Optimal Financial Portfolio Using Graphical Lasso Under Unstable Environment

Tae-Hwy Lee* Ekaterina Seregina†

Abstract
Unstable environments raise challenges for constructing a financial portfolio. In such scenarios, it is unrealistic to assume constant portfolio weights, whereas estimating weights using only post-break observations omits the information prior to the break point. This paper visualizes stock returns as a network of interacting entities and generalizes network inference in the presence of structural breaks. We estimate time-varying portfolio weights using pre- and post-break data when the stock returns are driven by common factors. Using the example of a strong structural break caused by the first wave of COVID-19 pandemic, we demonstrate that combining pre- and post-break observations for estimating portfolio weights improves portfolio return and Sharpe Ratio compared to constant weights and weights that use only post-break observations.

Key Words: Time-Varying Portfolio, Common Factors, Structural Break, Graphical Lasso, ADMM

1. Introduction

Precision matrix represents a network of interacting entities, such as corporations or genes. When the data is Gaussian, the sparsity in the precision matrix encodes the conditional independence graph - two variables are conditionally independent given the rest if and only if the entry corresponding to these variables in the precision matrix is equal to zero. Inferring the network is important for portfolio allocation problem. At the same time, the financial network changes over time, that is, the relationships between companies can change either smoothly, or abruptly (e.g. as a response to an unexpected policy shock, or in the times of economic downturns). Therefore, it is important to account for time-varying nature of stock returns.

There are two streams of literature that study time-varying networks. The first one models dynamics in the precision matrix locally. Zhou et al. [2010] develop a nonparametric method for estimating time-varying graphical structure for multivariate Gaussian distributions using an ℓ_1-penalized log-likelihood. They find out that if the covariances change smoothly over time, the covariance matrix can be estimated well in terms of predictive risk even in high-dimensional problems. Lu et al. [2015] introduce nonparanormal graphical models that allow to model high-dimensional heavy-tailed systems and the evolution of their network structure. They show that the estimator consistently estimates the latent inverse Pearson correlation matrix. The second stream of literature allows the network to vary with time by introducing two different frequencies. Hallac et al. [2017] study time-varying Graphical Lasso with smoothing evolutionary penalty.

In this paper we tackle two aspects of structural estimation of complex systems: (1) the presence of common factors in financial asset returns, and (2) the dynamical nature of a financial system that evolves in time due to structural breaks. In Lee and Seregina [2020] we showed the importance of incorporating the first aspect into graphical modeling. This paper argues that portfolio performance can be further improved when the second aspect is added. Our work provides a unified framework to generalize network inference in the

*University of California Riverside, 900 University Ave, Riverside, CA 92521. Email: taelee@ucr.edu
†Colby College, 4000 Mayflower Hill Dr, Waterville, ME 04901. Email: eseregin@colby.edu

presence of structural breaks. We propose to estimate time-varying precision matrix for portfolio optimization problem using pre- and post-break data when the stock returns are driven by common factors. Alternating direction method of multipliers (ADMM) is used to derive a closed-form solution for precision matrix. We use the example of a strong structural break caused by the first wave of COVID-19 pandemic and demonstrate that our estimator improves portfolio return and Sharpe Ratio compared to a constant precision matrix estimator and an estimator that uses only post-break observations.

2. Time-Varying Portfolio Using Graphical Methods

Suppose we observe p assets over T period of time. Let $\mathbf{r}_t = (r_{1t}, r_{2t}, \ldots, r_{pt})' \sim \mathcal{D}(\mathbf{m}, \boldsymbol{\Sigma})$ be a $p \times 1$ vector of stock returns drawn from a distribution $\mathcal{D}$, where $\mathbf{m}$ and $\boldsymbol{\Sigma}$ are the unconditional mean and covariance matrix of the returns.

Let daily stock returns follow a K-factor model:

$$\mathbf{r}_t = \boldsymbol{\alpha} + \mathbf{B}\mathbf{f}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, T, \tag{1}$$

where $\mathbf{f}_t = (f_{1t}, \ldots, f_{Kt})'$ are the factors, $\mathbf{B}$ is a $p \times K$ matrix of factor loadings, and $\boldsymbol{\varepsilon}_t$ is the idiosyncratic component that cannot be explained by the common factors.

We model the change in precision matrix due to N known structural breaks. Define $n_i \equiv t_i - t_{i-1}$ to be the sample between the i-th and $(i-1)$-th break points, where $i = 1, \ldots, N$, $\sum_{i=1}^{N} n_i = T$, $N \leq T$.

We propose that the dynamics of the system evolves through the precision matrix of the idiosyncratic component. Let $\boldsymbol{\Sigma}_{\varepsilon,i}$, $\boldsymbol{\Sigma}_f$, and $\boldsymbol{\Sigma}_i$ be covariance matrices of idiosyncratic part, factors and stock returns in regime i. Define the corresponding precision matrices to be $\boldsymbol{\Theta}_{\varepsilon,i} \equiv \boldsymbol{\Sigma}_{\varepsilon,i}^{-1}$, $\boldsymbol{\Theta}_f \equiv \boldsymbol{\Sigma}_f^{-1}$, and $\boldsymbol{\Theta}_i \equiv \boldsymbol{\Sigma}_i^{-1}$.

Unknown factors, loadings and the constant term are estimated using PCA in a standard way (Stock and Watson [2002]), and the estimators are denoted $\widehat{\mathbf{f}}_t$, $\widehat{\mathbf{B}}$, and $\widehat{\boldsymbol{\alpha}}$. Given a sample of the estimated factors $\{\widehat{\mathbf{f}}_t\}_{t=1}^{T}$, let $\widehat{\boldsymbol{\Sigma}}_f = (1/T)\sum_{t=1}^{T} \widehat{\mathbf{f}}_t \widehat{\mathbf{f}}_t'$ be the sample counterpart of the covariance matrix and $\widehat{\boldsymbol{\Theta}}_f = \widehat{\boldsymbol{\Sigma}}_f^{-1}$. Also, we can obtain $\widehat{\boldsymbol{\varepsilon}}_t = \mathbf{r}_t - \widehat{\boldsymbol{\alpha}} - \widehat{\mathbf{B}}\widehat{\mathbf{f}}_t$. To model dynamics in $\boldsymbol{\Theta}_{\varepsilon,i}$ we use the following optimization problem:

$$\widehat{\boldsymbol{\Theta}}_{\varepsilon,i} = \arg\min_{\boldsymbol{\Theta}_{\varepsilon,i} \succ 0} \sum_{i=1}^{N} n_i \Big[\operatorname{trace}\Big(\widehat{\boldsymbol{\Sigma}}_{\varepsilon,i}\boldsymbol{\Theta}_{\varepsilon,i}\Big) - \log\det \boldsymbol{\Theta}_{\varepsilon,i}\Big] + \lambda \|\boldsymbol{\Theta}_{\varepsilon,i}\|_{\text{od},1} + \beta \sum_{i=2}^{N} \psi(\boldsymbol{\Theta}_{\varepsilon,i} - \boldsymbol{\Theta}_{\varepsilon,i-1}). \tag{2}$$

where $\widehat{\boldsymbol{\Sigma}}_{\varepsilon,i} = \frac{1}{n_i}\sum_{k=1}^{n_i} \widehat{\varepsilon}_{i,k}\widehat{\varepsilon}_{i,k}'$ and $\|\boldsymbol{\Theta}_{\varepsilon,i}\|_{\text{od},1} = \sum_{l \neq q} |\boldsymbol{\Theta}_{lq,\varepsilon,i}|$ where $\boldsymbol{\Theta}_{lq,\varepsilon,i}$ is the lq-th element of matrix $\boldsymbol{\Theta}_{\varepsilon,i}$. The optimization problem in (2) has two tuning parameters: λ, which determines the sparsity level of the network, and β, which controls the strength of resemblance between two neighboring precision estimators. In the empirical exercise we use the first 2/3 of the training data to estimate portfolio weights and jointly tune λ and β in the remaining 1/3 to yield the highest Sharpe Ratio. The smoothing function $\psi(\cdot)$ in (2) can be Lasso ($\psi = \sum_{l,q} |\cdot|$), Group Lasso ($\psi = \sum_q \|\cdot_q\|_2$), Laplacian ($\psi = \sum_{l,q} (\cdot_{lq})^2$), Max norm penalty ($\psi = \sum_q \max_l |\cdot_{lq}|$).

To estimate (2) we use the ADMM algorithm. Once $\boldsymbol{\Theta}_{\varepsilon,i}$ is estimated, we combine estimated factors, loadings and precision matrix of the idiosyncratic components using Sherman-Morrison-Woodbury formula to estimate the final precision matrix of excess returns:

$$\widehat{\boldsymbol{\Theta}}_i = \widehat{\boldsymbol{\Theta}}_{\varepsilon,i} - \widehat{\boldsymbol{\Theta}}_{\varepsilon,i}\widehat{\mathbf{B}}[\widehat{\boldsymbol{\Theta}}_f + \widehat{\mathbf{B}}'\widehat{\boldsymbol{\Theta}}_{\varepsilon,i}\widehat{\mathbf{B}}]^{-1}\widehat{\mathbf{B}}'\widehat{\boldsymbol{\Theta}}_{\varepsilon,i}. \tag{3}$$

We can use precision matrix estimated using (3) to compute optimal portfolio weights $\widehat{\mathbf{w}}_i = f(\widehat{\mathbf{\Theta}}_i)$. Possible choice for $f(\mathbf{\Theta}_i)$ is a global minimum-variance portfolio (GMV, $f(\mathbf{\Theta}_i) = (\boldsymbol{\iota}'\mathbf{\Theta}_i\boldsymbol{\iota})^{-1}\mathbf{\Theta}_i\boldsymbol{\iota}$), where $\boldsymbol{\iota}$ is a $p \times 1$ vector of ones. We call the aforementioned procedure Time-Varying Factor Graphical Lasso (TVFGL) and summarize it in Algorithm 1.

Algorithm 1

1: **(FM)** Estimate $\widehat{\mathbf{f}}_t$ and $\widehat{\mathbf{B}}_t$ in (1). Get $\widehat{\mathbf{\Sigma}}_f$, $\widehat{\mathbf{\Theta}}_f$ and $\widehat{\varepsilon}_t = \widehat{\mathbf{B}}\widehat{\mathbf{f}}_t - \mathbf{r}_t$.
2: **(TVGL)** Solve (2) using ADMM to get $\widehat{\mathbf{\Theta}}_{\varepsilon,i}$.
3: **(TVFGL)** Use $\widehat{\mathbf{\Theta}}_{\varepsilon,i}$, $\widehat{\mathbf{\Theta}}_f$ and $\widehat{\mathbf{B}}$ from Steps 1-2 to get $\widehat{\mathbf{\Theta}}_i$ in Equation (3).
4: Use $\widehat{\mathbf{\Theta}}_i$ to get portfolio weights $\widehat{\mathbf{w}}_i = f(\widehat{\mathbf{\Theta}}_i)$.

3. Empirical Application

The full sample of the time-series of the daily returns has 1,089 observations on 500 components of the S&P500 and runs from January 3, 2017 to April 30, 2021. Using the behavior of the composite S&P500 Index, we identified February 10, 2020 as a break point. The training sample consists of 1,007 observations and runs from January 3, 2017 to December 31, 2020. This leaves 82 observations for the test sample which runs from January 1, 2021 to April 30, 2021. The following competing methods are considered: constant precision matrix estimated using FGL (Lee and Seregina [2020]); FGL-postbreak that uses only post-break observations; and TVFGL which accounts for the break. Table 1 reports the out-of-sample (OOS) portfolio mean return, standard deviation (risk), and the Sharpe Ratio.

Table 1: Daily portfolio returns, risk and Sharpe ratio.

	GMV		
	Mean	**Risk**	**Sharpe Ratio**
FGL	0.0006	0.0056	0.1155
FGL-postbreak	0.0009	0.0062	0.1390
TVFGL	0.0019	0.0133	0.1441

As evidenced by the results in Table 1, accounting for the break improves the OOS Sharpe Ratio and return of the GMV portfolio.

4. Summary

Dynamic nature of the financial system requires an approach to estimate time-varying portfolio weights. This paper addresses this issue by focusing on a precision matrix, the main object used for estimating portfolio weights. We visualize stock returns as a network of interacting entities and generalize network inference in the presence of structural breaks. We estimate time-varying portfolio weights using pre- and post-break data when the stock returns are driven by common factors. Using the example of a strong structural break caused by the first wave of COVID-19 pandemic, we demonstrate that combining pre- and post-break observations improves portfolio return and Sharpe Ratio compared to constant weights and weights that use only post-break observations.

References

Aielli, G. P. (2013). Dynamic conditional correlation: On properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Bauwens, L., Storti, G., and Violante, F. S. (2012). Dynamic conditional correlation models for realized covariance matrices. In *Proceedings of the Center for Operations Research and Econometrics, UCL.*

Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Optimization*, 3(1):1–122.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Danaher, P., Wang, P., and Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(2):373–397.

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

Friedman, J., Hastie, T., and Tibshirani, R. (2007). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441.

Greenewald, K. H., Park, S., Zhou, S., and Giessing, A. (2017). Time-dependent spatially varying graphical models, with application to brain fmri data analysis. In *NIPS*.

Guo, J., Levina, E., Michailidis, G., and Zhu, J. (2011). Joint estimation of multiple graphical models. *Biometrika*, 98(1):1–15.

Hallac, D., Park, Y., Boyd, S., and Leskovec, J. (2017). Network inference via the time-varying graphical lasso. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, pages 205–213, New York, NY, USA. ACM.

Lee, T.-H. and Seregina, E. (2020). Optimal portfolio using factor graphical lasso. *arXiv:2011.00435*.

Lu, J., Kolar, M., and Liu, H. (2015). Post-regularization inference for time-varying nonparanormal graphical models. *Journal of Machine Learning Research*, 18:203:1–203:78.

Pakel, C., Shephard, N., Sheppard, K., and Engle, R. (2018). Fitting vast dimensional time-varying covariance models. Technical Report FIN-08-009, NYU Working Paper.

Parikh, N. and Boyd, S. (2014). Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179.

Zhou, S., Lafferty, J., and Wasserman, L. (2010). Time varying undirected graphs. *Machine Learning*, 80(2):295–319.