INFORMATION
Volume 29, Number 1, pp.51-54

ISSN 1343-4500 eISSN 1344-8994

High-Dimensional Volatility Modeling: A Mixed-Factor Approach

Runyu Dai* and Yasumasa Matsuda

Graduate School of Economics and Management, Tohoku University, Sendai, Japan

*Corresponding author: dai.runyu.a3@tohoku.ac.jp

Abstract

Estimating high-dimensional volatility matrices is an important challenge in financial econometrics, particularly for portfolio allocation and systemic risk monitoring. This paper studies the empirical implementation of a volatility estimator that integrates observable economic factors with sparsity-induced weak latent components. We outline a sequential estimation procedure and apply the method to S&P 500 daily returns. The results demonstrate that incorporating mixed-factor structures into dynamic volatility modeling leads to consistently strong empirical performance.

Keywords: Volatility matrix; factor model; high-dimensional data; portfolio allocation.

1. Introduction

Modeling large volatility matrices is central to financial econometrics. Classical MGARCH models, extending ARCH (Engle, 1982) and GARCH (Bollerslev, 1986), quickly become infeasible in high dimensions due to the explosion of parameters. To improve tractability, Bollerslev (1990) introduced CCC-GARCH, and Engle (2002) proposed DCC-GARCH to allow time-varying correlations, with high-dimensional nonlinear extensions in Engle et al. (2019).

A complementary line of research employs low-rank factor structures. The PC-GARCH model of Ding (1994) and its O-GARCH and GO-GARCH variants (Alexander, 2000;Van der Weide, 2002) capture co-volatility through principal components. Subsequent developments include Full-factor GARCH (Vrontos et al., 2003), asymptotic analysis (Hafner and Preminger, 2009), and DCC-Factor-GARCH (Zhang and Chan, 2009). More recent studies incorporate observable-factor GARCH with sparse residual dependence (Li et al., 2022a,b), while asset-pricing evidence highlights omitted-variable concerns when relying solely on observable factors (Feng et al., 2020; Giglio and Xiu, 2021). Residual-based latent-factor extraction (Shi et al., 2022) mitigates this to some degree, though empirical evidence shows that many residual factors are weak and diverge slowly with cross-sectional size (Dai et al., 2024).

Motivated by these findings, we propose a DCC-embedded mixed-factor volatility model that integrates observable factors, sparsity-induced weak latent factors (Uematsu and Yamagata, 2022), and a dynamically evolving idiosyncratic component. By embedding DCC into all components and allowing for weak latent structures, the framework addresses information loss from unobservable risks and the limitations of static or purely observable-factor specifications.

2. The Estimation Framework

We consider an N-dimensional vector of asset returns generated by the mixed-factor structure

$$y_t = Ax_t + u_t, \quad u_t = Bf_t + e_t,$$

where x_t denotes observable economic factors, f_t means latent statistical factors, and e_t denotes idiosyncratic errors. A and B represent the factor loadings. The key innovation lies in explicitly modeling the volatilities of all components through a sequential and numerically stable estimation procedure.

2.1 Step 1: Filtering Observable Components

We first isolate variation driven by known factors. Standard asset-pricing factors (e.g., Fama–French) are modeled through a two-stage procedure: univariate GARCH for conditional variances and DCC-GARCH for dynamic correlations. Regressing asset returns on these observable factors yields loadings A and residuals

$$\widehat{u_t} = y_t - \hat{A}x_t,$$

which contain remaining common factors and idiosyncratic noise. This step produces the observable-factor volatility component $\hat{A}\widehat{\Sigma_x}(t)\hat{A}'$.

2.2 Step 2: Extraction of Latent Factors

We extract remaining common factors from u_t using the Sparse Orthogonal Factor Regression (SOFAR) method of Uematsu et al. (2019), which is well suited to identifying weak latent factors that influence only subsets of assets. The extracted latent factors f_t are then modeled by a secondary DCC specification, yielding the latent-factor volatility component $\hat{B}\widehat{\Sigma_f}(t)\hat{B}'$.

2.3 Step 3: Capturing Idiosyncratic Errors

The idiosyncratic component is obtained after removing observable and latent factors. The third DCC model is then fitted to $\widehat{e_t}$, with the unconditional covariance estimated via a sparse technique, yielding the idiosyncratic volatility matrix $\widehat{\Sigma_e}(t)$. The final estimated volatility matrix of y_t is

$$\widehat{\Sigma_y}(t) = \hat{A}\widehat{\Sigma_x}(t)\hat{A}' + \hat{B}\widehat{\Sigma_f}(t)\hat{B}' + \widehat{\Sigma_e}(t).$$

3. Empirical Studies

We use daily returns of S&P 500 constituents from April 2002 to April 2022. After data alignment and cleaning, we conduct a real-time portfolio allocation based on a monthly rolling-window procedure: model parameters are estimated using the most recent 252 trading days (about one year trading days), the volatility matrix $\widehat{\Sigma_y}(t+1)$ is then forecasted, and the Minimum Variance Portfolio (MVP) is constructed under the investment constraint $\sum_i w_i = 1$ using $\widehat{\Sigma_y}(t+1)$, where w_i is the unknown parameter representing the portfolio share of asset i. Across the sample, the portfolio constructed by our proposed method consistently achieves lower risk compared with several benchmark alternatives.

4. Conclusion Remarks

We propose a Factor-GARCH framework embedding DCC for both observable and latent factors, coupled with a sparse, dynamically evolving idiosyncratic component, to model the volatility matrices of high dimensional data. The approach mitigates omitted variable and singularity issues and performs well in simulations and empirical MVP applications. Future work may include: (i) richer GARCH formulations (T-/GJR-/E-GARCH) and heavy tailed QMLE (Fan et al., 2014); (ii) time-varying loadings for observable and latent factors; and (iii) applications to high-dimensional connectedness and risk forecasting.

5. Acknowledgements

Runyu Dai acknowledges support from the Japan Society for the Promotion of Science (Grant 24K22603). Yasumasa Matsuda acknowledges support from the Japan Society for the Promotion of Science (Grant 23K21647).

References

[1] Alexander, C. (2000). A primer on the orthogonal garch model. *Manuscript ISMA Centre*, University of Reading, UK, 2.

[2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327.

[3] Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *The Review of Economics and Statistics*, pages 498–505.

[4] Dai, R., Uematsu, Y., and Matsuda, Y. (2024). Estimation of large covariance matrices with mixed factor structures. *The Econometrics Journal*, 27(1):62–83.

[5] Ding, Z. (1994). *Time series analysis of speculative returns.* University of California, San Diego, Department of Economics.

[6] Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

[7] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987– 1007.

[8] Engle, R. F., Ledoit, O., and Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, 37(2):363–375.

[9] Fan, J., Qi, L., and Xiu, D. (2014). Quasi-maximum likelihood estimation of garch models with heavy-tailed likelihoods. *Journal of Business & Economic Statistics*, 32(2):178–191.

[10] Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370.

[11] Giglio, S. and Xiu, D. (2021). Asset pricing with omitted factors. *Journal of Political Economy*, 129(7):1947–1990.

[12] Hafner, C. M. and Preminger, A. (2009). Asymptotic theory for a factor garch model. *Econometric Theory*, 25(2):336–363.

[13] Li, X., Li, Y., Pan, J., and Zhang, X. (2022a). A factor-garch model for high dimensional volatilities. *Acta Mathematicae Applicatae Sinica, English Series*, 38(3):635–663.

[14] Li, X., Zhang, X., and Li, Y. (2022b). High-dimensional conditional covariance matrices estimation using a factor-garch model. *Symmetry*, 14(1):158.

[15] Shi, F., Shu, L., and Gu, X. (2022). An enhanced factor model for portfolio selection in high dimensions. *Journal of Financial Econometrics*.

[16] Uematsu, Y., Fan, Y., Chen, K., Lv, J., and Lin, W. (2019). SOFAR: large-scale association network learning. *IEEE transactions on information theory*, 65(8):4924–4939.

[17] Uematsu, Y. and Yamagata, T. (2022). Estimation of sparsity-induced weak factor models. *Journal of Business & Economic Statistics*, 41(1):213–227.

[18] Van der Weide, R. (2002). Go-garch: A multivariate generalized orthogonal garch model. *Journal of Applied Econometrics*, 17(5):549–564.

[19] Vrontos, I. D., Dellaportas, P., and Politis, D. N. (2003). A full-factor multivariate garch model. *The Econometrics Journal*, 6(2):312–334.

[20] Zhang, K. and Chan, L. (2009). Efficient factor garch models and factor-dcc models. *Quantitative Finance*, 9(1):71–91.

Transparency Note: This paper proposal is submitted primarily for the purpose of academic exchange in *The 11th International Conference on Information*. We sincerely seek to gather interdisciplinary feedback, particularly from information science, engineering, and computational perspectives, to further refine our method. This proposal outlines the computational framework and practical application logic, while the rigorous theoretical foundations, model descriptions, and comprehensive numerical analyses are detailed in a separate full paper currently under review at a specialized finance journal.