Approaching Mean-Variance Efficiency for Large Portfolios

Yingying Li

Department of ISOM & Department of Finance
Hong Kong University of Science and Technology

Based on Joint Work with Mengmeng Ao and Xinghua Zheng

Outline

Outline

Markowitz Mean-Variance Optimization

- Markowitz (mean-variance) optimization:

  maximize portfolio return given risk constraint

  $\Leftrightarrow$ minimize portfolio risk given return constraint

- The solution to Markowitz optimization is *mean-variance efficient*

Markowitz Mean-Variance Optimization

- Markowitz (mean-variance) optimization:

  maximize portfolio return given risk constraint

  $\Leftrightarrow$ minimize portfolio risk given return constraint

- The solution to Markowitz optimization is *mean-variance efficient*

Markowitz Mean-Variance Optimization

- Markowitz (mean-variance) optimization:

  maximize portfolio return given risk constraint

  $\Leftrightarrow$ minimize portfolio risk given return constraint

- The solution to Markowitz optimization is *mean-variance efficient*

Markowitz Mean-Variance Optimization

- Markowitz (mean-variance) optimization:

  maximize portfolio return given risk constraint

  $\Leftrightarrow$ minimize portfolio risk given return constraint

- The solution to Markowitz optimization is *mean-variance efficient*

The “Plug-in” Portfolio

- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\mu' \Sigma^{-1} \mu}} \Sigma^{-1} \mu$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ “plug-in” portfolio

The "Plug-in" Portfolio

- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\mu' \Sigma^{-1} \mu}} \Sigma^{-1} \mu$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ "plug-in" portfolio

The "Plug-in" Portfolio

- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ "plug-in" portfolio

The "Plug-in" Portfolio

- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ "plug-in" portfolio

The "Plug-in" Portfolio

- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ "plug-in" portfolio

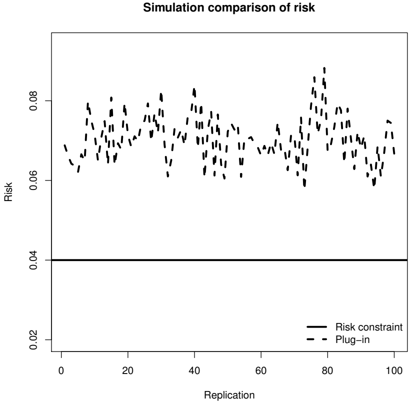
The "Plug-in" Portfolio

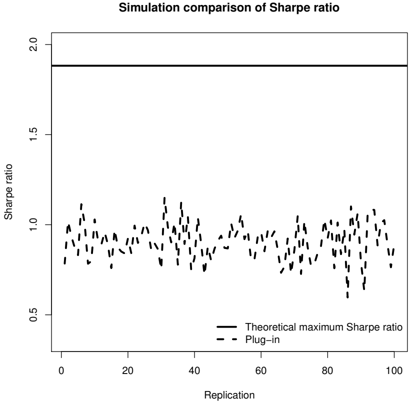
- *If the mean and covariance matrix of returns were known*
  $\Rightarrow$ optimal portfolio $\surd$

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

- We know this is impossible
- Natural/Naive approach: plug in the sample mean and sample covariance matrix
  $\Rightarrow$ "plug-in" portfolio

How well does the plug-in portfolio perform?





Challenges for Large Portfolios

- Poor performance of the plug-in portfolio
  - "Markowitz optimization enigma": Michaud (1989)
  - Best and Grauer (1991), Chopra and Ziemba (1993), Kan and Zhou (2007) etc.
- The situation worsens *as the number of assets increases*

♦ Key reason: (High) Dimensionality

$$\frac{SR(\text{plug-in})}{SR^*} \xrightarrow{P} \sqrt{\frac{1-\rho}{1+\rho/(SR^*)^2}} < \sqrt{1-\rho} < 1, \quad \text{as } \frac{N}{T} \to \rho \in (0,1)$$

Challenges for Large Portfolios

- Poor performance of the plug-in portfolio
  - "Markowitz optimization enigma": Michaud (1989)
  - Best and Grauer (1991), Chopra and Ziemba (1993), Kan and Zhou (2007) etc.
- The situation worsens *as the number of assets increases*

♦ Key reason: (High) Dimensionality

$$\frac{SR(\text{plug-in})}{SR^*} \xrightarrow{P} \sqrt{\frac{1-\rho}{1+\rho/(SR^*)^2}} < \sqrt{1-\rho} < 1, \quad \text{as } \frac{N}{T} \to \rho \in (0,1)$$

Challenges for Large Portfolios

- Poor performance of the plug-in portfolio
  - "Markowitz optimization enigma": Michaud (1989)
  - Best and Grauer (1991), Chopra and Ziemba (1993), Kan and Zhou (2007) etc.
- The situation worsens *as the number of assets increases*

♦ Key reason: (High) Dimensionality

$$\frac{SR(\text{plug-in})}{SR^*} \xrightarrow{P} \sqrt{\frac{1-\rho}{1+\rho/(SR^*)^2}} < \sqrt{1-\rho} < 1, \quad \text{as } \frac{N}{T} \to \rho \in (0,1)$$

Challenges for Large Portfolios

- Poor performance of the plug-in portfolio
  - "Markowitz optimization enigma": Michaud (1989)
  - Best and Grauer (1991), Chopra and Ziemba (1993), Kan and Zhou (2007) etc.
- The situation worsens *as the number of assets increases*

♦ Key reason: (High) Dimensionality

$$\frac{SR(\text{plug-in})}{SR^*} \xrightarrow{P} \sqrt{\frac{1-\rho}{1+\rho/(SR^*)^2}} < \sqrt{1-\rho} < 1, \quad \text{as } \frac{N}{T} \to \rho \in (0,1)$$

Alternative Methods

- Adjusting inputs
  - Regularized covariance matrix or its inverse:
    - shrinkage (Ledoit and Wolf (2004), Ledoit and Wolf (2017));
    - thresholding (Bickel and Levina (2008), Cai and Liu (2011)); CLIME (Cai, Liu and Luo (2011), Cai, Liu and Zhou (2016));
    - POET (Fan, Fan and Lv (2008), Fan, Liao and Mincheva (2013));
    - and many others...
  - Mean estimation: Black and Litterman (1991)
- Imposing constraints:
  - No-short-sale constraint (Jagannathan and Ma (2003));
  - gross-exposure/ℓ_1 constraint (Brodie, Daubechies, De Mol, Giannone and Loris (2009), Fan, Zhang and Yu (2012), Fan, Li and Yu (2012));
  - 2-norm-constrained minimum variance portfolio (DeMiguel, Garlappi, Nogales and Uppal (2009));
  - other non-convex constraints (Fastrich, Paterlini and Winker (2012))

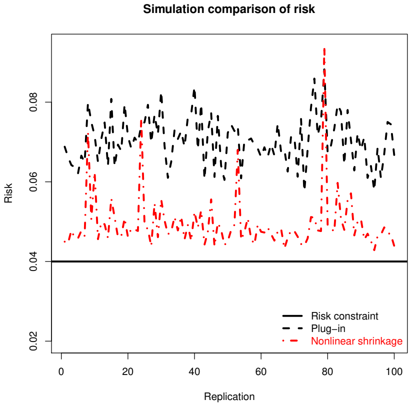
Alternative Methods

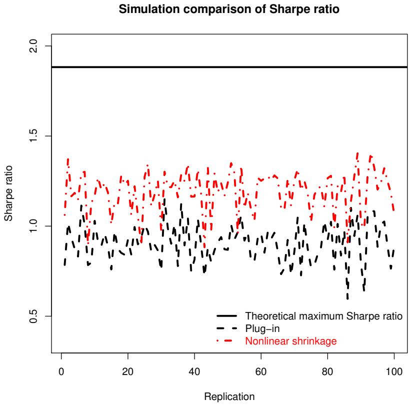
- Adjusting inputs
  - Regularized covariance matrix or its inverse:
    - shrinkage (Ledoit and Wolf (2004), Ledoit and Wolf (2017));
    - thresholding (Bickel and Levina (2008), Cai and Liu (2011)); CLIME (Cai, Liu and Luo (2011), Cai, Liu and Zhou (2016));
    - POET (Fan, Fan and Lv (2008), Fan, Liao and Mincheva (2013));
    - and many others...
  - Mean estimation: Black and Litterman (1991)
- Imposing constraints:
  - No-short-sale constraint (Jagannathan and Ma (2003));
  - gross-exposure/ℓ_1 constraint (Brodie, Daubechies, De Mol, Giannone and Loris (2009), Fan, Zhang and Yu (2012), Fan, Li and Yu (2012));
  - 2-norm-constrained minimum variance portfolio (DeMiguel, Garlappi, Nogales and Uppal (2009));
  - other non-convex constraints (Fastrich, Paterlini and Winker (2012))

Alternative Methods

- Adjusting inputs
  - Regularized covariance matrix or its inverse:
    - shrinkage (Ledoit and Wolf (2004), Ledoit and Wolf (2017));
    - thresholding (Bickel and Levina (2008), Cai and Liu (2011)); CLIME (Cai, Liu and Luo (2011), Cai, Liu and Zhou (2016));
    - POET (Fan, Fan and Lv (2008), Fan, Liao and Mincheva (2013));
    - and many others...
  - Mean estimation: Black and Litterman (1991)
- Imposing constraints:
  - No-short-sale constraint (Jagannathan and Ma (2003));
  - gross-exposure/ℓ_1 constraint (Brodie, Daubechies, De Mol, Giannone and Loris (2009), Fan, Zhang and Yu (2012), Fan, Li and Yu (2012));
  - 2-norm-constrained minimum variance portfolio (DeMiguel, Garlappi, Nogales and Uppal (2009));
  - other non-convex constraints (Fastrich, Paterlini and Winker (2012))

A Competitive Alternative: Nonlinear Shrinkage (Ledoit and Wolf (2017), RFS)





Two Objectives

1. Meet risk constraint
2. Attain the maximum Sharpe ratio

Q: Is it possible to achieve both objectives simultaneously?

Answer: Yes ! $\rightarrow$ MAXSER !

Two Objectives

1. Meet risk constraint
2. Attain the maximum Sharpe ratio

Q: Is it possible to achieve both objectives simultaneously?

Answer: Yes ! $\rightarrow$ MAXSER !

Two Objectives

1. Meet risk constraint
2. Attain the maximum Sharpe ratio

Q: Is it possible to achieve both objectives simultaneously?

Answer: Yes ! $\rightarrow$ MAXSER !

Two Objectives

1. Meet risk constraint
2. Attain the maximum Sharpe ratio

Q: Is it possible to achieve both objectives simultaneously?

Answer: Yes ! $\rightarrow$ MAXSER !

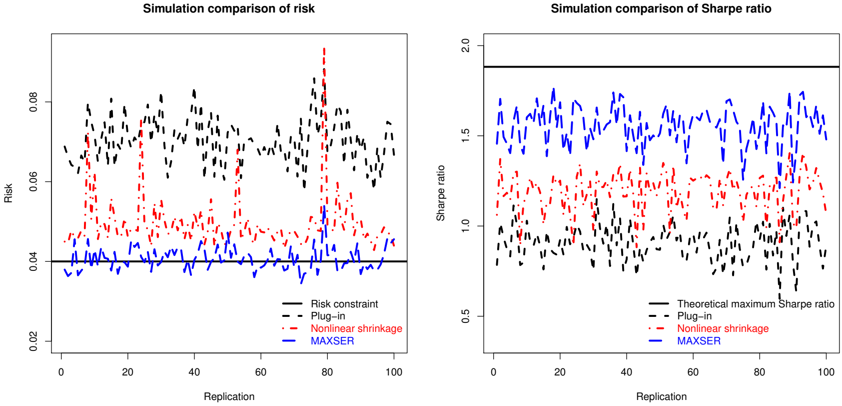
Two Objectives

1. Meet risk constraint
2. Attain the maximum Sharpe ratio

Q: Is it possible to achieve both objectives simultaneously?

Answer: Yes ! $\rightarrow$ MAXSER !

Our Portfolio: MAXSER



Our Contributions

- MAXSER
  - a bias-corrected *unconstrained* regression **equivalent to Markowitz**
  - consistent estimation of maximum Sharpe ratio
  - **consistency** of return & risk

$\rightarrow$ Approaches mean-variance efficiency for large portfolios!

Our Contributions

- MAXSER
  - a bias-corrected *unconstrained* regression **equivalent to Markowitz**
  - consistent estimation of maximum Sharpe ratio
  - **consistency** of return & risk

$\rightarrow$ Approaches mean-variance efficiency for large portfolios!

Our Contributions

- MAXSER
  - a bias-corrected *unconstrained* regression **equivalent to Markowitz**
  - consistent estimation of maximum Sharpe ratio
  - **consistency** of return & risk

$\rightarrow$ Approaches mean-variance efficiency for large portfolios!

Our Contributions

- MAXSER
  - a bias-corrected *unconstrained* regression **equivalent to Markowitz**
  - consistent estimation of maximum Sharpe ratio
  - **consistency** of return & risk

$\rightarrow$ Approaches mean-variance efficiency for large portfolios!

Our Contributions

- MAXSER
  - a bias-corrected *unconstrained* regression **equivalent to Markowitz**
  - consistent estimation of maximum Sharpe ratio
  - **consistency** of return & risk

$\rightarrow$ Approaches mean-variance efficiency for large portfolios!

Outline

Outline

Start From the Origin

- For a given level of risk constraint σ, the mean-variance optimization problem is

$$\max E(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\mu \quad \text{subject to} \quad \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\Sigma\boldsymbol{w} \leq \sigma^2. \tag{1}$$

- Denote by $\theta = \mu'\Sigma^{-1}\mu$ the squared maximum Sharpe ratio of the tangency portfolio, the dual form with return constraint $r^* = \sigma\sqrt{\theta}$ is

$$\min \boldsymbol{w}'\Sigma\boldsymbol{w} \quad \text{subject to} \quad \boldsymbol{w}'\mu = r^*. \tag{2}$$

- The optimal portfolio $\boldsymbol{w}^*$ admits

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\theta}}\Sigma^{-1}\mu. \tag{3}$$

Start From the Origin

- For a given level of risk constraint σ, the mean-variance optimization problem is

$$\max E(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\boldsymbol{\mu} \quad \text{subject to} \quad \text{Var}(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \leq \sigma^2. \tag{1}$$

- Denote by $\theta = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ the squared maximum Sharpe ratio of the tangency portfolio, the dual form with return constraint $r^* = \sigma\sqrt{\theta}$ is

$$\min \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \quad \text{subject to} \quad \boldsymbol{w}'\boldsymbol{\mu} = r^*. \tag{2}$$

- The optimal portfolio $\boldsymbol{w}^*$ admits

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\theta}}\Sigma^{-1}\mu. \tag{3}$$

Start From the Origin

- For a given level of risk constraint σ, the mean-variance optimization problem is

$$\max E(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\boldsymbol{\mu} \quad \text{subject to} \quad \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \leq \sigma^2. \tag{1}$$

- Denote by $\theta = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ the squared maximum Sharpe ratio of the tangency portfolio, the dual form with return constraint $r^* = \sigma\sqrt{\theta}$ is

$$\min \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} \quad \text{subject to} \quad \boldsymbol{w}'\boldsymbol{\mu} = r^*. \tag{2}$$

- The optimal portfolio $\boldsymbol{w}^*$ admits

$$\boldsymbol{w}^* = \frac{\sigma}{\sqrt{\theta}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}. \tag{3}$$

Existing Regression Formulations

- Constrained regression (e.g., Brodie, Daubechies, De Mol, Giannone and Loris (2009)):

$$\arg\min_{\boldsymbol{w}} E(r^* - \boldsymbol{w}'\boldsymbol{r})^2 \quad \text{subject to} \quad E(\boldsymbol{w}'\boldsymbol{r}) = r^* \text{ or } \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \sigma^2$$

$\rightarrow$ constraints have to be replaced with sample version, introducing errors/biases

- Britten-Jones (1999), arbitrary response (e.g. the number "1"):

$$\arg\min_{\boldsymbol{w}} E(1 - \boldsymbol{w}'\boldsymbol{r})^2$$

$\rightarrow$ yields a multiple of the suboptimal plug-in portfolio & needs a challenging scaling

Existing Regression Formulations

- Constrained regression (e.g., Brodie, Daubechies, De Mol, Giannone and Loris (2009)):

$$\arg\min_{\boldsymbol{w}} E(r^* - \boldsymbol{w}'\boldsymbol{r})^2 \quad \text{subject to} \quad E(\boldsymbol{w}'\boldsymbol{r}) = r^* \text{ or } \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \sigma^2$$

$\rightarrow$ constraints have to be replaced with sample version, introducing errors/biases

- Britten-Jones (1999), arbitrary response (e.g. the number “1”):

$$\arg\min_{\boldsymbol{w}} E(1 - \boldsymbol{w}'\boldsymbol{r})^2$$

$\rightarrow$ yields a multiple of the suboptimal plug-in portfolio & needs a challenging scaling

Existing Regression Formulations

- Constrained regression (e.g., Brodie, Daubechies, De Mol, Giannone and Loris (2009)):

$$\arg\min_{\boldsymbol{w}} E(r^* - \boldsymbol{w}'\boldsymbol{r})^2 \quad \text{subject to} \quad E(\boldsymbol{w}'\boldsymbol{r}) = r^* \text{ or } \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \sigma^2$$

$\rightarrow$ constraints have to be replaced with sample version, introducing errors/biases

- Britten-Jones (1999), arbitrary response (e.g. the number "1"):

$$\arg\min_{\boldsymbol{w}} E(1 - \boldsymbol{w}'\boldsymbol{r})^2$$

$\rightarrow$ yields a multiple of the suboptimal plug-in portfolio & needs a challenging scaling

Existing Regression Formulations

- Constrained regression (e.g., Brodie, Daubechies, De Mol, Giannone and Loris (2009)):

$$\arg\min_{\boldsymbol{w}} E(r^* - \boldsymbol{w}'\boldsymbol{r})^2 \quad \text{subject to} \quad E(\boldsymbol{w}'\boldsymbol{r}) = r^* \text{ or } \mathrm{Var}(\boldsymbol{w}'\boldsymbol{r}) = \sigma^2$$

  $\rightarrow$ constraints have to be replaced with sample version, introducing errors/biases

- Britten-Jones (1999), arbitrary response (e.g. the number “1”):

$$\arg\min_{\boldsymbol{w}} E(1 - \boldsymbol{w}'\boldsymbol{r})^2$$

  $\rightarrow$ yields a multiple of the suboptimal plug-in portfolio & needs a challenging scaling

Our Unconstrained Equivalent Regression Representation

Proposition 1

The unconstrained regression

$$\arg\min_{\boldsymbol{w}} E(r_c - \boldsymbol{w}'\boldsymbol{r})^2, \quad \textit{where} \quad r_c := \frac{1+\theta}{\theta} r^* \equiv \sigma \frac{1+\theta}{\sqrt{\theta}}, \tag{4}$$

is equivalent to the mean-variance optimization.

- **Unconstrained!**
- **Equivalent** to the mean-variance optimization!
- **Response r_c is crucial!**

Our Unconstrained Equivalent Regression Representation

Proposition 1

The unconstrained regression

$$\arg\min_{\boldsymbol{w}} E(r_c - \boldsymbol{w}'\boldsymbol{r})^2, \quad \textit{where} \quad r_c := \frac{1+\theta}{\theta} r^* \equiv \sigma \frac{1+\theta}{\sqrt{\theta}}, \tag{4}$$

is equivalent to the mean-variance optimization.

- **Unconstrained!**
- **Equivalent** to the mean-variance optimization!
- **Response r_c is crucial!**

Our Unconstrained Equivalent Regression Representation

Proposition 1

The unconstrained regression

$$\arg\min_{\boldsymbol{w}} E(r_c - \boldsymbol{w}'\boldsymbol{r})^2, \quad \textit{where} \quad r_c := \frac{1+\theta}{\theta} r^* \equiv \sigma \frac{1+\theta}{\sqrt{\theta}}, \tag{4}$$

is equivalent to the mean-variance optimization.

- **Unconstrained!**
- **Equivalent** to the mean-variance optimization!
- **Response r_c is crucial!**

Our Unconstrained Equivalent Regression Representation

Proposition 1

The unconstrained regression

$$\arg\min_{\boldsymbol{w}} E(r_c - \boldsymbol{w}'\boldsymbol{r})^2, \quad \textit{where} \quad r_c := \frac{1+\theta}{\theta} r^* \equiv \sigma \frac{1+\theta}{\sqrt{\theta}}, \tag{4}$$

is equivalent to the mean-variance optimization.

- **Unconstrained!**
- **Equivalent** to the mean-variance optimization!
- **Response r_c is crucial!**

Outline

High-dimensional Issues

- Proposition 1:

  MV optimization $\Rightarrow$ equivalent unconstrained regression
- Sample version in practice:

$$\arg\min_{\boldsymbol{w}} \frac{1}{T}\sum_{t=1}^{T}\left(r_c - \boldsymbol{w}'\boldsymbol{R_t}\right)^2,$$

  where $\boldsymbol{R_t} = (R_{t1}, \ldots, R_{tN})'$, $t = 1, \ldots, T$, are T i.i.d. copies of the return vector $\boldsymbol{r}$.
- In general it is *impossible* to consistently estimate the coefficients in a high-dimensional regression where $N/T = O(1)$

High-dimensional Issues

- Proposition 1:

  MV optimization $\Rightarrow$ equivalent unconstrained regression
- Sample version in practice:

$$\arg\min_{\boldsymbol{w}} \frac{1}{T}\sum_{t=1}^{T}\left(r_c - \boldsymbol{w}'\boldsymbol{R_t}\right)^2,$$

  where $\boldsymbol{R_t} = (R_{t1}, \dots, R_{tN})'$, $t = 1, \dots, T$, are T i.i.d. copies of the return vector $\boldsymbol{r}$.
- In general it is *impossible* to consistently estimate the coefficients in a high-dimensional regression where $N/T = O(1)$

High-dimensional Issues

- Proposition 1:

  MV optimization $\Rightarrow$ equivalent unconstrained regression
- Sample version in practice:

$$\arg\min_{\boldsymbol{w}} \frac{1}{T}\sum_{t=1}^{T}\left(r_c - \boldsymbol{w}'\boldsymbol{R_t}\right)^2,$$

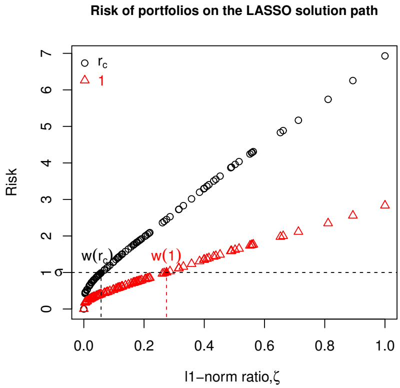
  where $\boldsymbol{R_t} = (R_{t1}, \ldots, R_{tN})'$, $t = 1, \ldots, T$, are T i.i.d. copies of the return vector $\boldsymbol{r}$.
- In general it is *impossible* to consistently estimate the coefficients in a high-dimensional regression where $N/T = O(1)$

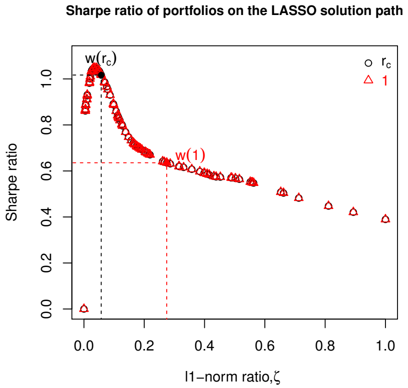
Sparse Regression

- We adopt the sparse regression technique LASSO:

$$\boldsymbol{w}(r_c) := \arg\min_{\boldsymbol{w}} \frac{1}{T}\sum_{t=1}^{T}\left(r_c - \boldsymbol{w}'\boldsymbol{R_t}\right)^2 \quad \text{subject to} \quad ||\boldsymbol{w}||_1 \leq \lambda$$

Importance of Using the Correct Response r_c





(ℓ_1-norm ratio: $\zeta = ||\boldsymbol{w}||_1/||\boldsymbol{w}_{ols}||_1$)

Outline

Estimator of the Maximum Sharpe Ratio and r_c

Proposition 2

Define the following estimators of θ:

$$\widehat{\theta} := \frac{(T-N-2)\widehat{\theta}_s - N}{T}, \tag{5}$$

where $\widehat{\theta}_s := \widehat{\boldsymbol{\mu}}'\widehat{\boldsymbol{\Sigma}}^{-1}\widehat{\boldsymbol{\mu}}$ is the sample estimate of θ. If $N/T \to \rho \in (0,1)$, under normality assumption we have

$$|\widehat{\theta} - \theta| \xrightarrow{P} 0.$$

Furthermore, our estimator of the response r_c is

$$\widehat{r_c} := \frac{1+\widehat{\theta}}{\sqrt{\widehat{\theta}}}, \tag{6}$$

which satisfies

$$|\widehat{r_c} - r_c| \xrightarrow{P} 0.$$

A LASSO-type Estimator

- Our estimator of $\boldsymbol{w}^*$:

$$\widehat{\boldsymbol{w}^*} = \arg\min_{\boldsymbol{w}} \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{r_c} - \boldsymbol{w}'\boldsymbol{R_t}\right)^2 \quad \text{subject to} \quad ||\boldsymbol{w}||_1 \leq \lambda. \qquad (7)$$

- $\widehat{\boldsymbol{w}^*}$ is our MAXimum - Sharpe ratio Estimated & sparse Regression (MAXSER) portfolio.

Main Result I: MAXSER Without Factor Structure

Theorem 1

Under normality and sparsity assumptions on the optimal portfolio, the MAXSER portfolio $\widehat{\boldsymbol{w}^*}$ defined in (7) with $\widehat{r_c}$ given by (6) satisfies that, as $N \to \infty$,

$$|\mu' \widehat{\boldsymbol{w}^*} - r^*| \xrightarrow{P} 0, \tag{8}$$

and

$$\left| \sqrt{\widehat{\boldsymbol{w}^*}' \Sigma \widehat{\boldsymbol{w}^*}} - \sigma \right| \xrightarrow{P} 0. \tag{9}$$

▲ The MAXSER asymptotically *achieves the maximum expected return* and meanwhile *satisfies the risk constraint*, therefore **approaches mean-variance efficiency**!

♦ **First method** ever that achieves both objectives for large portfolios

Main Result I: MAXSER Without Factor Structure

Theorem 1

Under normality and sparsity assumptions on the optimal portfolio, the MAXSER portfolio $\widehat{\mathbf{w}^*}$ defined in (7) with $\widehat{r_c}$ given by (6) satisfies that, as $N \to \infty$,

$$|\mu' \widehat{\mathbf{w}^*} - r^*| \xrightarrow{P} 0, \tag{8}$$

and

$$\left| \sqrt{\widehat{\mathbf{w}^*}' \Sigma \widehat{\mathbf{w}^*}} - \sigma \right| \xrightarrow{P} 0. \tag{9}$$

▲ The MAXSER asymptotically *achieves the maximum expected return* and meanwhile *satisfies the risk constraint*, therefore **approaches mean-variance efficiency**!

♦ **First method** ever that achieves both objectives for large portfolios

Main Result I: MAXSER Without Factor Structure

Theorem 1

Under normality and sparsity assumptions on the optimal portfolio, the MAXSER portfolio $\widehat{\mathbf{w}^*}$ *defined in* (7) *with* $\widehat{r_c}$ *given by* (6) *satisfies that, as* $N \to \infty$,

$$|\mu' \widehat{\mathbf{w}^*} - r^*| \xrightarrow{P} 0, \quad (8)$$

and

$$\left| \sqrt{\widehat{\mathbf{w}^*}' \Sigma \widehat{\mathbf{w}^*}} - \sigma \right| \xrightarrow{P} 0. \quad (9)$$

▲ The MAXSER asymptotically *achieves the maximum expected return* and meanwhile *satisfies the risk constraint*, therefore **approaches mean-variance efficiency**!

♦ **First method** ever that achieves both objectives for large portfolios

Main Result I: MAXSER Without Factor Structure

Theorem 1

Under normality and sparsity assumptions on the optimal portfolio, the MAXSER portfolio $\widehat{\mathbf{w}^*}$ *defined in* (7) *with* $\widehat{r_c}$ *given by* (6) *satisfies that, as* $N \to \infty$,

$$\left|\mu' \widehat{\mathbf{w}^*} - r^*\right| \xrightarrow{P} 0, \tag{8}$$

and

$$\left|\sqrt{\widehat{\mathbf{w}^*}' \Sigma \widehat{\mathbf{w}^*}} - \sigma\right| \xrightarrow{P} 0. \tag{9}$$

▲ The MAXSER asymptotically *achieves the maximum expected return* and meanwhile *satisfies the risk constraint*, therefore **approaches mean-variance efficiency**!

♦ **First method** ever that achieves both objectives for large portfolios

Main Result I: MAXSER Without Factor Structure

Theorem 1

Under normality and sparsity assumptions on the optimal portfolio, the MAXSER portfolio $\widehat{\mathbf{w}^*}$ *defined in* (7) *with* $\widehat{r_c}$ *given by* (6) *satisfies that, as* $N \to \infty$,

$$\left|\mu' \widehat{\mathbf{w}^*} - r^*\right| \xrightarrow{P} 0, \tag{8}$$

and

$$\left|\sqrt{\widehat{\mathbf{w}^*}' \Sigma \widehat{\mathbf{w}^*}} - \sigma\right| \xrightarrow{P} 0. \tag{9}$$

▲ The MAXSER asymptotically *achieves the maximum expected return* and meanwhile *satisfies the risk constraint*, therefore **approaches mean-variance efficiency**!

♦ **First method** ever that achieves both objectives for large portfolios

Outline

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation

- Consider the following model of returns:

$$r_i = \alpha_i + \sum_{j=1}^{K} \beta_{ij} f_j + e_i := \sum_{j=1}^{K} \beta_{ij} f_j + u_i, \qquad i = 1, \cdots, N,$$

- Special features of the model:
  - The K included factors need NOT to be the full set of factors
  - u_i's, the "idiosyncratic returns", are allowed to have factor structure
- Compact form:

$$\boldsymbol{r} = \beta \boldsymbol{f} + \boldsymbol{u}$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation

- Consider the following model of returns:

$$r_i = \alpha_i + \sum_{j=1}^{K} \beta_{ij} f_j + e_i := \sum_{j=1}^{K} \beta_{ij} f_j + u_i, \qquad i = 1, \cdots, N,$$

- Special features of the model:
  - The K included factors need NOT to be the full set of factors
  - u_i's, the "idiosyncratic returns", are allowed to have factor structure
- Compact form:

$$\boldsymbol{r} = \boldsymbol{\beta f} + \boldsymbol{u}$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation

- Consider the following model of returns:

$$r_i = \alpha_i + \sum_{j=1}^{K} \beta_{ij} f_j + e_i := \sum_{j=1}^{K} \beta_{ij} f_j + u_i, \qquad i = 1, \cdots, N,$$

- Special features of the model:
  - The K included factors need NOT to be the full set of factors
  - u_i's, the "idiosyncratic returns", are allowed to have factor structure
- Compact form:

$$\boldsymbol{r} = \beta \boldsymbol{f} + \boldsymbol{u}$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation

- Consider the following model of returns:

$$r_i = \alpha_i + \sum_{j=1}^{K} \beta_{ij} f_j + e_i := \sum_{j=1}^{K} \beta_{ij} f_j + u_i, \qquad i = 1, \cdots, N,$$

- Special features of the model:
    - The K included factors need NOT to be the full set of factors
    - u_i's, the "idiosyncratic returns", are allowed to have factor structure
- Compact form:

$$\boldsymbol{r} = \boldsymbol{\beta} \boldsymbol{f} + \boldsymbol{u}$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation, *ctd*

- We will invest in the N assets and the K factors
- Question: How to estimate the optimal portfolio weight $(w_1^f, \ldots, w_K^f; w_1, \ldots, w_N) := (\boldsymbol{w}_f, \boldsymbol{w})$

Proposition 3

For any given risk constraint level σ, the optimal portfolio $\boldsymbol{w}_{all} := (\boldsymbol{w}_f, \boldsymbol{w})$ is given by

$$\left(\sqrt{\frac{\theta_f}{\theta_{all}}}\sigma \boldsymbol{w}_f^* - \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma \beta' \boldsymbol{w}_u^*, \quad \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma \boldsymbol{w}_u^*\right),$$

where $\theta_f = \mu_f' \Sigma_f^{-1} \mu_f$, $\theta_u = \alpha' \Sigma_u^{-1} \alpha$, and $\theta_{all} = \mu_{all}' \Sigma_{all}^{-1} \mu_{all}$. $\boldsymbol{w}_f^*$ and $\boldsymbol{w}_u^*$ are optimal portfolio weights on factors and idiosyncratic components with one unit of risk:

$$\boldsymbol{w}_f^* = \frac{1}{\sqrt{\theta_f}} \Sigma_f^{-1} \mu_f, \quad \boldsymbol{w}_u^* = \frac{1}{\sqrt{\theta_u}} \Sigma_u^{-1} \alpha.$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation, *ctd*

- We will invest in the N assets and the K factors
- Question: How to estimate the optimal portfolio weight $(w_1^f, \ldots, w_K^f; w_1, \ldots, w_N) := (\boldsymbol{w}_f, \boldsymbol{w})$

Proposition 3

For any given risk constraint level σ, the optimal portfolio $\mathbf{w}_{all} := (\mathbf{w}_f, \mathbf{w})$ is given by

$$\left(\sqrt{\frac{\theta_f}{\theta_{all}}}\sigma \mathbf{w}_f^* - \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma\beta' \mathbf{w}_u^*, \quad \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma \mathbf{w}_u^*\right),$$

where $\theta_f = \mu_f'\Sigma_f^{-1}\mu_f$, $\theta_u = \alpha'\Sigma_u^{-1}\alpha$, and $\theta_{all} = \mu_{all}'\Sigma_{all}^{-1}\mu_{all}$. $\mathbf{w}_f^*$ and $\mathbf{w}_u^*$ are optimal portfolio weights on factors and idiosyncratic components with one unit of risk:

$$\mathbf{w}_f^* = \frac{1}{\sqrt{\theta_f}}\Sigma_f^{-1}\mu_f, \quad \mathbf{w}_u^* = \frac{1}{\sqrt{\theta_u}}\Sigma_u^{-1}\alpha.$$

The Optimal Portfolio: A Factor-Idiosyncratic Component Separation, *ctd*

- We will invest in the N assets and the K factors
- Question: How to estimate the optimal portfolio weight $(w_1^f, \ldots, w_K^f; w_1, \ldots, w_N) := (\boldsymbol{w}_f, \boldsymbol{w})$

Proposition 3

For any given risk constraint level σ, the optimal portfolio $\mathbf{w}_{all} := (\mathbf{w}_f, \mathbf{w})$ is given by

$$\left(\sqrt{\frac{\theta_f}{\theta_{all}}}\sigma \mathbf{w}_f^* - \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma \boldsymbol{\beta}' \mathbf{w}_u^*, \quad \sqrt{\frac{\theta_u}{\theta_{all}}}\sigma \mathbf{w}_u^*\right),$$

where $\theta_f = \boldsymbol{\mu}_f' \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\mu}_f$, $\theta_u = \boldsymbol{\alpha}' \boldsymbol{\Sigma}_u^{-1} \boldsymbol{\alpha}$, and $\theta_{all} = \boldsymbol{\mu}_{all}' \boldsymbol{\Sigma}_{all}^{-1} \boldsymbol{\mu}_{all}$. $\mathbf{w}_f^*$ and $\mathbf{w}_u^*$ are optimal portfolio weights on factors and idiosyncratic components with one unit of risk:

$$\mathbf{w}_f^* = \frac{1}{\sqrt{\theta_f}} \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\mu}_f, \quad \mathbf{w}_u^* = \frac{1}{\sqrt{\theta_u}} \boldsymbol{\Sigma}_u^{-1} \boldsymbol{\alpha}.$$

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    △ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    △ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    $\triangle$ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    $\triangle$ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    $\triangle$ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    $\triangle$ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    △ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    △ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    $\triangle$ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    $\triangle$ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Tasks & Challenges

- To estimate the optimal portfolio $\boldsymbol{w}_{all}$, we need to estimate
  - θ_f & $\boldsymbol{w}_f^*$

    △ low-dimensional nature

    $\Rightarrow$ the standard plug-in estimators work
  - θ_u & $\boldsymbol{w}_u^*$

    △ **high-dimensional** nature !

    $\Rightarrow$ main challenges

Estimator of Response r_c

- Based on the factor model, we have

$$\theta_{all} = \theta_f + \theta_u$$

- θ_u can be consistently estimated by $\widehat{\theta}_u := \widehat{\theta}_{all} - \widehat{\theta}_f$, where $\widehat{\theta}_{all}$ and $\widehat{\theta}_f$ are computed by applying (5) to all assets and factors
- Estimator of the response r_c: $\widehat{r_c} := (1 + \widehat{\theta}_u)/\sqrt{\widehat{\theta}_u}$

Estimator of Response r_c

- Based on the factor model, we have

$$\theta_{all} = \theta_f + \theta_u$$

- θ_u can be consistently estimated by $\widehat{\theta}_u := \widehat{\theta}_{all} - \widehat{\theta}_f$, where $\widehat{\theta}_{all}$ and $\widehat{\theta}_f$ are computed by applying (5) to all assets and factors
- Estimator of the response r_c: $\widehat{r}_c := (1 + \widehat{\theta}_u)/\sqrt{\widehat{\theta}_u}$

Estimator of Response r_c

- Based on the factor model, we have

$$\theta_{all} = \theta_f + \theta_u$$

- θ_u can be consistently estimated by $\widehat{\theta}_u := \widehat{\theta}_{all} - \widehat{\theta}_f$, where $\widehat{\theta}_{all}$ and $\widehat{\theta}_f$ are computed by applying (5) to all assets and factors
- Estimator of the response r_c: $\widehat{r_c} := (1 + \widehat{\theta}_u)/\sqrt{\widehat{\theta}_u}$

The MAXSER Portfolio

- Plug-in estimator of $\boldsymbol{w}_f^*$: $\widehat{\boldsymbol{w}_f^*} := \frac{1}{\sqrt{\widehat{\theta}_f}} \widehat{\boldsymbol{\Sigma}}_f^{-1} \widehat{\boldsymbol{\mu}}_f$
- Estimator of $\boldsymbol{w}_u^*$:

$$\widehat{\boldsymbol{w}_u^*} = \arg\min_{\boldsymbol{w}} \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{r_c} - \boldsymbol{w}' \widehat{\boldsymbol{U}_t} \right)^2 \quad \text{subject to} \quad ||\boldsymbol{w}||_1 \leq \lambda$$

- Final estimator of the optimal portfolio $\boldsymbol{w}_{all}$:

$$\widehat{\boldsymbol{w}_{all}} := (\widehat{\boldsymbol{w}_f}, \widehat{\boldsymbol{w}}) = \left(\sigma \sqrt{\frac{\widehat{\theta}_f}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_f^*} - \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\beta}' \widehat{\boldsymbol{w}_u^*}, \ \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_u^*} \right).$$

The MAXSER Portfolio

- Plug-in estimator of $\boldsymbol{w}_f^*$: $\widehat{\boldsymbol{w}_f^*} := \frac{1}{\sqrt{\widehat{\theta}_f}} \widehat{\boldsymbol{\Sigma}}_f^{-1} \widehat{\boldsymbol{\mu}}_f$
- Estimator of $\boldsymbol{w}_u^*$:

$$\widehat{\boldsymbol{w}_u^*} = \arg\min_{\boldsymbol{w}} \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{r_c} - \boldsymbol{w}' \widehat{\boldsymbol{U_t}} \right)^2 \quad \text{subject to} \quad ||\boldsymbol{w}||_1 \leq \lambda$$

- Final estimator of the optimal portfolio $\boldsymbol{w}_{all}$:

$$\widehat{\boldsymbol{w}_{all}} := (\widehat{\boldsymbol{w}_f}, \widehat{\boldsymbol{w}}) = \left(\sigma \sqrt{\frac{\widehat{\theta}_f}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_f^*} - \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\beta}' \widehat{\boldsymbol{w}_u^*}, \ \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_u^*} \right).$$

The MAXSER Portfolio

- Plug-in estimator of $\boldsymbol{w}_f^*$: $\widehat{\boldsymbol{w}_f^*} := \frac{1}{\sqrt{\widehat{\theta}_f}} \widehat{\boldsymbol{\Sigma}}_f^{-1} \widehat{\boldsymbol{\mu}}_f$
- Estimator of $\boldsymbol{w}_u^*$:

$$\widehat{\boldsymbol{w}_u^*} = \arg\min_{\boldsymbol{w}} \frac{1}{T} \sum_{t=1}^{T} \left(\widehat{r_c} - \boldsymbol{w}' \widehat{\boldsymbol{U_t}} \right)^2 \quad \text{subject to} \quad ||\boldsymbol{w}||_1 \leq \lambda$$

- Final estimator of the optimal portfolio $\boldsymbol{w}_{all}$:

$$\widehat{\boldsymbol{w}_{all}} := (\widehat{\boldsymbol{w}_f}, \widehat{\boldsymbol{w}}) = \left(\sigma \sqrt{\frac{\widehat{\theta}_f}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_f^*} - \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{\beta}}' \widehat{\boldsymbol{w}_u^*},\ \sigma \sqrt{\frac{\widehat{\theta}_u}{\widehat{\theta}_{all}}} \widehat{\boldsymbol{w}_u^*} \right).$$

Main Result II: MAXSER with Factor Investing

Theorem 2

Under normality assumption on returns and a mild sparsity assumption on $\mathbf{w}_u^*$, as $N \to \infty$, the MAXSER portfolio $\widehat{\mathbf{w}_{all}}$ satisfies

$$|\widehat{\mathbf{w}_{all}}'\boldsymbol{\mu}_{all} - r^*| \xrightarrow{P} 0, \quad \text{and} \quad |\widehat{\mathbf{w}_{all}}'\boldsymbol{\Sigma}_{all}\widehat{\mathbf{w}_{all}} - \sigma^2| \xrightarrow{P} 0, \qquad (10)$$

where $r^* = \mathbf{w}_{all}'\boldsymbol{\mu}_{all}$ is the maximum expected return at risk level σ.

Outline

Simulation Setup

- Monthly returns simulated from a three-factor model (parameters calibrated from real data)
- 1,000 replications
- Sample size $T = 120/240$, 100 stocks + 3 factors
- Compare portfolio risk and Sharpe ratio

Simulation Setup

- Monthly returns simulated from a three-factor model (parameters calibrated from real data)
- 1,000 replications
- Sample size $T = 120/240$, 100 stocks + 3 factors
- Compare portfolio risk and Sharpe ratio

Simulation Setup

- Monthly returns simulated from a three-factor model (parameters calibrated from real data)
- 1,000 replications
- Sample size $T = 120/240$, 100 stocks + 3 factors
- Compare portfolio risk and Sharpe ratio

Simulation Setup

- Monthly returns simulated from a three-factor model (parameters calibrated from real data)
- 1,000 replications
- Sample size $T = 120/240$, 100 stocks + 3 factors
- Compare portfolio risk and Sharpe ratio

Portfolios Under Comparison

Portfolio	Abbreviation
Plug-in MV on factors	Factor
Three-fund portfolio by Kan and Zhou (2007)	KZ
MV/GMV with different covariance matrix estimates	
MV with sample cov	MV-P
MV with linear shrinkage cov	MV-LS
MV with nonlinear shrinkage cov	MV-NLS
MV with nonlinear shrinkage cov adjusted for factor models	MV-NLSF
GMV with linear shrinkage cov	GMV-LS
GMV with nonlinear shrinkage cov	GMV-NLS
MV with short-sale constraint & cross-validation	
MV with sample cov & short-sale-CV	MV-P-SSCV
MV with linear shrinkage cov & short-sale-CV	MV-LS-SSCV
MV with nonlinear shrinkage cov & short-sale-CV	MV-NLS-SSCV
MV with ℓ_1-norm constraint & cross-validation	
MV with sample cov & ℓ_1-CV	MV-P-L1CV
MV with linear shrinkage cov & ℓ_1-CV	MV-LS-L1CV
MV with nonlinear shrinkage cov & ℓ_1-CV	MV-NLS-L1CV

Simulation Results: Normal Distribution, $T = 120$

Normal Distribution	$\sigma = 0.04, \ SR^* = 1.882$	$T = 120$
Portfolio	Risk	Sharpe Ratio
Factor	0.041 (0.003)	0.401 (0.169)
KZ	0.052 (0.040)	0.329 (0.184)
MAXSER	0.043 (0.005)	**1.083** (0.302)
MV/GMV with different covariance matrix estimates		
MV-P	0.296 (0.072)	0.367 (0.168)
MV-LS	0.082 (0.006)	0.697 (0.160)
MV-NLS	0.054 (0.017)	0.945 (0.183)
MV-NLSF	0.044 (0.002)	0.837 (0.139)
GMV-LS	0.013 (0.001)	0.438 (0.132)
GMV-NLS	0.015 (0.003)	0.553 (0.148)
MV with short-sale constraint & cross-validation		
MV-P-SSCV	0.057 (0.035)	0.400 (0.112)
MV-LS-SSCV	0.039 (0.025)	0.666 (0.177)
MV-NLS-SSCV	0.035 (0.023)	0.850 (0.259)
MV with ℓ_1-norm constraint & cross-validation		
MV-P-L1CV	0.041 (0.011)	0.539 (0.215)
MV-LS-L1CV	0.032 (0.012)	0.726 (0.179)
MV-NLS-L1CV	0.029 (0.011)	0.973 (0.171)

Simulation Results: Normal Distribution, $T = 240$

Normal Distribution	$\sigma = 0.04,\ SR^* = 1.882$	$T = 240$
Portfolio	Risk	Sharpe Ratio
Factor	0.041 (0.002)	0.467 (0.108)
KZ	0.091 (0.031)	0.909 (0.130)
MAXSER	0.041 (0.003)	**1.422** (0.200)
MV/GMV with different covariance matrix estimates		
MV-P	0.070 (0.005)	0.911 (0.123)
MV-LS	0.061 (0.004)	0.943 (0.117)
MV-NLS	0.049 (0.004)	1.199 (0.117)
MV-NLSF	0.042 (0.001)	1.068 (0.104)
GMV-LS	0.009 (0.000)	0.450 (0.102)
GMV-NLS	0.009 (0.001)	0.539 (0.167)
MV with short-sale constraint & cross-validation		
MV-P-SSCV	0.038 (0.008)	0.754 (0.259)
MV-LS-SSCV	0.038 (0.008)	0.744 (0.275)
MV-NLS-SSCV	0.038 (0.010)	0.847 (0.396)
MV with ℓ_1-norm constraint & cross-validation		
MV-P-L1CV	0.036 (0.006)	1.057 (0.185)
MV-LS-L1CV	0.036 (0.005)	1.121 (0.177)
MV-NLS-L1CV	0.037 (0.005)	1.207 (0.154)

Simulation Results: Heavy-tailed Distribution, $T = 120$

$t(6)$ Distribution	$\sigma = 0.04,\ SR^* = 1.882$	$T = 120$
Portfolio	Risk	Sharpe Ratio
Factor	0.034 (0.003)	0.350 (0.202)
KZ	0.039 (0.031)	0.288 (0.191)
MAXSER	0.035 (0.005)	**0.913** (0.327)
MV/GMV with different covariance matrix estimates		
MV-P	0.246 (0.060)	0.321 (0.174)
MV-LS	0.062 (0.005)	0.635 (0.169)
MV-NLS	0.042 (0.009)	0.845 (0.179)
MV-NLSF	0.036 (0.002)	0.716 (0.150)
GMV-LS	0.013 (0.001)	0.459 (0.130)
GMV-NLS	0.014 (0.003)	0.572 (0.125)
MV with short-sale constraint & cross-validation		
MV-P-SSCV	0.045 (0.033)	0.372 (0.102)
MV-LS-SSCV	0.031 (0.020)	0.609 (0.175)
MV-NLS-SSCV	0.028 (0.018)	0.764 (0.232)
MV with ℓ_1-norm constraint & cross-validation		
MV-P-L1CV	0.034 (0.009)	0.456 (0.202)
MV-LS-L1CV	0.025 (0.010)	0.661 (0.186)
MV-NLS-L1CV	0.023 (0.009)	0.860 (0.181)

Simulation Results: Heavy-tailed Distribution, $T = 240$

$t(6)$ Distribution	$\sigma = 0.04,\ SR^* = 1.882$	$T = 240$
Portfolio	Risk	Sharpe Ratio
Factor	0.033 (0.002)	0.427 (0.141)
KZ	0.059 (0.023)	0.802 (0.154)
MAXSER	0.034 (0.003)	**1.281** (0.243)
MV/GMV with different covariance matrix estimates		
MV-P	0.058 (0.004)	0.807 (0.140)
MV-LS	0.048 (0.003)	0.847 (0.133)
MV-NLS	0.040 (0.004)	1.071 (0.138)
MV-NLSF	0.034 (0.001)	0.931 (0.117)
GMV-LS	0.010 (0.000)	0.469 (0.107)
GMV-NLS	0.010 (0.001)	0.538 (0.182)
MV with short-sale constraint & cross-validation		
MV-P-SSCV	0.030 (0.008)	0.566 (0.227)
MV-LS-SSCV	0.030 (0.008)	0.551 (0.223)
MV-NLS-SSCV	0.031 (0.009)	0.575 (0.293)
MV with ℓ_1-norm constraint & cross-validation		
MV-P-L1CV	0.028 (0.005)	0.980 (0.195)
MV-LS-L1CV	0.028 (0.005)	1.044 (0.179)
MV-NLS-L1CV	0.028 (0.005)	1.102 (0.173)

Outline

Data & Rolling-window Scheme

- Two asset universes
  - DJIA 30 constituents and Fama-French three factors
  - S&P 500 constituents and Fama-French three factors
- Rolling-window scheme
  - monthly rolling and rebalancing
  - risk constraint fixed to be the standard deviation of the index during the first training period
- Stock pool determination
  - DJIA 30: all constituents at each time of portfolio construction, updated monthly
  - S&P 500: yearly updated stock pools consisting of 100 randomly picked constituents

Data & Rolling-window Scheme

- Two asset universes
  - DJIA 30 constituents and Fama-French three factors
  - S&P 500 constituents and Fama-French three factors
- Rolling-window scheme
  - monthly rolling and rebalancing
  - risk constraint fixed to be the standard deviation of the index during the first training period
- Stock pool determination
  - DJIA 30: all constituents at each time of portfolio construction, updated monthly
  - S&P 500: yearly updated stock pools consisting of 100 randomly picked constituents

Data & Rolling-window Scheme

- Two asset universes
  - DJIA 30 constituents and Fama-French three factors
  - S&P 500 constituents and Fama-French three factors
- Rolling-window scheme
  - monthly rolling and rebalancing
  - risk constraint fixed to be the standard deviation of the index during the first training period
- Stock pool determination
  - DJIA 30: all constituents at each time of portfolio construction, updated monthly
  - S&P 500: yearly updated stock pools consisting of 100 randomly picked constituents

Data & Rolling-window Scheme

- Two asset universes
  - DJIA 30 constituents and Fama-French three factors
  - S&P 500 constituents and Fama-French three factors
- Rolling-window scheme
  - monthly rolling and rebalancing
  - risk constraint fixed to be the standard deviation of the index during the first training period
- Stock pool determination
  - DJIA 30: all constituents at each time of portfolio construction, updated monthly
  - S&P 500: yearly updated stock pools consisting of 100 randomly picked constituents

Data & Rolling-window Scheme

- Two asset universes
  - DJIA 30 constituents and Fama-French three factors
  - S&P 500 constituents and Fama-French three factors
- Rolling-window scheme
  - monthly rolling and rebalancing
  - risk constraint fixed to be the standard deviation of the index during the first training period
- Stock pool determination
  - DJIA 30: all constituents at each time of portfolio construction, updated monthly
  - S&P 500: yearly updated stock pools consisting of 100 randomly picked constituents

Compared Portfolios and Performance Measure

- Additional compared portfolios:
  - Index
  - The equally weighted portfolio (the “1/N” rule)
- We compare the risk and Sharpe ratio, and further perform test about Sharpe ratio
  - Test

$$H_0 : SR_{MAXSER} \leqslant SR_0 \quad vs \quad H_a : SR_{MAXSER} > SR_0,$$

where SR_{MAXSER} is the Sharpe ratio of MAXSER portfolio, and SR_0 is the Sharpe ratio of one of the compared portfolios

Compared Portfolios and Performance Measure

- Additional compared portfolios:
  - Index
  - The equally weighted portfolio (the “1/N” rule)
- We compare the risk and Sharpe ratio, and further perform test about Sharpe ratio
  - Test

$$H_0 : SR_{MAXSER} \leqslant SR_0 \quad vs \quad H_a : SR_{MAXSER} > SR_0,$$

where SR_{MAXSER} is the Sharpe ratio of MAXSER portfolio, and SR_0 is the Sharpe ratio of one of the compared portfolios

DJIA Constituents & FF3

DJIA 30 Constituents & FF3 (*Without Transaction Costs*)			$T = 60$			$\sigma = 0.05$
Period	1977–2016			1997–2016		
Portfolio	Risk	Sharpe Ratio	p-value	Risk	Sharpe Ratio	p-value
Index	0.043	0.270	0.000	0.043	0.310	0.001
Equally weighted	0.042	0.328	0.000	0.044	0.307	0.001
Factor	0.055	0.427	0.000	0.058	0.254	0.000
KZ	0.104	0.250	0.000	0.097	0.265	0.000
MAXSER	0.060	**0.556**	–	0.064	**0.567**	–
MV-P	0.116	0.196	0.000	0.132	0.292	0.000
MV-LS	0.070	0.132	0.000	0.077	0.376	0.003
MV-NLS	0.068	0.166	0.000	0.073	0.352	0.001
MV-NLSF	0.067	0.232	0.000	0.070	0.290	0.000
GMV-LS	0.016	0.453	0.030	0.018	0.307	0.000
GMV-NLS	0.016	0.364	0.000	0.018	0.274	0.000
MV-P-SSCV	0.045	0.407	0.001	0.045	0.448	0.042
MV-LS-SSCV	0.044	0.376	0.000	0.045	0.469	0.070
MV-NLS-SSCV	0.044	0.443	0.005	0.044	0.473	0.072
MV-P-L1CV	0.043	0.136	0.000	0.043	0.253	0.000
MV-LS-L1CV	0.041	0.102	0.000	0.040	0.366	0.002
MV-NLS-L1CV	0.040	0.131	0.000	0.038	0.317	0.000

S&P 500 Constituents & FF3

S&P 500 Constituents & FF3 (*Without Transaction Costs*)			$T = 120$		$\sigma = 0.04$	
Period	1977–2016			1997–2016		
Portfolio	Risk	Sharpe Ratio	p-value	Risk	Sharpe Ratio	p-value
Index	0.043	0.279	0.000	0.044	0.302	0.000
Equally weighted	0.047	0.332	0.000	0.049	0.344	0.001
Factor	0.040	0.517	0.002	0.045	0.409	0.005
KZ	0.081	0.369	0.000	0.087	0.331	0.001
MAXSER	0.047	**0.667**	–	0.053	**0.591**	–
MV-P	0.347	0.383	0.000	0.367	0.257	0.000
MV-LS	0.079	0.248	0.000	0.078	0.093	0.000
MV-NLS	0.061	0.232	0.000	0.064	0.091	0.000
MV-NLSF	0.054	0.348	0.000	0.057	0.141	0.000
GMV-LS	0.022	0.277	0.000	0.025	0.436	0.027
GMV-NLS	0.025	0.271	0.000	0.027	0.467	0.063
MV-P-SSCV	0.061	0.347	0.000	0.067	0.316	0.000
MV-LS-SSCV	0.054	0.120	0.000	0.058	0.157	0.000
MV-NLS-SSCV	0.054	0.096	0.000	0.058	0.139	0.000
MV-P-L1CV	0.047	0.318	0.000	0.047	0.128	0.000
MV-LS-L1CV	0.044	0.047	0.000	0.048	−0.053	0.000
MV-NLS-L1CV	0.043	0.060	0.000	0.048	−0.059	0.000

What about transaction costs?

- The portfolio return net of transaction cost in period t, $r_{net}(t)$ is calculated by

$$r_{net}(t) = \left(1 - \sum_j c_{t,j}|w_j(t+1) - w_j(t+)|\right)(1 + r(t)) - 1,$$

  - $c_{t,j}$: a cost level that measures transaction cost per dollar traded for trading asset j
  - $w_j(t+1)$: weight on asset j at the beginning of period $t+1$
  - $w_j(t+)$: weight of asset j at the end of period t
  - $r(t)$: portfolio return without transaction cost in period t
- Based on Brandt et al. (2009) and Engle et al. (2012), we set $c_{t,j}$ to be time-varying, different for individual stock and factor portfolio

What about transaction costs?

- The portfolio return net of transaction cost in period t, $r_{net}(t)$ is calculated by

$$r_{net}(t) = \left(1 - \sum_j c_{t,j} |w_j(t+1) - w_j(t+)|\right)(1 + r(t)) - 1,$$

  - $c_{t,j}$: a cost level that measures transaction cost per dollar traded for trading asset j
  - $w_j(t+1)$: weight on asset j at the beginning of period $t+1$
  - $w_j(t+)$: weight of asset j at the end of period t
  - $r(t)$: portfolio return without transaction cost in period t
- Based on Brandt et al. (2009) and Engle et al. (2012), we set $c_{t,j}$ to be time-varying, different for individual stock and factor portfolio

What about transaction costs?

- The portfolio return net of transaction cost in period t, $r_{net}(t)$ is calculated by

$$r_{net}(t) = \left(1 - \sum_j c_{t,j} |w_j(t+1) - w_j(t+)|\right)(1 + r(t)) - 1,$$

  - $c_{t,j}$: a cost level that measures transaction cost per dollar traded for trading asset j
  - $w_j(t+1)$: weight on asset j at the beginning of period $t+1$
  - $w_j(t+)$: weight of asset j at the end of period t
  - $r(t)$: portfolio return without transaction cost in period t
- Based on Brandt et al. (2009) and Engle et al. (2012), we set $c_{t,j}$ to be time-varying, different for individual stock and factor portfolio

What about transaction costs?

- The portfolio return net of transaction cost in period t, $r_{net}(t)$ is calculated by

$$r_{net}(t) = \left(1 - \sum_j c_{t,j}|w_j(t+1) - w_j(t+)|\right)(1 + r(t)) - 1,$$

  - $c_{t,j}$: a cost level that measures transaction cost per dollar traded for trading asset j
  - $w_j(t+1)$: weight on asset j at the beginning of period $t+1$
  - $w_j(t+)$: weight of asset j at the end of period t
  - $r(t)$: portfolio return without transaction cost in period t
- Based on Brandt et al. (2009) and Engle et al. (2012), we set $c_{t,j}$ to be time-varying, different for individual stock and factor portfolio

What about transaction costs?

- The portfolio return net of transaction cost in period t, $r_{net}(t)$ is calculated by

$$r_{net}(t) = \left(1 - \sum_j c_{t,j} |w_j(t+1) - w_j(t+)|\right)(1 + r(t)) - 1,$$

  - $c_{t,j}$: a cost level that measures transaction cost per dollar traded for trading asset j
  - $w_j(t+1)$: weight on asset j at the beginning of period $t+1$
  - $w_j(t+)$: weight of asset j at the end of period t
  - $r(t)$: portfolio return without transaction cost in period t
- Based on Brandt et al. (2009) and Engle et al. (2012), we set $c_{t,j}$ to be time-varying, different for individual stock and factor portfolio

DJIA constituents & FF3, transaction costs considered

DJIA 30 Constituents & FF3 (*With Transaction Costs*)				$T = 60$		$\sigma = 0.05$
Period	1977–2016			1997–2016		
Portfolio	Risk	Sharpe Ratio	*p*-value	Risk	Sharpe Ratio	*p*-value
Index[1]	0.043	0.270	0.002	0.043	0.310	0.011
Equally weighted	0.042	0.317	0.724	0.044	0.300	0.108
Factor	0.055	0.265	0.273	0.058	0.146	0.000
KZ	0.108	−0.134	0.000	0.098	0.040	0.000
MAXSER	0.061	0.284	–	0.064	**0.402**	–
MV-P	0.117	−0.073	0.000	0.132	0.101	0.000
MV-LS	0.071	−0.014	0.000	0.077	0.299	0.067
MV-NLS	0.069	−0.077	0.000	0.073	0.213	0.002
MV-NLSF	0.067	0.045	0.000	0.070	0.187	0.000
GMV-LS	0.016	0.313	0.716	0.018	0.213	0.005
GMV-NLS	0.017	0.079	0.000	0.018	0.066	0.000
MV-P-SSCV	0.046	−0.258	0.000	0.045	0.147	0.000
MV-LS-SSCV	0.045	−0.042	0.000	0.045	0.323	0.105
MV-NLS-SSCV	0.045	−0.099	0.000	0.044	0.299	0.047
MV-P-L1CV	0.044	−0.350	0.000	0.043	0.011	0.000
MV-LS-L1CV	0.042	−0.127	0.000	0.040	0.281	0.036
MV-NLS-L1CV	0.041	−0.232	0.000	0.038	0.172	0.000

S&P 500 & FF3, transaction costs considered

S&P 500 Constituents & FF3 (*With Transaction Costs*)				$T = 120$		$\sigma = 0.04$
Period	1977–2016			1997–2016		
Portfolio	Risk	Sharpe Ratio	p-value	Risk	Sharpe Ratio	p-value
Index[2]	0.043	0.279	0.003	0.044	0.302	0.012
Equally weighted	0.047	0.307	0.012	0.049	0.330	0.030
Factor	0.040	0.408	0.228	0.045	0.330	0.013
KZ	0.082	0.009	0.000	0.087	0.160	0.000
MAXSER	0.048	**0.445**	–	0.053	**0.483**	–
MV-P	0.349	−0.185	0.000	0.357	−0.018	0.000
MV-LS	0.079	0.066	0.000	0.078	0.011	0.000
MV-NLS	0.061	0.099	0.000	0.064	0.022	0.000
MV-NLSF	0.054	0.175	0.000	0.057	0.054	0.000
GMV-LS	0.022	0.104	0.000	0.025	0.350	0.044
GMV-NLS	0.025	0.142	0.000	0.027	0.398	0.139
MV-P-SSCV	0.062	0.059	0.000	0.068	0.174	0.000
MV-LS-SSCV	0.054	−0.043	0.000	0.059	0.083	0.000
MV-NLS-SSCV	0.054	−0.028	0.000	0.058	0.075	0.000
MV-P-L1CV	0.047	0.040	0.000	0.047	−0.013	0.000
MV-LS-L1CV	0.044	−0.101	0.000	0.048	−0.112	0.000
MV-NLS-L1CV	0.043	−0.059	0.000	0.048	−0.110	0.000

Outline

Summary

- MAXSER asymptotically achieves the maximum Sharpe ratio and meanwhile satisfies the risk constraint
- First method ever that achieves both objectives
- Outstanding performance confirmed by comprehensive simulation and empirical studies

Summary

- MAXSER asymptotically achieves the maximum Sharpe ratio and meanwhile satisfies the risk constraint
- First method ever that achieves both objectives
- Outstanding performance confirmed by comprehensive simulation and empirical studies

Summary

- MAXSER asymptotically achieves the maximum Sharpe ratio and meanwhile satisfies the risk constraint
- First method ever that achieves both objectives
- Outstanding performance confirmed by comprehensive simulation and empirical studies

Summary

- MAXSER asymptotically achieves the maximum Sharpe ratio and meanwhile satisfies the risk constraint
- First method ever that achieves both objectives
- Outstanding performance confirmed by comprehensive simulation and empirical studies

Thank you!

Ao, M., Li, Y., and Zheng, X. “Approaching Mean-Variance Efficiency for Large Portfolios” (2017).

Bai, Z., Liu, H., and Wong, W.-K. “Enhancement of the applicability of Markowitz’s portfolio optimization by utilizing random matrix theory.” *Mathematical Finance*, 19(4):639–667 (2009).

Best, M. J. and Grauer, R. R. “On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results.” *Review of Financial Studies*, 4(2):315–342 (1991).

Bickel, P. J. and Levina, E. “Covariance regularization by thresholding.” *Ann. Statist.*, 36(6):2577–2604 (2008).

Black, F. and Litterman, R. B. “Asset allocation: combining investor views with market equilibrium.” *The Journal of Fixed Income*, 1(2):7–18 (1991).

Brandt, M. W., Santa-Clara, P., and Valkanov, R. “Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns.” *The Review of Financial Studies*, 22(9):3411–3447 (2009).

Britten-Jones, M. "The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights." *The Journal of Finance*, 54(2):655–671 (1999).

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., and Loris, I. "Sparse and stable Markowitz portfolios." *Proceedings of the National Academy of Sciences*, 106(30):12267–12272 (2009).

Cai, T. and Liu, W. "A direct estimation approach to sparse linear discriminant analysis." *J. Amer. Statist. Assoc.*, 106(496):1566–1577 (2011).

Cai, T. T., Liu, W., and Luo, X. "A Constrained ℓ_1 Minimization Approach to Sparse Precision Matrix Estimation." *Journal of the American Statistical Association*, 106:594–607 (2011).

Cai, T. T., Liu, W., and Zhou, H. H. "Estimating sparse precision matrix: optimal rates of convergence and adaptive estimation." *Ann. Statist.*, 44(2):455–488 (2016).

Chopra, V. K. and Ziemba, W. T. "The effect of errors in means, variances, and covariances on optimal portfolio choice." *The Journal of Portfolio Management*, 19(2):6–11 (1993).

DeMiguel, V., Garlappi, L., Nogales, F. J., and Uppal, R. "A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms." *Management Science*, 55(5):798–812 (2009).

El Karoui, N. "High-dimensionality effects in the Markowitz problem and other quadratic programs with linear constraints: Risk underestimation." *The Annals of Statistics*, 38(6):3487–3566 (2010).

Engle, R., Ferstenberg, R., and Russell, J. "Measuring and Modeling Execution Cost and Risk." *The Journal of Portfolio Management*, 38(2):14–28 (2012).

Fan, J., Fan, Y., and Lv, J. "High dimensional covariance matrix estimation using a factor model." *J. Econometrics*, 147(1):186–197 (2008).

Fan, J. and Li, R. "Variable selection via nonconcave penalized likelihood and its oracle properties." *J. Amer. Statist. Assoc.*, 96(456):1348–1360 (2001).

Fan, J., Zhang, J., and Yu, K. "Vast portfolio selection with gross-exposure constraints." *Journal of the American Statistical Association*, 107(498):592–606 (2012).

Fan, J., Li, Y., and Yu, K. "Vast volatility matrix estimation using high-frequency data for portfolio selection." *J. Amer. Statist. Assoc.*, 107(497):412–428 (2012).

Fan, J., Liao, Y., and Mincheva, M. "High-dimensional covariance matrix estimation in approximate factor models." *Ann. Statist.*, 39(6):3320–3356 (2011).

—. "Large covariance estimation by thresholding principal orthogonal complements." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(4):603–680 (2013).

Fastrich, B., Paterlini, S., and Winker, P. "Constructing Optimal Sparse Portfolios Using Regularization Methods."

Jagannathan, R. and Ma, T. "Risk reduction in large portfolios: Why imposing the wrong constraints helps." *The Journal of Finance*, 58(4):1651–1684 (2003).

Jobson, J. D. and Korkie, B. M. “Performance hypothesis testing with the Sharpe and Treynor measures.” *The Journal of Finance*, 36(4):889–908 (1981).

Kan, R. and Zhou, G. “Optimal portfolio choice with parameter uncertainty.” *Journal of Financial and Quantitative Analysis*, 42(3):621–656 (2007).

Ledoit, O. and Wolf, M. “A well-conditioned estimator for large-dimensional covariance matrices.” *Journal of Multivariate Analysis*, 88(2):365–411 (2004).

—. “Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks.” *The Review of Financial Studies* (2017).

Memmel, C. “Performance hypothesis testing with the Sharpe ratio.” *Finance Letters*, 1(1) (2003).

Michaud, R. O. “The Markowitz optimization enigma: is “optimized" optimal?” *Financial Analysts Journal*, 31–42 (1989).