High-dimensional minimum variance portfolio estimation based on high frequency data

Journal of Econometrics (2020)

Author: T. Tony Cai[1] Jianchang Hu[2] Yingying Li[3] Xinghua Zheng[4]

January 10, 2023

[1]Department of Statistics, The Wharton School, University of Pennsylvania

[2]Department of Statistics, University of Wisconsin, Madison

[3]Department of ISOM and Department of Finance, Hong Kong University of Science and Technology

[4]Department of ISOM, Hong Kong University of Science and Technology

Outline

Background

Estimation of MVP

We want to find out the estimation of high-dimensional MVP using high-frequency data, i.e., given p assets, we aim to find ω such that

$$\arg\min_{\omega} \boldsymbol{\omega^T \Sigma \omega} \quad \text{subject to} \quad \boldsymbol{\omega^T 1} = 1 \tag{1}$$

where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_p)^T$ represents the weights put on different assets. And the optimal solution is given by

$$\boldsymbol{\omega_{opt}} = \frac{\boldsymbol{\Sigma^{-1} 1}}{\boldsymbol{1^T \Sigma^{-1} 1}} \tag{2}$$

which also yields the minimum risk

$$R_{min} = \boldsymbol{\omega_{opt}^T \Sigma \omega_{opt}} = \frac{1}{\boldsymbol{1^T \Sigma^{-1} 1}} \tag{3}$$

Background

Since we do not have true covariance, we use the sample covariance instead. This leads to several issues.

- The actual risk of the "plug-in" portfolio (the portfolio that uses the sample covariance) can be devastatingly higher than the theoretical minimum risk.
- The perceived risk can be lower than the theoretical minimum risk.

To solve these problems, Fan et al (2012, [3]) add a gross-exposure constraint

$$\arg\min_{w} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} \quad \text{subject to} \quad \boldsymbol{w}^\top \mathbf{1} = 1 \text{ and } \|\boldsymbol{w}\|_1 \leq \lambda \tag{4}$$

where $\|w\|_i = \Sigma_{i=1}^{p} |w_i|$ and λ is a chosen constant.

Background

Since the difference between the risk of an estimated portfolio and the minimum risk going to zero **may not be sufficient to guarantee** optimality. (under rather general assumptions, the minimum risk $R_{\min} = 1/\mathbf{1}^T\Sigma^{-1}\mathbf{1}$ may go to zero as the number of assets $p \to \infty$.

Based on this consideration, we turn to find $\hat{\boldsymbol{w}}$ which satisfies a stronger sense of consistency in that the ratio between the risk of the estimated portfolio and the minimum risk goes to 1, i.e.,

$$\frac{R(\hat{\boldsymbol{w}})}{R_{min}} \xrightarrow{p} 1 \quad \text{as} \quad p \to \infty$$

Research Questions

- What is the estimator of minimum variance portfolio that can accommodate stochastic volatility and market microstructure noise?
- What is the estimator of the minimum risk?

High-frequency Data Model

We assume that the latent p-dimensional log-price process $(\boldsymbol{X}_t)$ follows a diffusion model:

$$d\mathbf{X}_t = \boldsymbol{\mu}_t dt + \Theta_t d\mathbf{W}_t, \qquad \text{for } t \geq 0 \tag{5}$$

where $\boldsymbol{\mu}_t$ is the drift process, $\boldsymbol{\Theta}_t$ is a $p \times p$ matrix-valued process called co-volatility process, and $\boldsymbol{W}_t$ is a p-dimensional Brownian motion.

Both $\boldsymbol{\mu}_t$ and $\boldsymbol{\Theta}_t$ are stochastic, càdlàg, and dependent on $\boldsymbol{W}_t$, all defined on a common filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$

High-frequency Data Model

Let

$$\mathbf{\Sigma}_t = \mathbf{\Theta}_t \mathbf{\Theta}_t^\top := \left(\sigma_t^{ij}\right)$$

be the spot covariance matrix process. The *ex-post* integrated covariance (ICV) matrix over an integral, say $[0, 1]$, is

$$\mathbf{\Sigma}_{\mathrm{ICV}} = \mathbf{\Sigma}_{\mathrm{ICV},1} = \left(\sigma^{ij}\right) := \int_0^1 \mathbf{\Sigma}_t dt$$

And in this slide, the inverse of $\mathbf{\Sigma}_{\mathrm{ICV}}$ is denoted as $\mathbf{\Omega}_{\mathrm{ICV}}$, i.e., $\mathbf{\Omega}_{\mathrm{ICV}} := \mathbf{\Sigma}_{\mathrm{ICV}}^{-1}$

High-frequency Data Model

Σ_{ICV} is only measurable to $\mathcal{F}_1$, and so is $R_{\min}$. So it is impossible to construct a portfolio that is measurable to $\mathcal{F}_0$ to achieve the minimum risk $R_{\min}$

The **practical implementation of MVP** relies on making forecasts of Σ_{ICV} based on historical data.

The simplest approach is to assume $\Sigma_{\mathsf{ICV,\ t}} \approx \Sigma_{\mathsf{ICV,t+1}}$, where $\Sigma_{\mathsf{ICV,t}}$ stands for the ICV matrix in period $[t-1,t]$. (The volatility process is often found to be nearly unit root, in which case the one-step ahead prediction is approximately the current value)
If we can construct a portfolio $\boldsymbol{w}$ based on the observations during $[t-1,t]$ that can approximately minimize the **ex-post** risk $\boldsymbol{w}^T\Sigma_{\mathsf{ICV,t}}\boldsymbol{w}$, then if we hold the portfolio during the next period $[t,t+1]$, the actual risk is **still approximately minimized**.

High-frequency Case with no Microstructure Noise

Intuitive approach: We use the constrained ℓ_1- minimization for inverse matrix estimation (CLIME, Cai et al, [2])(under i.i.d. observation setting) to estimate the MVP. Denote the covariance matrix as Σ. Let $\Omega := \Sigma^{-1}$ be the precision matrix. The CLIME estimator is defined as,

$$\widehat{\mathbf{\Omega}}_{\mathrm{CLIME}} := \arg\min_{\Omega'} \left\|\mathbf{\Omega}'\right\|_1 \text{ subject to } \left\|\widehat{\mathbf{\Sigma}}\mathbf{\Omega}' - \mathbf{I}\right\|_\infty \leq \lambda \tag{6}$$

where $\hat{\mathbf{\Sigma}}$ is the sample covariance matrix. The λ is a tuning parameter and is usually chosen via cross-validation.

CLIME Method

The CLIME method is designed for the following uniformity class of precision matrices. For any $0 \leq q < 1, s_0 = s_0(p) < \infty$ and $M = M(p) < \infty$, let

$$\begin{aligned}\mathcal{U}(q, s_0, M) =& \{\mathbf{\Omega} = (\mathbf{\Omega}_{ij})_{p \times p} : \mathbf{\Omega} \text{ positive definite}, \\ & \|\mathbf{\Omega}\|_{L_1} \leq M, \max_{1 \leq i \leq p} \sum_{j=1}^{p} |\mathbf{\Omega}_{ij}|^q \leq s_0\}\end{aligned} \tag{7}$$

When the observations are i.i.d. sub-Gaussian and Ω belongs to $\mathcal{U}(q, s_0(p), M(p))$, this method establishes consistency of $\hat{\Omega}_{\text{CLIME}}$ when $M^{2-2q} s_0 (\log p/n)^{(1-q)/2} \to 0$ holds.

However, in the high-frequency setting, the returns are not *i.i.d.*, so **the results above cannot apply**.

Sparsity Assumption

The sparsity assumption $\mathbf{\Omega} = \mathbf{\Sigma}^{-1} \in \mathcal{U}(q, s_0, M)$ (See Equation (7)) appears to be reasonable in financial applications. The following gives an example.

If the returns follow a conditional multivariate normal distribution matrix $\mathbf{\Sigma}$, then the (i, j)-th element in $\mathbf{\Omega}$ being 0 is equivalent to that the returns of the i-th and j-th assets are conditionally independent given the other asset returns. And for stocks in different sectors, many pairs might be **conditionally independent or only weak dependent**.

CLIME Method in the High-Frequency Setting

Our goal is to estimate $\mathbf{\Omega}_{\mathsf{ICV}} := \mathbf{\Sigma}_{\mathsf{ICV}}^{-1}$. And we use $\hat{\mathbf{\Sigma}}$ in Equation (6). When the true log-price are observed, one of the most commonly used estimators for $\mathbf{\Sigma}_{\mathsf{ICV}}$ is the realized covariance matrix (RCV).

For each asset i, the observations at stage n are $(x^i_{t_\ell^{i,n}})$, where $0 = t_0^{i,n} < t_1^{i,n} < \cdots < t_{N_i}^{i,n} = 1$ are the observation times. The n characterizes the observation frequency. As $n \to \infty$, $N_i \to \infty$. The synchronous observation case corresponds to

$$t_\ell^{i,n} \equiv t_\ell^n \quad \text{for all} \quad i = 1, \ldots, p$$

which reduces to,

$$t_\ell^{i,n} = t_\ell^n = \ell/n, \ell = 0, 1, \ldots, n \tag{8}$$

MVP Estimator with no Microstructure Noise

The **resulting MVP estimator** is,

$$\widehat{\boldsymbol{w}}_{\mathrm{CLIME-SV}} = \frac{\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-SV}}\mathbf{1}}{\mathbf{1}^{\top}\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-SV}}\mathbf{1}} \tag{9}$$

which is associated with a risk of,

$$\begin{aligned} R_{\mathrm{CLIME-SV}} &= \widehat{\boldsymbol{w}}_{\mathrm{CLIME-SV}}^{\top}\boldsymbol{\Sigma}_{\mathrm{ICV}}\widehat{\boldsymbol{w}}_{\mathrm{CLIME-SV}} \\ &= \frac{\left(\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-SV}}\mathbf{1}\right)^{\top}\boldsymbol{\Sigma}_{\mathrm{ICV}}\left(\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-SV}}\mathbf{1}\right)}{\left(\mathbf{1}^{\top}\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-SV}}\mathbf{1}\right)^{2}} \end{aligned} \tag{10}$$

Induction of MVP Estimator in Equation (9)

In the synchronous observation (Equation 8), let

$$\Delta \boldsymbol{X}_\ell := \boldsymbol{X}_{t_\ell^n} - \boldsymbol{X}_{t_{\ell-1}^n}$$

be the log-return vector over the time period $[t_{\ell-1}^n, t_\ell^n]$. Then, the RCV matrix is defined as,

$$\widehat{\mathbf{\Sigma}}_{\mathrm{RCV}} = \sum_{\ell=1}^{n} \Delta \mathbf{X}_\ell \left(\Delta \mathbf{X}_\ell\right)^\top \tag{11}$$

We now can conduct the **constrained ℓ_1-minimization for inverse matrix estimation with stochastic volatility** (CLIME-SV), $\mathbf{\Omega}_{\text{CLIME-SV}}$,

$$\widehat{\mathbf{\Omega}}_{\mathrm{CLIME-SV}} := \arg\min_{\mathbf{\Omega}'} \left\|\mathbf{\Omega}'\right\|_1 \text{ subject to } \left\|\widehat{\mathbf{\Sigma}}_{\mathrm{RCV}}\mathbf{\Omega}' - \mathbf{I}\right\|_\infty \leq \lambda \tag{12}$$

High-frequency Case with Microstructure Noise

However, in general, the observation prices are believed to be contaminated by microstructure noise. The true log-price $(X^i_{t^{i,n}_\ell})$, for each asset i, at stage n are

$$Y^i_{t^{i,n}_\ell} = X^i_{t^{i,n}_\ell} + \varepsilon^i_\ell \tag{13}$$

where the last term, ε^i_ℓ's represent microstructure noise.

In this case, if we simply plug $(Y^i_{t^{i,n}_\ell})$ into the formula of RCV in Equation (11), the resulting estimator is not **consistent** even when the dimension p is fixed.

Consistent Estimators in the Univariate Case

- Two-scales realized volatility (TSRV, Zhang et al. (2005))
- Multi-scale realized volatility (MSRV, Zhang (2006))
- Pre-averaging estimator (PAV, Jacod et al. (2009), Podolskij and Vetter (2009) and Jacod et al. (2019))
- Realized kernels (RK, Barndorff-Nielsen et al. (2008))
- Quasi-maximum likelihood estimator (QMLE, Xiu (2010))
- Estimated-price realized volatility (ERV, Li et al. (2016))
- Unified volatility estimator (UV, Li et al. (2018))

Note: These estimators are **not consistent** in the high-dimensional setting.

In this paper, we choose to work with **PAV**, with the equidistant time setting (8). **Asynchronicity** can be dealt with by using existing data synchronization techniques.

Implementation of PAV Estimator

To implement PAV estimator, we fix a constant $\theta > 0$ and let $k_n = [\theta n^{1/2}]$ be the window length over which the averaging takes place. Define

$$\overline{\mathbf{Y}}_k^n = \frac{\sum_{i=k_n/2}^{k_n-1} \mathbf{Y}_{t_{k+i}} - \sum_{i=0}^{k_n/2-1} \mathbf{Y}_{t_{k+i}}}{k_n}$$

The PAV with weight function $g(x) = x \wedge (1-x)$ for $x \in (0,1)$ is defined as,

$$\widehat{\mathbf{\Sigma}}_{\mathrm{PAV}} = \frac{12}{\theta\sqrt{n}} \sum_{k=0}^{n-k_n+1} \overline{\mathbf{Y}}_k^n \cdot \left(\overline{\mathbf{Y}}_k^n\right)^\top - \frac{6}{\theta^2 n} \operatorname{diag}\left(\sum_{k=1}^{n} \left(\Delta Y_{t_k}^i\right)^2\right)_{i=1,\ldots,p} \quad (14)$$

We now define the **constrained ℓ_1-minimization for inverse matrix estimation with stochastic volatility and microstructure noise**, $\widehat{\mathbf{\Omega}}_{\mathrm{CLIME-SVMN}}$, as

$$\widehat{\mathbf{\Omega}}_{\mathrm{CLIME-SVMN}} := \arg\min_{\mathbf{\Omega}'} \left\|\mathbf{\Omega}'\right\|_1 \text{ subject to } \left\|\widehat{\mathbf{\Sigma}}_{\mathrm{PAV}}\mathbf{\Omega}' - \mathbf{I}\right\|_\infty \leq \lambda \quad (15)$$

Minimum Risk with CLIME Based Estimators with Sparsity Assumption

Recall Equation (3), $R_{min} = \boldsymbol{\omega}_{\boldsymbol{opt}}^{\boldsymbol{T}} \boldsymbol{\Sigma} \boldsymbol{\omega}_{\boldsymbol{opt}} = \frac{1}{\mathbf{1}^{\boldsymbol{T}} \boldsymbol{\Sigma}^{-1} \mathbf{1}}$. Our estimator under sparsity assumption with no microstructure noise is,

$$\widehat{R}_{\text{CLIME-SV}} = \frac{1}{\mathbf{1}^{\top} \widehat{\Omega}_{\text{CLIME-SV}} \mathbf{1}} \tag{16}$$

where $\widehat{\Omega}_{\text{CLIME-SV}}$ is given in Equation (12). If there's microstructure noise, we have the corresponding minimum risk,

$$\widehat{R}_{\text{CLIME-SVMN}} = \frac{1}{\mathbf{1}^{\top} \widehat{\Omega}_{\text{CLIME-SVMN}} \mathbf{1}} \tag{17}$$

where $\widehat{\Omega}_{\text{CLIME-SVMN}}$ is given in Equation (15)

Without Sparsity Assumption: Low-Frequency i.i.d. Returns

This estimator is hence more suitable for the **low-frequency setting** and can be used to estimate the minimum risk over a long time period.

Suppose we observe n i.i.d. returns $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ (at low frequency). Let $\boldsymbol{S}$ be the sample covariance matrix, and $\boldsymbol{w}_p = \frac{\boldsymbol{S}^{-1}\mathbf{1}}{\mathbf{1}^T\boldsymbol{S}^{-1}\mathbf{1}}$ be the "plug-in" portfolio. The corresponding perceived risk is $\hat{R}_p = \boldsymbol{w}_p^T \boldsymbol{S} \boldsymbol{w}_p$.

We have the following results on the relationship between $\hat{R}_p$ and the minimum risk $R_{\min}$ based on which a **consistent estimator** of the minimum risk is constructed.

Relationship between $\hat{R}_p$ and $R_{\min}$

Suppose that the returns $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose that both n and $p \to \infty$, in such a way that $p_n := p/n \to \rho \in (0,1)$. Then

$$\left| \frac{\widehat{R}_p}{R_{\min}} - (1 - \rho_n) \right| \xrightarrow{p} 0 \tag{18}$$

Therefore, if we define

$$\widehat{R}_{\min} = \frac{1}{1 - \rho_n} \widehat{R}_p \tag{19}$$

then, we have

$$\frac{\widehat{R}_{\min}}{R_{\min}} \xrightarrow{p} 1 \tag{20}$$

and furthermore,

$$\sqrt{n - p} \left(\frac{\widehat{R}_{\min}}{R_{\min}} - 1 \right) \Rightarrow \mathcal{N}(0, 2) \tag{21}$$

Interpretation of above Relationship

Convergence (18) explains why the perceived risk is systematically lower than the minimum risk.
Convergence (21) shows the "blessing" of dimensionality: **the higher the dimension, the more accurate the estimation.**

In Basak et al (2009, [1]), it is shown that the risk of the "plug-in" portfolio is on average a higher-than-one multiple of the minimum risk.
Since n and $p \to \infty$ and $p_n := p/n \to \rho \in (0,1)$ their result can be strengthened to be that the risk of the "plug-in" portfolio is, with probability approaching one, a larger-than-one multiple of the minimum risk.
Using the results of Basak et al, we can show that,

$$\frac{R\left(\boldsymbol{w}_p\right)}{R_{\min}} \xrightarrow{p} \frac{1}{1-\rho} \tag{22}$$

where $R(\boldsymbol{w}_p) = \boldsymbol{w}_p^T \boldsymbol{\Sigma} \boldsymbol{w}_p$ is the risk of the "plug-in" portfolio.

High-Frequency Factor Structure

Let $\boldsymbol{r}_k = (r_{k,1}, \ldots, r_{k,p})^T, k = 1, \ldots, n$ be asset returns, which can either be high-frequency or low-frequency returns. Assume $\boldsymbol{r}_k$ admits a factor structure as follows,

$$\mathbf{r}_k = \alpha + \Gamma \mathbf{f}_k + \mathbf{z}_k \tag{23}$$

where α is a $p \times 1$ unknown vector, Γ is a $p \times m$ unknown matrix, $\boldsymbol{f}_k = (f_{k,1}, \ldots, f_{k,m})^T$ are factor returns, and $\boldsymbol{z}_k$ is $p \times 1$ random vector with mean 0 and covariance matrix $\boldsymbol{\Sigma}_{k,0} = (\sigma_{ij}^{k,0})$

We assume that, for each k, $\boldsymbol{f}_k$'s and $\boldsymbol{z}_k$'s are independent, and the pairs $(\boldsymbol{r}_k, \boldsymbol{f}_k)$ are mutually independent. Let $\boldsymbol{\Sigma}_{\boldsymbol{r},k} = \text{Cov}(\boldsymbol{r}_k)$ and $\boldsymbol{\Sigma}_{\boldsymbol{f},k} = \text{Cov}(\boldsymbol{f}_k)$, and let them be dependent on the time index k, **to accommodate the stochastic (co-)volatility**.

It's impossible to estimate individual $\boldsymbol{\Sigma}_{\boldsymbol{r},k}$ and $\boldsymbol{\Sigma}_{\boldsymbol{f},k}$, but it's possible to estimate their means $\boldsymbol{\Sigma}_{\mathbf{r}} = \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{\Sigma}_{\mathbf{r},k}$, $\boldsymbol{\Sigma}_{\mathbf{f}} = \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{\Sigma}_{\mathbf{f},k}$, and $\boldsymbol{\Sigma}_0 = \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{\Sigma}_{k,0}$, and the corresponding $\boldsymbol{\Omega}_0 = \boldsymbol{\Sigma}_{\mathbf{0}}^{-1}$

CLIME Method in Factor Structure

Target: Estimate the precision matrix $\boldsymbol{\Omega_r} := \boldsymbol{\Sigma_r}^{-1}$

We define the **constrained ℓ_1-minimization for inverse matrix estimation adjusted for factor** (CLIME-F), $\hat{\boldsymbol{\Omega}}_{\text{CLIME-F}}$, as

$$\widehat{\boldsymbol{\Omega}}_{\text{CLIME-F}} = \widehat{\boldsymbol{\Omega}}_0 - \widehat{\boldsymbol{\Omega}}_0 \widehat{\Gamma} \left(\mathbf{S}_{\mathbf{f}}^{-1} + \widehat{\Gamma}^\top \widehat{\boldsymbol{\Omega}} \widehat{0} \widehat{\Gamma} \right)^{-1} \widehat{\Gamma}^\top \widehat{\boldsymbol{\Omega}}_0 \tag{24}$$

where $\boldsymbol{S_f}$ is the sample covariance of $\boldsymbol{f_k}$. And also, the **MVP estimator** is

$$\widehat{\boldsymbol{w}}_{\mathrm{CLIME-F}} = \frac{\widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-F}} \mathbf{1}}{\mathbf{1}^\top \widehat{\boldsymbol{\Omega}}_{\mathrm{CLIME-F}} \mathbf{1}} \tag{25}$$

Induction of CLIME Method in Factor Structure

To estimate the model, we calculate the least square estimators of α and Γ, denoted by $\hat{\alpha}$ and $\hat{\Gamma}$. The residuals are $\boldsymbol{e}_k := \boldsymbol{r}_k - \hat{\alpha} - \hat{\Gamma}\boldsymbol{f}_k$. Then we have the CLIME estimator of $\boldsymbol{\Omega}_0$, denoted by $\hat{\boldsymbol{\Omega}}_0$ based on residuals.

$$\widehat{\boldsymbol{\Omega}}_0 := \arg\min_{\Omega'} \left\|\boldsymbol{\Omega}'\right\|_1 \text{ subject to } \left\|\mathbf{S_e}\boldsymbol{\Omega}' - \mathbf{I}\right\|_\infty \leq \lambda \tag{26}$$

where $\boldsymbol{S_e}$ is the sample covariance matrix of residuals. Note that, $\boldsymbol{\Sigma}_\mathbf{r} = \Gamma\boldsymbol{\Sigma}_\mathbf{f}\Gamma^\top + \boldsymbol{\Sigma}_0$, we have,

$$\boldsymbol{\Omega}_\mathbf{r} = \left(\Gamma\boldsymbol{\Sigma}_\mathbf{f}\Gamma^\top + \boldsymbol{\Sigma}_0\right)^{-1} = \boldsymbol{\Omega}_0 - \boldsymbol{\Omega}_0\Gamma\left(\boldsymbol{\Sigma}_\mathbf{f}^{-1} + \Gamma^\top\boldsymbol{\Omega}_0\Gamma\right)^{-1}\Gamma^\top\boldsymbol{\Omega}_0 \tag{27}$$

Therefore, we can define the CLIME-F in (24)

Conclusions

Conclusions

- Propose estimators of the MVP in the high-dimensional setting based on high-frequency data.
- Propose consistent estimators of minimum risk with and without sparsity assumption

Note: For the details of simulation studies and empirical studies, please refer to Sections (4) and (5) in High-dimensional minimum variance portfolio estimation based on high-frequency data.

References I

[1] Gopal K Basak, Ravi Jagannathan, and Tongshu Ma. "Jackknife estimator for tracking error variance of optimal portfolios". In: *Management Science* 55.6 (2009), pp. 990–1002.

[2] Tony Cai, Weidong Liu, and Xi Luo. "A constrained 1 minimization approach to sparse precision matrix estimation". In: *Journal of the American Statistical Association* 106.494 (2011), pp. 594–607.

[3] Jianqing Fan, Jingjin Zhang, and Ke Yu. "Vast portfolio selection with gross-exposure constraints". In: *Journal of the American Statistical Association* 107.498 (2012), pp. 592–606.