with $\mathcal{F}_t$ standing for the past information up to and including the time t, Bollerslev (1986) defined the classes of *generalized autoregressive conditional heteroscedastic* (GARCH) models for a univariate returns series $\{y_t\}$ by

$$\begin{cases} y_t|\mathcal{F}_{t-1} \sim N(\mu_t, \sigma_t^2), \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \end{cases} \quad (1)$$

where the constraints $\alpha_0 > 0$ and $\alpha_i \geq 0$, $\beta_i \geq 0$, ensure a positive variance. Note that this general class of GARCH models reduces to the simpler class of ARCH models when all $\beta_j = 0$. For a discussion and review of progress in continuous-time stochastic volatility models see Stelzer (2010).

Emboldened by the ease of use and success of univariate GARCH models, many early variants of multivariate GARCH models (Engle and Kroner 1995) were defined simply as difference equations of the form (1) either for the vectorized sequence of covariance matrices $\{vec\, \Sigma_t\}$ or the sequence $\{\Sigma_t\}$ itself with suitable matrix coefficients. The number of free parameters of such models is known to grow proportional to N^4 and N^2, respectively. Consequently, for large covariance matrices the use of full multivariate GARCH models has proved impractical (Engle 2002; Tse and Tsui 2002). Simplification occurs (Alexander 2001, Chap. 7) when the coefficients are diagonal matrices, in which case, each variance/covariance term in Σ_t follows a univariate GARCH model with the lagged variance/covariance terms and squares and cross products of the data (Ledoit et al. 2003). However, complicated restrictions on the coefficient parameters are needed to guarantee their positive-definiteness. These restrictions are often too difficult to satisfy in the course of iterative optimization of the likelihood function even when the number of assets is about five.

In this paper, we provide parsimonious models for the time-varying covariance matrices $\{\Sigma_t\}$, by writing difference equations for the components of the Cholesky decomposition of Σ_t (Pourahmadi 1999; Tsay 2005; Smith and Kohn 2002). The new class of models are closely related to the standard and familiar factor models (Diebold and Nerlove 1989; Vrontos et al. 2003), and orthogonal GARCH models (Alexander 2001); for an extensive discussion on the interrelationship among these models and the role of ordering the stocks in a portfolio see Dellaportas and Pourahmadi (2004, Sect. 2), Chou et al. (2009) and Chang and Tsay (2010). Highly desirable and practical features of our approach compared to some of the existing methods are:

(a) Forecast consistency, in the sense that when new assets are added to the portfolio, the volatility forecasts of the original assets will be unchanged.
(b) The estimation of the volatility and dependence parameters are not separated.
(c) The Cholesky decomposition reduces multivariate volatility modeling to separate estimation of N univariate regression GARCH models.
(d) By viewing our models as full factor models (Sect. 2.1), exploitation of (c) above allows an efficient methodology to estimate the full factor models (Vrontos et al. 2003; Aguilar and West 2000), a fact that has not been used before.

The outline of the paper is as follows. In Sect. 2, we present the Cholesky decomposition of a covariance matrix and provide examples of structured and parsimonious models for the Cholesky factors. The roles of such models in handling volatility of financial data are discussed. It is shown that the multivariate conditional covariance estimation can be reduced to estimating the $3N$ parameters of univariate GARCH models and a few "dependence" parameters. A maximum likelihood procedure for estimation of the "dependence" and volatility parameters as well as the issue of ordering the stocks are presented in Sect. 3. In Sect. 4, for a data set of $N = 7$ daily exchange rates we study the impact of ordering the variables on the forecast error of the volatility and compare the performance of several traditional multivariate volatility models. Section 5 concludes the paper.

2 The Cholesky decomposition and GARCH models

We rely on the notion of regression to derive the Cholesky decomposition of a covariance matrix and hence motivate the use of a lower triangular matrix with unconstrained entries, instead of an orthogonal matrix in the orthogonal GARCH models (Alexander 2001). For the time being, we drop the subscript t in Y_t, Σ_t and focus on the contemporaneous covariance structure of a generic random vector $Y = (y_1, \dots, y_N)'$ and view $y_1, y_2, \dots, y_j, \dots, y_N$ as an ordered set of random variables indexed by j. Consider regressing y_j on its predecessors $y_1, \dots, y_{j-1}$:

$$y_j = \sum_{k=1}^{j-1} \phi_{jk} y_k + \varepsilon_j, \quad j = 1, 2, \dots, N, \tag{2}$$

where ϕ_{jk} and $\sigma_j^2 = var(\varepsilon_j)$ are the unique regression coefficients and residual variances with obvious statistical interpretations; by convention $\sum_{\emptyset} = 0$. Indeed, with $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_N)'$ and $\nu = cov(\boldsymbol{\varepsilon}) = diag(\sigma_1^2, \dots, \sigma_N^2)$, one can write (2) in the matrix form $TY = \boldsymbol{\varepsilon}$, where T is a unit lower triangular matrix with $-\phi_{jk}$ in the (j,k)th position, then it follows that the unit lower triangular matrix T diagonalizes Σ:

$$T \Sigma T' = \nu. \tag{3}$$

The pair of matrices (T, ν) are the components of the modified Cholesky decomposition of Σ (Pourahmadi 1999). For an unstructured covariance matrix, the nonredundant entries of T and $\log \nu$ are unconstrained and referred to as its *generalized autoregressive parameters* (GARP) and log-*innovation variances* (IV), respectively.

Since the inverse of T is also a unit lower triangular matrix, setting $T^{-1} = B = (\theta_{ij})$ it follows from (3) that

$$\Sigma = B\nu B' = \begin{pmatrix} \sigma_1^2 & \theta_{21}\sigma_1^2 & \theta_{31}\sigma_1^2 & \cdots & \theta_{N1}\sigma_1^2 \\ \theta_{21}\sigma_1^2 & \sum_{i=1}^2 \theta_{2i}^2\sigma_i^2 & \sum_{i=1}^2 \theta_{3i}\theta_{21}\sigma_i^2 & \cdots & \sum_{i=1}^2 \theta_{21}\theta_{Ni}\sigma_i^2 \\ \theta_{31}\sigma_1^2 & \sum_{i=1}^2 \theta_{2i}\theta_{3i}\sigma_i^2 & \sum_{i=1}^3 \theta_{3i}^2\sigma_i^2 & \cdots & \sum_{i=1}^3 \theta_{3i}\theta_{Ni}\sigma_i^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \theta_{N1}\sigma_1^2 & \sum_{i=1}^2 \theta_{2i}\theta_{Ni}\sigma_i^2 & \sum_{i=1}^3 \theta_{3i}\theta_{Ni}\sigma_i^2 & \cdots & \sum_{i=1}^N \theta_{Ni}^2\sigma_i^2 \end{pmatrix}. \quad (4)$$

This alternative representation of the Cholesky decomposition of Σ and its parameters are closely related to the moving average and factor models discussed next.

2.1 The factor model interpretation

Now we discuss the possibility of viewing the Cholesky decomposition (4) as a factor model. In fact, solving for Y from $TY = \varepsilon$, or regressing y_j on the past innovations $\varepsilon_1, \ldots, \varepsilon_{j-1}$, it follows that

$$y_j = \varepsilon_j + \sum_{k=1}^{j-1} \theta_{jk}\varepsilon_k, \quad j = 1, \ldots, N, \quad (5)$$

or in matrix form,

$$Y = B\boldsymbol{\varepsilon}. \quad (6)$$

It turns out that the representation (6) is closely related to the factor models. To this end, we partition the innovation vector $\boldsymbol{\varepsilon}$ and the matrix B so that (6) becomes

$$Y = B\boldsymbol{\varepsilon} = (B_1 \vdots B_2)\begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \cdots \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} = B_1\boldsymbol{\varepsilon}_1 + B_2\boldsymbol{\varepsilon}_2. \quad (7)$$

Now, think of $\boldsymbol{\varepsilon}_1$ as a $k \times 1$ vector of latent factors, B_1 the corresponding matrix of factor loadings, and $B_2\varepsilon_2$ as the vector of idiosyncratic errors. Then, the latent factors ε_1 have clear statistical interpretations as the first k innovations of Y and (7) has the appearance of a factor model. Note that for the extreme case of $k = N$, the vector of idiosyncratic errors in (7) is zero and it reduces to the full-factor representation of Y (Aguilar and West 2000; Vrontos et al. 2003).

2.2 Structured T matrices: parsimony

In this section we provide a class of structured covariance matrices with a small number of parameters. These matrices can be used, for example, in Bollerslev's (1990) constant-correlation models to reduce the number of correlation parameters from the maximum of $N(N-1)/2$ to as low as one or two. Similar structures for the lower triangular matrix T containing the GARPs in (2)–(3) will reduce the high number of parameters in the AR structures.

Consider a situation where T is a Toeplitz matrix or the entries along its subdiagonals are constant:

$$\phi_{i,i-j} = \phi_j, \quad j = 1, 2, \ldots, N-1;\ i = j+1, \ldots, N. \quad (8)$$

This structure reduces the number of parameters in T to $N-1$, which still could be large. If needed, one could further reduce the dimension of $(\phi_1, \phi_2, \ldots, \phi_{N-1})$ via parametric models formulated using the regressogram introduced in Pourahmadi (1999) or other graphical tools. When there are indications that ϕ_j's are monotone in j, then one may set, for example,

$$\phi_j = \gamma_0 + \gamma_1 j^{\pm k}, \quad j = 1, \ldots, N-1, \quad (9)$$

where γ_0, γ_1 are the two new parameters and k is a known positive integer. In applications where constancy along the subdiagonals of T is deemed inappropriate, one could exponentiate ϕ_j's by the Box-Cox transformation of the (time) index i along those subdiagonals. Namely, a non-Toeplitz T can be obtained by setting

$$\phi_{i,i-j} = \phi_j^{f(i;\lambda_j) - f(i-j;\lambda_j)},$$
$$j = 1, 2, \ldots, N-1;\ i = j+1, \ldots, N, \quad (10)$$

where

$$f(x;\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \log x & \text{if } \lambda = 0. \end{cases}$$

For example, if $0 < \phi_j < 1$, then the entries along the jth subdiagonal of T are monotone increasing if $\lambda_j < 1$, monotone decreasing if $\lambda_j > 1$, or constant if $\lambda_j = 1$. For other range of values of ϕ_j, similar nonconstant patterns could be prescribed depending on the values of the exponent λ_j. Here, the number of parameters $(\phi_1, \ldots, \phi_{N-1}, \lambda_1, \ldots, \lambda_{N-1})$ could be as small as 2 or as large as $2(N-1)$. For more examples of this type see Dellaportas and Pourahmadi (2004).

2.3 The Cholesky-GARCH models

Returning to the multivariate time series of returns $\{Y_t\}$, note that from the matrix form of (2) we have $\boldsymbol{\varepsilon}_t = T_t Y_t$, where the ε_{jt}'s are uncorrelated. Thus, motivated by the idea of orthogonal GARCH (Alexander 2001) and constant-correlation GARCH (Bollerslev 1990) models, we refer to the case when $T_t \equiv T$, and GARCH (p, q) models for the time-varying innovation variances $\{\nu_t\}$ are used, as *Cholesky-GARCH models* for the returns $\{Y_t\}$. Note that even in this simple case the number of parameters is $N(N-1)/2 + (p+q+1)N$ which is quite large for large data sets with, say, $N = 100$. This can be reduced considerably by selecting a structured T from the previous section, or from a longer list of such examples in Dellaportas and Pourahmadi (2004, Sect. 3). In the sequel for simplicity we fix $p = q = 1$.

A noteworthy feature of our Cholesky-GARCH models, which is immediate from (4), is that their correlation matrices are time-varying, and hence could be more flexible compared to the Bollerslev's (1990) constant-correlation GARCH models.

3 Estimation

In this section we present some of the conceptual underpinnings for estimation of the parameters of the Cholesky-GARCH models and the issue of ordering the stocks in a portfolio which is required in our approach.

3.1 The likelihood function

Assuming normality for the returns, the log-likelihood function, up to ignoring some irrelevant constants, is given by

$$\begin{aligned} L(\theta) &= \sum_{t=1}^{n}\left(\sum_{j=1}^{N}\log\sigma_{jt}^2 + Y_t' T_t' \nu_t^{-1} T_t Y_t\right) \\ &= \sum_{t=1}^{n}\sum_{j=1}^{N}\left(\log\sigma_{jt}^2 + \frac{\varepsilon_{jt}^2}{\sigma_{jt}^2}\right) \\ &= \sum_{j=1}^{N}\left\{\sum_{t=1}^{n}\left(\log\sigma_{jt}^2 + \frac{\varepsilon_{jt}^2}{\sigma_{jt}^2}\right)\right\}, \qquad (11) \end{aligned}$$

where θ is the vector of parameters in a Cholesky-GARCH model.

The last representation is the most convenient to use for computational purposes since it can be viewed as a sum of N univariate Gaussian likelihoods for the mutually uncorrelated (transformed) returns $\{\varepsilon_{jt}\}$, $j = 1, \ldots, N$. It is the source of a rather unique and appealing property of the Cholesky-GARCH models that is not available in any other multivariate time series models with time-varying innovation variances. Thus, estimation of the $N(N-1)/2$ parameters in T proceeds by estimating independently N univariate regression models with time-varying innovation variances. Details of an algorithm for finding MLE of the parameters and the asymptotic properties of the estimators can be found in Pourahmadi (2000, Sect. 2).

By contrast, in the orthogonal GARCH models the estimation of the orthogonal matrix and the volatilities are achieved in two steps, the matrix is usually estimated first as if the volatilities were time-independent, followed by estimation of the latter. Similarly, Vrontos et al. (2003) focused on representation (4) and estimated simultaneously all the parameters, without exploiting the fact that the GARP parameters $T = B^{-1}$ could be estimated faster via (11). From a practical perspective, Cholesky- GARCH models provide an easy and very competitive alternative to the existing multivariate models (Dellaportas and Pourahmadi 2004, Sect. 2), allowing quick and easy estimation for larger values of N. There is also the possibility of developing a sequential algorithm over time or when n gets large following the recent result in Chang and Tsay (2010). At the moment, the only drawback of the Cholesky-GARCH models seems to be the need for ordering the stocks in a portfolio which is discussed next.

3.2 Ordering the stocks

The Cholesky-GARCH models we proposed here is based on some sort of a priori ordering of the components of the return vector. While for a portfolio of N stocks there are $N!$ choices, the degree of non-uniqueness here is similar to the factor models and the choice of factor rotations (Geweke and Zhou 1996). As a statistical decision problem, ordering variables is quite challenging. In the Bayesian literature this very same problem has been investigated by Webb and Forster (2008) who presented an efficient reversible jump MCMC algorithm that searches over all models with different orderings. This could be readily implemented here by applying a Laplace approximation to each row of (2) as described by Vrontos et al. (2003) who provided a way to find the best ordering in a full factor model.

A default and simple method is to order the stocks according to the sample variances of their returns. As another possible alternative, we suggest a search algorithm based on some criterion such as AIC or BIC that can be readily implemented. First, note that GARP parameterization leads to the Cholesky-GARCH models in which the ordering problem is of order N^2 and not the usual order $N!$. To see this, note that for any $N > 1$ in (2), if y_j is picked up as a response variable the order of $y_1, y_2, \ldots, y_{j-1}$ is irrelevant. Thus, starting from $j = N$, there are N possible comparisons, via AIC or BIC, required to choose y_j. Subsequently, since y_N has been chosen and does not appear in the rest of the equations, y_{N-1} is chosen as the best model among the $N-1$ possible ways to write (2) for $k = 1, \ldots, N-1$, and so forth. Therefore, the number of required comparisons is $N(N-1)/2 - 1$.

Although we perform this ordering algorithm in our real data example in the next Section, we emphasize that this model-fitting exercise can only give the model that best describes the data generation mechanism, but it may fail in providing the best model that optimally predicts covariance matrices. This important fact had been also noted by Aguilar and West (2000), but they did not deal with the problem of ordering the stocks in a portfolio.

4 Exchange rates data

The predictive performance of multivariate time-varying volatility models can be judged based on their forecasting powers. Here we concentrate on comparing the forecast performance of our procedure with six of the existing multivariate volatility models. We also study the impact of various ordering of the stocks in a portfolio on the forecast performance of the various models. Some additional experiments that focus on the ability of our Cholesky-GARCH models and the competing models to construct portfolios or calculate Value at Risk (VaR) can be found in Dellaportas and Pourahmadi (2004).

4.1 The data and the competing models

We obtained (source: DATASTREAM) 7 daily exchange rates of the US dollar against UK pound, EURO, Swedish krona, Australian dollar, Canadian dollar, Swiss franc and Japanese YEN, recorded from 2/1/1999 up to 28/10/2003. We also obtained (source: REUTERS) two-minute intraday data from the same exchange rates for the following 3 days 29–31/10/2003. From the 7 series, we created all ${}^7C_5 = 21$ portfolios of 5 exchange rates, and used them as replications in our computational experiment. Finally, for the purposes of comparison we used the following six widely used multivariate volatility models:

(i) The multivariate diagonal-vec model of order (1, 1); see Bollerslev et al. (1994).
(ii) The matrix-diagonal model of order (1, 1); see Bollerslev et al. (1994).
(iii) An exponentially weighted moving average model of the form

$$\Sigma_t = \alpha(\varepsilon_{t-1}\varepsilon'_{t-1}) + (1-\alpha)\Sigma_{t-1},$$

where $0 \le \alpha \le 1$ and $\{\varepsilon_t\}$ is the vector of shocks. Note that RiskMetrics uses this model where the smoothing parameter α is not estimated, but set to 0.06.
(iv) The constant-correlation of order (1, 1); see Bollerslev (1990).
(v) The orthogonal GARCH model of order (1, 1); see Alexander (2001).
(vi) The dynamic conditional correlation (DCC) model; see Engle (2002).

For each of the 21 replications of the exchange rates, and for all 5! = 120 possible orderings within each replication, we predicted the conditional covariance matrices for the next three days of 29–31/10/2003 using the new Cholesky-GARCH models and the six popular multivariate models (i)–(vi) mentioned above.

4.2 Measures of forecast performance and the empirical results

Although the true realized covariance matrix is unavailable, recent developments in the analysis of realized covariation (Andersen et al. 1999; Barndorff-Nielsen and Shephard 2004) allow us to replace it by a reliable proxy, the realized covariation matrix. The realized covariation matrix with elements σ_{ij}, is calculated for each of the three days as the cumulative cross-products of intraday returns over each day. Following Ledoit et al. (2003), for the corresponding forecasts σ^*_{ij} derived from the daily data series the following two measures of forecast performance will be used here:

Mean absolute deviation

$$MAD = N^{-2}\sum_{i,j} E|\sigma^*_{ij} - \sigma_{ij}|.$$

Root mean square error

$$RMSE = \left[N^{-2}\sum_{i,j} E(\sigma^*_{ij} - \sigma_{ij})^2\right]^{1/2}.$$

Table 1 presents MAD and RMSE averaged over all days and all 21 datasets. We also report results for the orderings that give the lowest and highest RMSE and MAD. Note that these orderings can only be chosen after the calculation of RMSE and MAD for all 120 orderings so they have no practical value, but they provide an indication of the range of achievable predictive ability of all Cholesky-GARCH models. It is clear that if we were able to find the best ordering in terms of forecasting power we would have outperformed any other model. Moreover, if we were randomly chosen the worst possible ordering, we would have not have achieved the worst performance across all models.

Our best-fit ordering algorithm of Sect. 3.2 does not perform, in terms of forecasting power, as well as the simple ordering based on the unconditional variances of the stocks in a portfolio. This is not surprising since it is usually the case that the model that fits best does not necessarily predicts best. Indeed, the constant conditional correlation model which makes the simplest assumption of equal correlation beats all models with respect to both MAD and RMSE. Our Cholesky-GARCH simple ordering based on unconditional variance beats all other models except the matrix-diagonal model which is better with respect to

Table 1 Average root mean square error (RMSE) and mean absolute deviation (MAD) across 21 datasets and 3 days. All values have been multiplied by 10^4

MODELS	RMSE	MAD
Cholesky-GARCH, simple ordering	0.2364	0.1829
Cholesky-GARCH, AIC-based ordering	0.2431	0.1845
Cholesky-GARCH, best ordering	0.2123	0.1637
Cholesky-GARCH, worst ordering	0.2664	0.2044
Dynamic conditional correlation (DCC)	0.3301	0.2135
Diagonal-vec	0.2380	0.1833
Matrix diagonal	0.2354	0.1835
Exponentially weighted moving average	0.2473	0.2008
Constant conditional correlation (CCC)	0.2257	0.1680
Orthogonal GARCH	0.2739	0.1988

RMSE but worst with respect to MAD. For practical purposes, the Cholesky-GARCH model based on ordering the unconditional variances is simple to run, quick when N is large, and performs well compared to competing models.

We find these results to be very promising, especially because our simple ordering which requires no extra computational cost performs well in terms of forecasting. We conjecture that in portfolios containing financial products more diverse than the exchange rates the constant correlation model will fail to capture the empirical dynamics of all series. Moreover, we emphasize the extraordinary ease in applying our method in high dimensions compared to the exponential moving average and the orthogonal GARCH models. For example, in Dellaportas and Pourahmadi (2004) we apply our algorithm in a 100-dimensional problem by just running a series of 100 univariate regression GARCH(1, 1) models.

Since one of the strengths of the Cholesky-GARCH models is their ease of computation, we report CPU results based on 100 stocks and 250 days from a dataset used in Dellaportas and Pourahmadi (2004). A simple stocks ordering based on unconditional variances and 100 calls to the MATLAB routine garchfit took 7.3 minutes in a 2 GHz Pentium laptop. For a smaller portfolio of 10 stocks the corresponding time was less than 9 seconds.

5 Conclusion and future work

Our preliminary empirical work shows the great promise of the proposed Cholesky-GARCH models in providing parsimonious models for conditional covariances while guaranteeing the positive-definiteness of their estimators. Detailed empirical results from an experimental study based on exchange rates indicates that choosing a particular ordering of the stocks in a portfolio does not alter by much the forecasting power of our proposed models, a similar conclusion in a slightly different context is arrived at by Chang and Tsay (2010). Nevertheless the problem of developing a computationally efficient algorithm for "optimal" ordering of the stocks remains open and will be studied in our subsequent work. Extension of Chang and Tsay's (2010) recent sequential algorithm in the context of penalized likelihood estimation involving the Cholesky decomposition to our setup will be studied where it can speed up the computation as more data become available over time.

Acknowledgements Research of the second author was supported in part by NSF grants DMS-0505696 and DMS-0906252. Comments from a referee have greatly helped to improve the presentation and focus of the paper.

References

Aguilar, O., West, M.: Bayesian dynamic factor models and portfolio allocation. J. Bus. Econ. Stat. **18**, 338–357 (2000)

Alexander, C.: Market Models: A Guide to Financial Data Analysis. Wiley, New York (2001)

Andersen, T.G., Bollerslev, Lange, S.: Forecasting financial market volatility: sample frequency vis-à-vis forecast horizon. J. Empir. Finance **6**, 457–477 (1999)

Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics. Econometrica **72**, 885–925 (2004)

Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. J. Econom. **31**, 307–327 (1986)

Bollerslev, T.: Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. Rev. Econ. Stat. **72**, 498–505 (1990)

Bollerslev, T., Engle, R., Nelson, D.: ARCH models. In: Engle, R., McFadden, D. (eds.) Handbook of Econometrics, pp. 2959–3038 (1994)

Chang, C., Tsay, R.S.: Estimation of covariance matrix via sparse Cholesky factor with Lasso. J. Stat. Plan. Inference **40**, 3858–3873 (2010)

Chou, R.Y., Wu, C.C., Liu, N.: Forecasting time-varying covariance with a range-based dynamic conditional correlation model. Rev. Quant. Finance Account. **33**, 327–345 (2009)

Dellaportas, P., Pourahmadi, M.: Large time-varying covariance matrices with applications to finance. Technical report, Athens University of Economics, Department of Statistics (2004)

Diebold, F.X., Nerlove, M.: The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. J. Appl. Econom. **4**, 1–21 (1989)

Engle, R.F.: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica **50**, 987–1008 (1982)

Engle, R.F.: Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. **20**, 339–350 (2002)

Engle, R.F., Kroner, K.F.: Multivariate simultaneous generalized ARCH. Econom. Theory **11**, 122–150 (1995)

Geweke, J.F., Zhou, G.: Measuring the pricing error of the arbitrage pricing theory. Rev. Financ. Stud. **9**, 557–587 (1996)

Ledoit, O., Santa-Clara, P., Wolf, M.: Flexible multivariate GARCH modeling with an application to international stock markets. Rev. Econ. Stat. **85**, 735–747 (2003)

Pourahmadi, M.: Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation. Biometrika **86**, 677–690 (1999)

Pourahmadi, M.: Maximum likelihood estimation of generalized linear models for multivariate normal covariance matrix. Biometrika **87**, 425–435 (2000)

Smith, M., Kohn, R.: Parsimonious covariance matrix estimation for longitudinal data. J. Am. Stat. Assoc. **97**, 1141–1153 (2002)

Stelzer, R.J.: Multivariate COGARCH (1, 1) processes. Bernoulli **16**, 80–115 (2010)

Tsay, R.: Analysis of Financial Time Series, 3rd edn. Wiley, New York (2005)

Tse, Y.K., Tsui, A.K.: A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J. Bus. Econ. Stat. **20**, 351–362 (2002)

Vrontos, I.D., Dellaportas, P., Politis, D.N.: A full-factor multivariate GARCH model. Econom. J. **6**, 312–334 (2003)

Webb, E.L., Forster, J.J.: Bayesian model determination for multivariate ordinal and binary data. Comput. Stat. Data Anal. **52**, 2632–2649 (2008)