10-2020

Forecasting large covariance matrix with high-frequency data: A factor approach for the correlation matrix

Yingjie DONG
University of International Business and Economics

Yiu Kuen TSE
Singapore Management University, yktse@smu.edu.sg

Forecasting large covariance matrix with high-frequency data using factor approach for the correlation matrix

Yingjie Dong[a], Yiu-Kuen Tse[b*]

[a] Business School, University of International Business and Economics, Beijing, China

[b] School of Economics, Singapore Management University, Singapore 178903, Singapore
* Corresponding author. E-mail address:yktse@smu.edu.sg

Published in Economics Letters, 2020, 195, 109465. DOI: 10.1016/j.econlet.2020.109465

Abstract: We apply the factor approach to the correlation matrix to forecast large covariance matrix of asset returns using high-frequency data, using the principal component method to model the underlying latent factors of the correlation matrix. The realized variances are separately forecasted using the Heterogeneous Autoregressive model. The forecasted variances and correlations are then combined to forecast large covariance matrix. Our proposed method is found to perform better in reporting smaller forecast errors than some selected competitors. Empirical application to a portfolio of 100 NYSE and NASDAQ stocks shows that our method provides lower out-of-sample realized variance in selecting global minimum variance portfolio.

Keywords: Dimension reduction, Eigenanalysis, Factor model, High-frequency data, Large correlation matrix, Nonlinear shrinkage

1. Introduction

Modeling time varying covariance matrix of asset returns plays a crucial role in modern financial risk management and asset allocation. Multivariate GARCH (MGARCH) models are useful tools to deal with this problem. These models, however, are usually applied to low-dimension portfolios and problems arise when the number of assets is large. Aielli (2013) proposes a consistent corrected Dynamic Conditional Correlation (cDCC) method for high dimension portfolios. Pakel et al. (2020) propose a composite quasi-likelihood estimate to tackle the computational issue.

MGARCH models are typically applied to daily, weekly or monthly data. With the availability of high-frequency intraday data, researchers can model and forecast the variance and covariance of asset returns using tick-by-tick transaction or quotation data. Johnstone (2001), among others, point out that as the size of the portfolio covariance matrix goes to infinity, the naive sample covariance/correlation estimator is inconsistent and the eigenvalues and eigenvectors of the estimated covariance matrix may deviate substantially from the true values. Fan et al. (2016) review some large covariance matrix estimation methods. Compared against the MGARCH family of models, these methods can deal with the curse of dimensionality quite successfully. Although these methods can provide good estimates of covariance/correlation matrices, how to model the dynamics of the large covariance matrices and produce reliable out-of-sample forecasts is still a very challenging task.

The focus of this paper is to examine the performance of a dynamic method to forecast large covariance matrix with high-frequency data using a factor approach for the correlation matrix. First, we modify the latent factor model of Tao et al. (2011) and apply it to correlation matrix. The dynamic high-dimension correlation matrix is assumed to be driven by a low-dimension latent process. We model the dynamic structure of the latent factors using a vector autoregressive (VAR) model. This captures the short-memory dynamics of the latent factors. Forecasts for these factors are then used to generate forecasts for the full correlation matrix. Second, we forecast the volatility of individual asset returns using the Heterogeneous Autoregressive (HAR) model of Corsi (2009). This model approximates well the long-memory properties of realized variances. Finally, we combine the realized volatility forecasts with the large correlation matrix forecasts to obtain large covariance matrix forecasts.

Our method differs from current large covariance forecast methods in the literature in two aspects. First, we model the correlation matrix process and the univariate volatility processes separately. We assume a short-memory structure for the vectorized latent factors of the correlation matrix and a long-memory structure for the volatility processes. In contrast, existing papers often assume short-memory dynamic structures for the covariance (see Tao et al. (2011) and Callot et al. (2016)), while we allow for a more flexible set-up. For each univariate volatility process, we can adopt models with persistent dynamics and/or asymmetric effect in positive versus negative returns. For the correlation matrix process, we adopt factor models as currently used for large covariance matrices. The idea of separately modeling the correlation matrix process and the univariate volatility processes dates back to the Dynamic Conditional Correlation (DCC) model of Engle (2002) and the Time-Varying Correlation model of Tse and Tsui (2002). In this paper, instead of using the MGARCH models, we explore whether we can obtain better large covariance matrix estimates with high-frequency data by using the factor approach.

Second, to obtain raw large covariance matrix using high-frequency data for the eigen-analysis, we calculate the raw large correlation matrix by regulating the eigenvalues of the matrix using the nonlinear shrinkage method of Ledoit and Wolf (2012).[1] The nonlinear shrinkage method is a successful approach even in the absence of any knowledge about the structure of the latent covariance matrix. In contrast, the threshold averaging realized volatility matrix (TARVM) method of Tao et al. (2011) imposes sparsity assumption on elements of the covariance matrix, which may be inappropriate as stock returns are largely correlated due to the presence of market risks.

We perform an empirical comparison of our method against the following methods: the factor covariance matrix method of Tao et al. (2011), the cDCC method of Aielli (2013), and the DCC-shrinkage method of Engle et al. (2017). Our method has better performance in terms of out-of-sample portfolio allocation for constructing both the global minimum variance (GMV) portfolio and the Markowitz portfolio with momentum signal using a portfolio of NYSE and NASDAQ stocks. Also, it performs the best in reporting smaller forecast errors in our Monte Carlo simulation study.

The plan of the rest of this paper is as follows. In Section 2, we describe the construction of our latent factor approach of the large correlation matrix. Section 3 describes an empirical investigation of the performance of different large covariance matrix forecasts in terms of out-of-sample asset allocation. Some concluding remarks are given in Section 4. A summary of the implementation procedure of our method and some Monte Carlo results for the performance of our method are separately reported in the Online Appendix.

2. Forecasting large covariance matrix

2.1. Model set-up

Let $\mathbf{X}(t) = (X_1(t), \ldots, X_d(t))'$ be an Itô process given by

$$d\mathbf{X}(t) = \boldsymbol{\mu}(t)dt + \boldsymbol{\sigma}'(t)d\mathbf{B}_t, \qquad t \in (0, T], \tag{1}$$

where the stochastic processes $\mathbf{B}_t$, $\boldsymbol{\mu}(t)$, and $\boldsymbol{\sigma}(t)$ are defined on the filtered probability space denoted by $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \in [0, T]\}, P)$. $\mathbf{B}_t$ is a d-dimensional standard Brownian motion with respect to $\mathcal{F}_t$, $\boldsymbol{\mu}(t)$ is a d-dimensional drift vector, $\boldsymbol{\sigma}(t)$ is a $d \times d$ matrix, and $\boldsymbol{\mu}(t)$ and $\boldsymbol{\sigma}(t)$ are assumed to be predictable processes with respect to the filtration $\mathcal{F}_t$. We assume d to be large, typically in the hundreds.

The *integrated covariance matrix* of $\mathbf{X}_t$ for the tth period (from time $t-1$ to time t) is defined as the $d \times d$ matrix

$$\boldsymbol{\Sigma}_t = \int_{t-1}^{t} \boldsymbol{\sigma}(s)'\boldsymbol{\sigma}(s)\,ds, \qquad t = 1, \ldots, T, \tag{2}$$

and the *integrated correlation matrix* $\boldsymbol{\Gamma}_t$ for the tth period is the correlation matrix transformation (CMT) of $\boldsymbol{\Sigma}_t$ as follows

$$\boldsymbol{\Gamma}_t = \widetilde{\boldsymbol{\Sigma}}_t^{-\frac{1}{2}}\boldsymbol{\Sigma}_t\widetilde{\boldsymbol{\Sigma}}_t^{-\frac{1}{2}}, \qquad t = 1, \ldots, T, \tag{3}$$

where $\widetilde{\boldsymbol{\Sigma}}_t$ is obtained from $\boldsymbol{\Sigma}_t$ by replacing the off-diagonal elements of $\boldsymbol{\Sigma}_t$ by zero.

Empirically, we observe the trading price that is the contaminated price of the efficient price $\mathbf{X}(t)$ due to market microstructure noise. Our objective is to forecast the integrated covariance matrix of $\mathbf{X}(t)$ using high-frequency data.

2.2. Estimation of latent factors of the correlation matrix

We adopt the matrix factor model of Tao et al. (2011) for high-frequency covariance matrix estimation and apply it to large correlation matrix. Specifically, we assume

$$\boldsymbol{\Gamma}_t = \mathbf{A}\boldsymbol{\Gamma}_t^f\mathbf{A}' + \boldsymbol{\Gamma}_0, \tag{4}$$

where $\boldsymbol{\Gamma}_t^f$, $t = 1, \ldots, T$, are $r \times r$ ($r \ll d$) positive definite matrices treated as dynamical latent factors of the correlation process, $\mathbf{A}$ is a $d \times r$ factor loading matrix with $\mathbf{A}'\mathbf{A} = \mathbf{I}_r$, and $\boldsymbol{\Gamma}_0$ is a $d \times d$ positive definite time invariant matrix. The dynamical structure of the $d \times d$ correlation matrices $\boldsymbol{\Gamma}_t$ is driven by a lower-dimensional $r \times r$ latent process $\boldsymbol{\Gamma}_t^f$, while $\boldsymbol{\Gamma}_0$ represents the static part of $\boldsymbol{\Gamma}_t$. Thus, to capture the dynamics of $\boldsymbol{\Gamma}_t$, we control the parametric dimension by modeling $\boldsymbol{\Gamma}_t^f$.[2]

We first estimate the covariance matrix $\boldsymbol{\Sigma}_t$, from which the correlation matrix $\boldsymbol{\Gamma}_t$ can be calculated using the CMT. To estimate $\boldsymbol{\Sigma}_t$, we adopt the nonlinear shrinkage method proposed by Ledoit and Wolf (2012), which rectifies the overfitting problem of the sample covariance matrix by recovering the population eigenvalues from the sample eigenvalues using a nonlinear shrinkage formula.[3] The corresponding estimate of the correlation matrix through the CMT will then be denoted by $\widehat{\boldsymbol{\Gamma}}_t$.

To calculate the time invariant matrices $\mathbf{A}$ and $\boldsymbol{\Gamma}_0$ in (4), we use the method of Tao et al. (2011) for covariance matrices and apply it to our model. Thus, we define

$$\widehat{\mathbf{S}} = \frac{1}{T}\sum_{t=1}^{T}(\widehat{\boldsymbol{\Gamma}}_t - \widehat{\boldsymbol{\Gamma}})^2, \tag{5}$$

where $\widehat{\boldsymbol{\Gamma}} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{\Gamma}}_t$. We use the r orthonormal eigenvectors corresponding to the r largest eigenvalues of $\widehat{\mathbf{S}}$ as the columns of the factor loading matrix $\mathbf{A}$, and denote this estimate by $\widehat{\mathbf{A}}$. The estimated factor matrix is then computed as

$$\widehat{\boldsymbol{\Gamma}}_t^f = \widehat{\mathbf{A}}'\widehat{\boldsymbol{\Gamma}}_t\widehat{\mathbf{A}}, \tag{6}$$

and the estimate of $\boldsymbol{\Gamma}_0$ is

$$\widehat{\boldsymbol{\Gamma}}_0 = \widehat{\boldsymbol{\Gamma}} - \widehat{\mathbf{A}}\widehat{\mathbf{A}}'\widehat{\boldsymbol{\Gamma}}\widehat{\mathbf{A}}\widehat{\mathbf{A}}'. \tag{7}$$

2.3. Forecasting factor matrix and large correlation matrix

We use the Vector Autoregressive (VAR) Model to capture the short-run dynamics of the latent factors. For a $r \times r$ matrix $\boldsymbol{\Gamma}$, let vech$(\boldsymbol{\Gamma})$ denote the vector obtained by stacking together all

[1] We thank Ledoit and Wolf for providing the codes (www.econ.uzh.ch/en/people/faculty/wolf/publications.html).

[2] Note that $\boldsymbol{\Gamma}_t^f$ need not be a well-defined correlation matrix. We assume, however, this is a latent factor matrix generating the large correlation matrix $\boldsymbol{\Gamma}_t$.

[3] Engle, Ledoit and Wolf (2012) show that the nonlinear shrinkage method has superior out-of-sample properties when applied to the Dynamic Conditional Correlation (DCC) Model. An alternative method to calculate $\boldsymbol{\Sigma}_t$ is the TARVM estimator proposed by Tao et al. (2011). This method will also be considered in our empirical application.

elements on and below the diagonal of $\boldsymbol{\Gamma}$. The VAR model for $\boldsymbol{\Gamma}_t^f$ is given by

$$\mathrm{vech}(\boldsymbol{\Gamma}_t^f) = \boldsymbol{\alpha}_0 + \sum_{j=1}^{q} \boldsymbol{\alpha}_j \mathrm{vech}(\boldsymbol{\Gamma}_{t-j}^f) + \boldsymbol{e}_t, \tag{8}$$

where $\boldsymbol{\alpha}_0$ is a $\tilde{r} \times 1$ vector with $\tilde{r} = r(r+1)/2$, and $\boldsymbol{\alpha}_j$, for $j = 1, \ldots, q$, are $\tilde{r} \times \tilde{r}$ square matrices. $\boldsymbol{e}_t$ is a $\tilde{r} \times 1$ vector white noise process with zero mean and finite fourth moments. Empirically, we fit equation (8) using $\widehat{\boldsymbol{\Gamma}}_k^f$ as observed values of $\boldsymbol{\Gamma}_k^f$, for $k = 1, \ldots, t-1$, to obtain the estimated coefficients $\widehat{\boldsymbol{\alpha}}_j$ for $j = 0, 1, \ldots, q$, and then $\widehat{\boldsymbol{\alpha}}_j$ are used to compute the out-of-sample forecasted latent factors matrix for the tth period.

We denote the forecast of $\boldsymbol{\Gamma}_t^f$ conditional upon information up to time $t-1$ using the estimated VAR model by $\check{\boldsymbol{\Gamma}}_t^f$. Then, the forecast of the $d \times d$ large correlation matrix $\check{\boldsymbol{\Gamma}}_t$ is computed as

$$\check{\boldsymbol{\Gamma}}_t = \widehat{\mathbf{A}} \check{\boldsymbol{\Gamma}}_t^f \widehat{\mathbf{A}}' + \widehat{\boldsymbol{\Gamma}}_0. \tag{9}$$

Note that $\check{\boldsymbol{\Gamma}}_t$ may not be a well defined correlation matrix (positive definite matrix with unit diagonal elements).[4] To resolve this problem, we apply the CMT on $\check{\boldsymbol{\Gamma}}_t$ to obtain $\check{\boldsymbol{\Gamma}}_t^*$ as the forecasted correlation matrix. On the other hand, if $\check{\boldsymbol{\Gamma}}_t$ is not positive definite, we project the matrix onto the space of positive definite matrices using methods in Fan et al. (2016).

2.4. Forecasting realized variance and large covariance matrix

We further forecast the variance of individual assets separately using the Heterogeneous Autoregressive (HAR) model of realized volatility proposed by Corsi (2009). We estimate the HAR equation as follows

$$RV_{i,t} = \omega_i + \alpha_i RV_{i,t-1} + \beta_i RV_{i,t-1}^{w} + \gamma_i RV_{i,t-1}^{m}, \qquad i = 1, \ldots, d, \tag{10}$$

where $RV_{i,t}$ is the calculated realized variance of asset i in period t, $RV_{i,t-1}^w = \frac{1}{5}\sum_{s=1}^{5} RV_{i,t-s}$, $RV_{i,t-1}^m = \frac{1}{22}\sum_{s=1}^{22} RV_{i,t-s}$. To compute $RV_{i,t}$, we use the subsampling method of Zhang et al. (2005) at 3-min intervals. The estimated models in Eq. (10) are used to forecast the realized variances. These forecasts are then collected to form the matrix $\check{\mathbf{D}}_t$, which is a $d \times d$ diagonal matrix with its ith diagonal element being the forecasted realized variance of asset i.

Finally, we compute the forecasted large covariance matrix as

$$\check{\boldsymbol{\Sigma}}_t = \check{\mathbf{D}}_t^{\frac{1}{2}} \check{\boldsymbol{\Gamma}}_t^* \check{\mathbf{D}}_t^{\frac{1}{2}}. \tag{11}$$

This forecast procedure will be called M1. We conduct a Monte Carlo study to investigate the performance of this method against some other alternatives in the literature. Our method performs the best in reporting smaller Frobenius norm errors and spectral norm errors. These results and a summary of the method M1 can be found in the Online Appendix.

3. Empirical comparison of portfolio selection

We compare the performance of various forecasts of the variance matrix based on out-of-sample asset allocation. We select 100 largest market capitalization stocks (as of 2015) that are listed in NYSE or NASDAQ, with at least 200 trading days in any calendar year between 2004 and 2016 (3171 trading days). Tick-by-tick millisecond transaction data are compiled and downloaded from the WRDS Daily TAQ (DTAQ) database.

We calculate the forecasted covariance matrices using our proposed method M1. For comparison, we also vary M1 by modeling the covariance matrix process instead of the correlation matrix process, and call this method M2. Similar to M1, M2 uses a VAR model to capture the dynamics of the latent covariances (not correlations). It differs from M1 in that the HAR forecast for the realized variance of individual assets is not performed and the dynamic covariances are directly modeled. For further comparison, we also include the TARVM method of Tao et al. (2011) and denote this estimate as M3, which is the same as M2 except for the method in estimating the raw large covariance matrix. Finally, we include the DCC model with nonlinear shrinkage of Engle et al. (2017) and the cDCC model of Aielli (2013), which are denoted as DCC-shrinkage and cDCC, separately.

3.1. Portfolio selection

We compare the performance of various covariance forecast methods based on the selection for the global minimum variance (GMV) portfolio and the Markowitz portfolio with momentum signal.

For the GMV portfolio, we choose the portfolio weights to minimize the portfolio variance. We also investigate the problem of choosing portfolio weights such that the portfolio variance is minimized given a specific expected rate of return r_p. We follow Engle et al. (2017) and treat the momentum factor as the required portfolio return r_p. We construct portfolios based on the calculated out-of-sample optimal weights, and then evaluate different methods by comparing the corresponding portfolio's realized variance and information ratio. The latter is defined as the ex-post realized portfolio return divided by the realized portfolio volatility and is particularly relevant as a performance measure for the Markowitz portfolio with momentum signal.[5]

3.2. Out-of-sample comparison of portfolio selection

We compare the performance of different covariance matrix estimates in terms of their ability to select portfolios with the lowest variance for the GMV portfolio and higher information ratio for the Markowitz portfolio with momentum signal. We calculate the optimal portfolio weights using the out-of-sample forecasted covariance matrices, and then construct optimal portfolios of the next period based on the calculated optimal weights. To avoid an excessive amount of turnover and thus transaction costs, we update all portfolios at biweekly frequency, that is, every 10 consecutive trading days.[6] We calculate all large covariance/correlation matrix at 15-min frequency and calculate volatility at 3-min frequency.[7] To calculate the volatility of the constructed portfolio, we use the subsampling method of Zhang et al. (2005) with portfolio returns at 15-min frequency.

To fit the factor correlation/covariance matrices we let the number of factors r be 3, 4 and 5. The number of coefficients of the VAR model increases quickly as the lag parameter q or the number of factors r increases. Thus, we fit the diagonal-VAR(q) models for the vectorized factor matrices, with $q = 1$. For comparison, We fit the cDCC model of Aielli (2013) and the shrinkage

[4] Note that this problem exists for all covariance/correlation matrix forecast methods.

[5] Note that focusing on the out-of-sample standard deviation is now inappropriate due to estimation error in the momentum signal.

[6] As there are 3171 trading days in our sample, we have a total of $T = 317$ periods. We start to calculate the out-of-sample portfolio weights at $t = 251$. To calculate the forecasted biweekly variance, we use a model similar to HAR and select daily RV, weekly RV and monthly RV as explanatory variables.

[7] See the Online Appendix for further discussions on this setting.

Table 1
Comparison of constructed portfolios.

Method	r	GMV		Markowitz portfolio with a signal	
		Volatility (%)	Information ratio	Volatility (%)	Information ratio
M1	3	9.7373	1.4246	9.8131	1.5702
	4	9.6986	1.5344	9.7853	1.6455
	5	9.6904	1.5157	9.7846	1.6292
M2	3	9.8345	0.9158	9.9533	1.0761
	4	10.1826	0.7693	10.2871	0.9157
	5	10.1832	0.7730	10.2874	0.9195
M3	3	10.9701	1.1198	11.0449	1.1279
	4	10.9777	1.1083	11.0519	1.1179
	5	10.9741	1.1298	11.0484	1.1379
DCC-Shrinkage		11.0258	1.1662	11.0845	1.1918
cDCC		20.0295	0.8213	27.1846	0.8643

Notes: The figures are the mean realized daily volatility (annualized standard deviation) and information ratio of the constructed portfolios. Out-of-sample optimal portfolio weights are calculated for the global minimum variance (GMV) portfolio and the Markowitz portfolio with momentum signal.

DCC method of Engle et al. (2017) using the biweekly close-to-close returns. For the cDCC model, we compute the DCC coefficients using the bivariate composite quasi-likelihood method of Pakel et al. (2020) based on contiguous pairs.[8]

We report the calculated mean portfolio realized variance and information ratio in Table 1, for both the GMV portfolio and the Markowitz portfolio with momentum signal. We observe that empirically M1 performs the best in reporting smaller mean portfolio realized variance and larger mean portfolio information ratio. The cDCC estimates have rather poor performance. The factor correlation/covariance matrix models are robust with respect to the choice of r, which coincides with the finding of Fan et al. (2013) that the covariance estimate is robust to the overestimation of r. The results confirm the advantage of using the factor model for the correlation matrix rather than the covariance matrix, as well as the use of high-frequency data. We also achieve better results by using the nonlinear shrinkage estimate for the covariance matrices. Interestingly, although the DCC-Shrinkage estimate of Engle et al. (2017) does not utilize high-frequency data, it performs quite well compared against M2 and M3 for the Markowitz portfolio with momentum signal. This may further suggests the good performance of the nonlinear shrinkage estimate of Ledoit and Wolf (2012) and the DCC structure.[9] Apart from the volatility and information ratio, we also calculate other statistics of the out-of-sample optimal portfolios as in Brito et al. (2018). The results are reported in Table A.6 of the Online Appendix. The cDCC portfolios have extreme short positions and our proposed method M1 tends to generate very balanced portfolios.

4. Conclusions

We apply the factor approach to the correlation matrix to model and forecast large covariance matrices using high-frequency data. The dynamical structure of the correlation matrices is assumed to be driven by a low-dimension latent process. The realized variance of individual assets is forecasted using the HAR model. We forecast the large covariance matrix by combining the short-memory estimated correlation matrix and the HAR realized volatilities. Our empirical study shows that our method performs the best among some alternative methods in the literature.

8 We also fit these models using daily close-to-close returns. But poorer results are obtained.

9 We perform some additional robustness checks for our empirical findings. We fit the diagonal-VAR(q) models for the vectorized factor matrices with $q = 2$. Results are similar to the case for $q = 1$ and can be found in the Online Appendix. Results are very poor if we let r be 1 or 2.

Acknowledgments

Yingjie Dong acknowledges partial support provided by the projects 71803021 and 71790604 of the National Natural Science Foundation of China, and the project 19ZDA101 of the National Social Science Fund of China.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2020.109465.

References

Aielli, Gian Piero, 2013. Dynamic conditional correlation: on properties and estimation. J. Bus. Econom. Statist. 31 (3), 282–299.

Brito, Diego, Medeiros, Marcelo C., Ribeiro, Ruy M., 2018. Forecasting large realized covariance matrices: The benefits of factor models and shrinkage. Available at SSRN: https://ssrn.com/abstract=3163668, Working paper.

Callot, Laurent A.F., Kock, Anders B., Medeiros, Marcelo C., 2016. Modeling and forecasting large realized covariance matrices and portfolio choice. J. Appl. Econometrics 32 (1), 140–158.

Corsi, Fulvio, 2009. A simple approximate long-memory model of realized volatility. J. Financ. Econ. 7 (2), 174–196.

Engle, Robert, 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econom. Statist. 20 (3), 339–350.

Engle, Robert, Ledoit, Olivier, Wolf, Michael, 2017. Large dynamic covariance matrices. J. Bus. Econom. Statist. 37 (2), 363–375.

Fan, Jianqing, Liao, Yuan, Liu, Han, 2016. An overview of the estimation of large covariance and precision matrices. Econom. J. 19 (1), C1–C32.

Fan, Jianqing, Liao, Yuan, Mincheva, Martina, 2013. Large covariance estimation by thresholding principal orthogonal complements. J. R. Stat. Soc. Ser. B Stat. Methodol. 75 (4), 603–680.

Johnstone, Iain M., 2001. On the distribution of the largest eigenvalue in principal components analysis. Ann. Statist. 29 (2), 295–327.

Ledoit, Olivier, Wolf, Michael, 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Statist. 40 (2), 1024–1060.

Pakel, Cavit, Shephard, Neil, Sheppard, Kevin, Engle, Robert F., 2020. Fitting vast dimensional time-varying covariance models. J. Bus. Econom. Statist. 1–17. http://dx.doi.org/10.1080/07350015.2020.1713795.

Tao, Minjing, Wang, Yazhen, Yao, Qiwei, Zou, Jian, 2011. Large volatility matrix inference via combining low-frequency and high-frequency approaches. J. Amer. Statist. Assoc. 106 (495), 1025–1040.

Tse, Yiu-Kuen, Tsui, Albert K.C., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. J. Bus. Econom. Statist. 20 (3), 351–362.

Zhang, Lan, Mykland, Per A., Aït-Sahalia, Yacine, 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. J. Amer. Statist. Assoc. 100 (472), 1394–1411.