Asset Allocation and Risk Assessment with Gross Exposure Constraints

Forrest Zhang

Bendheim Center for Finance
Princeton University

A joint work with Jianqing Fan and Ke Yu, Princeton
Introduction
Markowitz’s Mean-variance analysis

Problem: \( \min_w \ w^T \Sigma w, \) \quad s.t. \( w^T 1 = 1, \) and \( w^T \mu = r_0. \)

Solution: \( w = c_1 \Sigma^{-1} \mu + c_2 \Sigma^{-1} 1 \)

- Cornerstone of modern finance where CAPM and many portfolio theory is built upon.
- Too sensitive on input vectors and their estimation errors.
- Can result in extreme short positions (Green and Holdfield, 1992).
- More severe for large portfolio.
Markowitz’s Mean-variance analysis

Problem: \( \min_w w^T \Sigma w \), s.t. \( w^T 1 = 1 \), and \( w^T \mu = r_0 \).

Solution: \( w = c_1 \Sigma^{-1} \mu + c_2 \Sigma^{-1} 1 \)

- Cornerstone of modern finance where CAPM and many portfolio theory is built upon.
- Too sensitive on input vectors and their estimation errors.
- Can result in extreme short positions (Green and Holdfield, 1992).
- More severe for large portfolio.
Estimating **high-dim** cov-matrices is intrinsically challenging.

- Suppose we have 500 (2000) stocks to be managed. There are 125K (2 m) free parameters!

- Yet, 2-year daily returns yield only about sample size $n = 500$. Accurately estimating it poses significant challenges.

- Impact of dimensionality is large and poorly understood:
  - Risk: $w^T \hat{\Sigma} w$.
  - Allocation: $\hat{c}_1 \hat{\Sigma}^{-1} 1 + \hat{c}_2 \hat{\Sigma}^{-1} \hat{\mu}$.

- Accumulating of millions of estimation errors can have a devastating effect.
Estimating high-dim cov-matrices is intrinsically challenging.

- Suppose we have 500 (2000) stocks to be managed. There are 125K (2 m) free parameters!
- Yet, 2-year daily returns yield only about sample size $n = 500$. Accurately estimating it poses significant challenges.
- Impact of dimensionality is large and poorly understood:
  - Risk: $w^T \hat{\Sigma} w$.
  - Allocation: $\hat{c}_1 \hat{\Sigma}^{-1} 1 + \hat{c}_2 \hat{\Sigma}^{-1} \hat{\mu}$.
- Accumulating of millions of estimation errors can have a devastating effect.
Efforts in Remedy

- Reduce sensitivity of estimation.
  - Shrinkage and Bayesian: —Expected return (Klein and Bawa, 76; Chopra and Ziemba, 93; ) —Cov. matrix (Ledoit & Wolf, 03, 04)
  - Factor-model based estimation (Fan, Fan and Lv , 2008; Pesaran and Zaffaroni, 2008)

- Robust portfolio allocation (Goldfarb and Iyengar, 2003)

- No-short-sale portfolio (De Roon et al., 2001; Jagannathan and Ma, 2003; DeMiguel et al., 2008; Bordie et al., 2008)

- None of them are far enough; no theory.
Efforts in Remedy

- Reduce sensitivity of estimation.
  - Shrinkage and Bayesian: —Expected return (Klein and Bawa, 76; Chopra and Ziemba, 93; ) —Cov. matrix (Ledoit & Wolf, 03, 04)
  - Factor-model based estimation (Fan, Fan and Lv , 2008; Pesaran and Zaffaroni, 2008)

- Robust portfolio allocation (Goldfarb and Iyengar, 2003)

- No-short-sale portfolio (De Roon et al., 2001; Jagannathan and Ma, 2003; DeMiguel et al., 2008; Bordie et al., 2008)

- None of them are far enough; no theory.
About this talk

- Propose utility maximization with gross-sale constraint. It bridges no-short-sale constraint to no-constraint on allocation.

- Oracle (Theoretical), actual and empirical risks are very close.
  - No error accumulation effect.
  - Elements in covariance can be estimated separately; facilitates the use of non-synchronized high-frequency data.
  - Provide theoretical understanding why wrong constraint can even beat Markowitz’s portfolio (Jagannathan and Ma, 2003).

- Portfolio selection and tracking.
  - Select or track a portfolio with limited number of stocks.
  - Improve any given portfolio with modifications of weights on limited number of stocks.
About this talk

- Propose utility maximization with gross-sale constraint. It bridges no-short-sale constraint to no-constraint on allocation.

- Oracle (Theoretical), actual and empirical risks are very close.
  - No error accumulation effect.
  - Elements in covariance can be estimated separately; facilitates the use of non-synchronized high-frequency data.
  - Provide theoretical understanding why wrong constraint can even beat Markowitz’s portfolio (Jagannathan and Ma, 2003).

- Portfolio selection and tracking.
  - Select or track a portfolio with limited number of stocks.
  - Improve any given portfolio with modifications of weights on limited number of stocks.
Outline

1. Portfolio optimization with gross-exposure constraint.
2. Portfolio selection and tracking.
3. Simulation studies
4. Empirical studies:
Short-constrained portfolio selection

\[
\max_w \quad E[U(w^T R)] \\
\text{s.t.} \quad w^T 1 = 1, \quad \|w\|_1 \leq c, \quad Aw = a.
\]

**Equality Constraint:**
- \(A = \mu \implies\) expected portfolio return.
- \(A\) can be chosen so that we put constraint on sectors.

**Short-sale constraint:** When \(c = 1\), no short-sale allowed. When \(c = \infty\), problem becomes Markowitz’s.

- Portfolio selection: solution is usually sparse.
Short-constrained portfolio selection

$$\max_w E[U(w^T R)]$$

s.t.  $$w^T 1 = 1, \quad \|w\|_1 \leq c, \quad Aw = a.$$  

**Equality Constraint:**
- $A = \mu \rightarrow$ expected portfolio return.
- $A$ can be chosen so that we put constraint on sectors.

**Short-sale constraint:** When $c = 1$, no short-sale allowed. When $c = \infty$, problem becomes Markowitz’s.

- Portfolio selection: solution is usually sparse.
Risk optimization Theory

**Actual and Empirical risks:**
\[ R(w) = w^T \Sigma w, \quad R_n(w) = w^T \hat{\Sigma} w. \]

\[
\begin{align*}
\mathbf{w}_{opt} &= \text{argmin}_{\| \mathbf{w} \|_1 \leq c} R(\mathbf{w}), \\
\hat{\mathbf{w}}_{opt} &= \text{argmin}_{\| \mathbf{w} \|_1 \leq c} R_n(\mathbf{w})
\end{align*}
\]

• Risks: \( \sqrt{R(\mathbf{w}_{opt})} \) —oracle, \( \sqrt{R_n(\hat{\mathbf{w}}_{opt})} \) —empirical; \( \sqrt{R(\hat{\mathbf{w}}_{opt})} \) —actual risk of a selected portfolio.

**Theorem 1**: Let \( a_n = \| \hat{\Sigma} - \Sigma \|_\infty \). Then, we have

\[
\begin{align*}
|R(\hat{\mathbf{w}}_{opt}) - R(\mathbf{w}_{opt})| &\leq 2a_n c^2 \\
|R(\hat{\mathbf{w}}_{opt}) - R_n(\hat{\mathbf{w}}_{opt})| &\leq a_n c^2 \\
|R(\mathbf{w}_{opt}) - R_n(\hat{\mathbf{w}}_{opt})| &\leq a_n c^2.
\end{align*}
\]
Risk optimization Theory

**Actual and Empirical risks:**
\[ R(w) = w^T \Sigma w, \quad R_n(w) = w^T \hat{\Sigma} w. \]

\[ w_{opt} = \arg\min_{\|w\|_1 \leq c} R(w), \quad \hat{w}_{opt} = \arg\min_{\|w\|_1 \leq c} R_n(w) \]

- Risks: \( \sqrt{R(w_{opt})} \) — oracle, \( \sqrt{R_n(\hat{w}_{opt})} \) — empirical; \( \sqrt{R(\hat{w}_{opt})} \) — actual risk of a selected portfolio.

**Theorem 1:** Let \( a_n = \|\hat{\Sigma} - \Sigma\|_\infty \). Then, we have

\[
\begin{align*}
|R(\hat{w}_{opt}) - R(w_{opt})| & \leq 2a_n c^2 \\
|R(\hat{w}_{opt}) - R_n(\hat{w}_{opt})| & \leq a_n c^2 \\
|R(w_{opt}) - R_n(\hat{w}_{opt})| & \leq a_n c^2.
\end{align*}
\]
**Theorem 2**: If for a sufficiently large $x$,

$$\max_{i,j} P\{\sqrt{n}|\sigma_{ij} - \hat{\sigma}_{ij}| > x\} < \exp(-Cx^{1/a}),$$

for some two positive constants $a$ and $C$, then

$$\|\Sigma - \hat{\Sigma}\|_{\infty} = O_P\left(\frac{(\log p)^a}{\sqrt{n}}\right).$$

- Impact of dimensionality is limited.
Accuracy of Covariance: I

**Theorem 2:** If for a sufficiently large \( x \),

\[
\max_{i,j} \mathbb{P} \{ \sqrt{n} | \sigma_{ij} - \hat{\sigma}_{ij} | > x \} < \exp(-Cx^{1/a}),
\]

for some two positive constants \( a \) and \( C \), then

\[
\| \Sigma - \hat{\Sigma} \|_\infty = O_P \left( \frac{(\log p)^a}{\sqrt{n}} \right).
\]

- Impact of dimensionality is limited.
Algorithms

\[ \min_{\mathbf{w}^T \mathbf{1} = 1, \| \mathbf{w} \|_1 \leq c} \mathbf{w}^T \Sigma \mathbf{w}. \]

1. Quadratic programming for each given \( c \) (Exact).
2. Coordinatewise minimization.
3. LARS approximation.
Regression problem: Letting $Y = R_p$ and $X_j = R_p - R_j$,

$$\text{var}(w^T R) = \min_b E(w^T R - b)^2$$

$$= \min_b E(Y - w_1 X_1 - \cdots - w_{p-1} X_{p-1} - b)^2,$$

Gross exposure: $\|w\|_1 = \|w^*\|_1 + |1 - 1^T w^*| \leq c,$
not equivalent to $\|w^*\|_1 \leq d$.

$d = 0$ picks $X_p$, but $c = 1$ picks multiple stocks.
**Regression problem:** Letting $Y = R_p$ and $X_j = R_p - R_j$,

$$\text{var}(w^T R) = \min_b E(w^T R - b)^2$$

$$= \min_b E(Y - w_1 X_1 - \cdots - w_{p-1} X_{p-1} - b)^2,$$

**Gross exposure:** $\|w\|_1 = \|w^*\|_1 + |1 - 1^T w^*| \leq c$,

not equivalent to $\|w^*\|_1 \leq d$.

• $d = 0$ picks $X_p$, but $c = 1$ picks multiple stocks.
Approximate solution

**LARS**: to find solution path $w^*(d)$ for PLS

$$
\min_{b, \|w^*\|_1 \leq d} E(Y - w^*T X - b)^2,
$$

**Approximate solution**: PLS provides a suboptimal solution to risk optimization problem with

$$
c = d + |1 - 1^T w^*_{opt}(d)|.
$$

• Take $Y = \text{optimal no-short-sale constraint (} c = 1\).$
• Multiple $Y$ helps. e.g. Also take $Y = \text{solution to } c = 2$
Portfolio tracking and improvement

- PLS regarded as finding a portfolio to minimize the expected tracking error — **portfolio tracking**.

- PLS interpreted as modifying weights to improve the performance of $Y$ — **Portfolio improvements**.
  - with ♠ limited number of stocks   ♠ limited exposure.
  - empirical risk path $R_n(d)$ helps decision making.

**Remark:** PLS $\min_b, \|w^*\|_1 \leq d \sum_{t=1}^{n} (Y_i - w^*^T X_t^* - b)^2$ is equivalent to PLS using **sample covariance** matrix.
Data: \( Y = \text{CRSP}; \ X = 10 \) industrial portfolios. Today = 1/8/05. Sample Cov: one-year daily return. Actual: hold one year.
Fama-French three-factor model

**Model**: \( R_i = b_{i1} f_1 + b_{i2} f_2 + b_{i3} f_3 + \varepsilon_i \) or \( R = Bf + \varepsilon \).

★ \( f_1 \) = CRSP index; ★ \( f_2 \) = size effect; ★ \( f_3 \) = book-to-market effect

**Covariance**: \( \Sigma = B \text{cov}(f) B^T + \text{diag}(\sigma_1^2, \cdots, \sigma_p^2) \).

<table>
<thead>
<tr>
<th>Parameters for factor loadings</th>
<th>Parameters for factor returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_b )</td>
<td>( \mu_f )</td>
</tr>
<tr>
<td>.783</td>
<td>.0291</td>
</tr>
<tr>
<td>.518</td>
<td>.0239</td>
</tr>
<tr>
<td>.410</td>
<td>.0102</td>
</tr>
</tbody>
</table>

**Parameters**: Calibrated to market data (5/1/02–8/29/05, from Fan, Fan and Lv, 2008)

— **Parameters**:
  - Factor loadings: \( b_i \sim_{i.i.d.} N(\mu_b, \text{cov}_b) \)
  - Noise: \( \sigma_i \sim_{i.i.d.} \text{Gamma}(3.34, .19) \) conditioned on \( \sigma_i > .20 \).

— **Simulation**: Factor returns \( f_t \sim_{i.i.d.} N(\mu_f, \text{cov}_f) \),

\( \varepsilon_{it} \sim_{i.i.d.} \sigma_{it} t^* \)
Fama-French three-factor model

**Model:** $R_i = b_{i1}f_1 + b_{i2}f_2 + b_{i3}f_3 + \varepsilon_i$ or $R = Bf + \varepsilon$.

$\star f_1 =$ CRSP index; $\star f_2 =$ size effect; $\star f_3 =$ book-to-market effect

**Covariance:** $\Sigma = B\text{cov}(f)B^T + \text{diag}(\sigma_1^2, \cdots, \sigma_p^2)$.

<table>
<thead>
<tr>
<th>Parameters for factor loadings</th>
<th>Parameters for factor returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_b$</td>
<td>$\mu_f$</td>
</tr>
<tr>
<td>.783</td>
<td>.0291</td>
</tr>
<tr>
<td>.518</td>
<td>.0239</td>
</tr>
<tr>
<td>.410</td>
<td>.0102</td>
</tr>
<tr>
<td>$\text{cov}_b$</td>
<td>$\text{cov}_f$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Parameters:** Calibrated to market data (5/1/02–8/29/05, from Fan, Fan and Lv, 2008)

---

**Parameters:**

- Factor loadings: $b_i \sim i.i.d. \ N(\mu_b, \text{cov}_b)$
- Noise: $\sigma_i \sim i.i.d. \ \text{Gamma}(3.34, .19)$ conditioned on $\sigma_i > .20$.

---

**Simulation:** Factor returns $f_t \sim i.i.d. \ N(\mu_f, \text{cov}_f)$,

$\varepsilon_{it} \sim i.i.d. \ \sigma_i t_6^*$
Risk Improvements and decision making

- **Sample**
  - (a) Empirical and actual risks – sample cov
  - (b) Number of stocks – sample cov

- **Factor**
  - (c) Empirical and actual risks – factor cov
  - (d) Number of stocks – factor cov

- Factor-model based estimation is more accurate.
Empirical studies (I)
Some details

**Data**: 100 portfolios from the website of Kenneth French from 1998–2007 (10 years)

**Portfolios**: two-way sort according to the size and book-to-equity ratio, 10 categories each.

**Evaluation**: Rebalance monthly, and record daily returns.

**Covariance matrix**: Estimate by sample covariance matrix, factor model used last twelve months daily data, and RiskMetrics.
Risk, Sharpe-Ratio, Maximum Weight, Annualized return
# Short-constrained MV portfolio (Results I)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Mean</th>
<th>Std</th>
<th>Sharpe-R</th>
<th>Max-W</th>
<th>Min-W</th>
<th>Long</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Covariance Matrix Estimator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No short (c = 1)</td>
<td>19.51</td>
<td>10.14</td>
<td>1.60</td>
<td>0.27</td>
<td>-0.00</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Exact (c = 1.5)</td>
<td>21.04</td>
<td>8.41</td>
<td>2.11</td>
<td>0.25</td>
<td>-0.07</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Exact (c = 2)</td>
<td>20.55</td>
<td>7.56</td>
<td>2.28</td>
<td>0.24</td>
<td>-0.09</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Exact (c = 3)</td>
<td>18.26</td>
<td>7.13</td>
<td>2.09</td>
<td>0.24</td>
<td>-0.11</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>Approx. (c = 2)</td>
<td>21.16</td>
<td>7.89</td>
<td>2.26</td>
<td>0.32</td>
<td>-0.08</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Approx. (c = 3)</td>
<td>19.28</td>
<td>7.08</td>
<td>2.25</td>
<td>0.28</td>
<td>-0.11</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>GMV</td>
<td>17.55</td>
<td>7.82</td>
<td>1.82</td>
<td>0.66</td>
<td>-0.32</td>
<td>52</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unmanaged Index</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal-W</td>
<td>10.86</td>
<td>16.33</td>
<td>0.46</td>
<td>0.01</td>
<td>0.01</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>CRSP</td>
<td>8.2</td>
<td>17.9</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Empirical studies (II)
Some details

**Data**: 1000 stocks with missing data selected from Russell 3000 from 2003-2007 (5 years).

**Allocation**: Each month, pick 400 stocks at random and allocate them (mitigating survivor biases).

**Evaluation**: Rebalance monthly, and record daily returns.

**Covariance matrix**: Estimate by sample covariance matrix, factor model used last **twenty-four** months daily data, and RiskMetrics.
(a) Risk of portfolios (NS)

(b) Risk of portfolios (mkt)
Conclusion

- Utility maximization with gross-sale constraint bridges no-short-sale constraint to no-constraint on allocation.

- It makes oracle (theoretical), actual and empirical risks close:
  - No error accumulation effect for a range of $c$;
  - Elements in covariance can be estimated separately; facilitates use of non-synchronize high-frequency data.
  - Provide theoretical understanding why wrong constraint help.

- Portfolio selection, tracking, and improvement.
  - Select or track a portfolio with limited number of stocks.
  - Improve any given portfolio with modifications of weights on limited number of stocks.
  - Provide tools for checking efficiency of a portfolio.
Conclusion

- Utility maximization with gross-sale constraint bridges no-short-sale constraint to no-constraint on allocation.

- It makes oracle (theoretical), actual and empirical risks close:
  - No error accumulation effect for a range of $c$;
  - Elements in covariance can be estimated separately; facilitates use of non-synchronize high-frequency data.
  - Provide theoretical understanding why wrong constraint help.

- Portfolio selection, tracking, and improvement.
  - Select or track a portfolio with limited number of stocks.
  - Improve any given portfolio with modifications of weights on limited number of stocks.
  - Provide tools for checking efficiency of a portfolio.