IDENTIFYING BROAD AND NARROW FINANCIAL RISK FACTORS VIA CONVEX OPTIMIZATION: PART II

Alexander D. Shkolnik
ads2@berkeley.edu

joint with Jeffrey Bohn and Lisa Goldberg.
Overview

- Simulations (synthetic data).
- Performance metrics (finance).
  - benchmark: PCA.
- Theoretical considerations for $N > T$. 
Simulations (synthetic data)

- $N = 125$ and $N = 500$ securities.
- $T = 250$ observations (one year of daily data).
  - $K = 2$ broad factors (16% ann. vol., 4% ann. vol.)
  - $\kappa \approx N / \log N$ and $\kappa \approx \sqrt{N}$ countries.
- 400 simulations (realizations of a sample covariance matrix).
- Convex program parameters: $\theta = (\gamma, \lambda)$,
  - $\gamma = 1 / \sqrt{N}$ (see Candès, Li, Ma & Wright (2011)),
  - $\lambda$ set to control the number of recovered broad factors.
Input to algorithm ($N = 125, T = 250$)
Ordered by country
True vs recovered broad factor matrix $L_K$
True vs recovered narrow factor matrix $L_κ$
True vs recovered specific risk matrix
Measuring performance on simulated data

- Typically, for estimator $\Sigma(\theta)$, consider $\|\Sigma(\theta) - \Sigma\|$ for some norm.
- Instead, consider the variance of returns to portfolio $w$.

$$\text{Var}_{\Sigma}(w) = w^\top \Sigma w$$  \hspace{1cm} (1)

- The portfolio risk forecasting ratio is computed as

$$\mathcal{R}_{\Sigma}(w \mid \theta) = \frac{\text{Var}_{\Sigma(\theta)}(w)}{\text{Var}_{\Sigma}(w)}.$$  \hspace{1cm} (2)

- Ratios $\mathcal{R}_{L_K}$, $\mathcal{R}_{L_\kappa}$ and $\mathcal{R}_{\Delta}$ are defined analogously.
- Other estimators: $K(\theta)$ and $\kappa(\theta)$. 
Test portfolios

- Let $e = (1, \ldots, 1)^T$. Equally weighted portfolio is given by
  \[ w = e/N \]  
  (3)

- Minimum variance (long only) portfolio is the solution $w(\theta)$ of
  \[ \min_w w^T \Sigma(\theta) w \]  
  subject to $w^T e = 1$,  
  (4)
  (5)

- Minimum variance (long/short): solves (4) subject to $w^T e = 1$,
  \[ w(\theta) = \frac{\Sigma^{-1}(\theta) e}{e^T \Sigma^{-1}(\theta) e}. \]  
  (7)
Number of factors: equally weighted portfolio

<table>
<thead>
<tr>
<th>$N$</th>
<th>Metric</th>
<th>DSLR: $(\lambda \times 10^3)$</th>
<th>PCA</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>$K(\theta)$</td>
<td>8.51</td>
<td>2.45</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.83)</td>
<td>(0.44)</td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>$\kappa(\theta)$</td>
<td>20.3</td>
<td>22.4</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.21)</td>
<td>(0.86)</td>
<td>(0.00)</td>
</tr>
</tbody>
</table>

$\star T = 250$ (observations).
## Performance: equally weighted portfolio

<table>
<thead>
<tr>
<th>$N$</th>
<th>Metric</th>
<th>DSLR: ($\lambda \times 10^3$)</th>
<th>PCA</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>$\mathcal{R}_\Sigma$</td>
<td>1.00</td>
<td>0.96</td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>500</td>
<td>$\mathcal{R}_{LK}$</td>
<td>0.97</td>
<td>0.94</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>500</td>
<td>$\mathcal{R}_{L^\kappa}$</td>
<td>1.05</td>
<td>0.68</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.09)</td>
<td>(0.05)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>500</td>
<td>$\mathcal{R}_\Delta$</td>
<td>0.89</td>
<td>1.12</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
</tbody>
</table>
## Performance: optimized portfolio

<table>
<thead>
<tr>
<th>$N$</th>
<th>Metric</th>
<th>DSLR: ($\lambda \times 10^3$)</th>
<th>PCA</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>$R_{\Sigma}$</td>
<td>0.97</td>
<td>0.88</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>500</td>
<td>$R_{L_K}$</td>
<td>0.86</td>
<td>0.70</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>500</td>
<td>$R_{L_K}$</td>
<td>0.95</td>
<td>0.48</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.10)</td>
<td>(0.07)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>500</td>
<td>$R_{\Delta}$</td>
<td>0.96</td>
<td>1.22</td>
<td>1.31</td>
</tr>
<tr>
<td></td>
<td>(stdev)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>
Summary of findings

• **SLRD** ourperforms classical **PCA** on risk forecasts.

• Accurate sparse component recovery implies:
  – accurate $R_\Delta$ forecast,
  – accurate $\kappa(\theta)$ estimates.

• Inaccurate $K(\theta)$ estimates does not imply poor risk forecast.

• $N > T$?
\( \kappa (\theta) \) estimates \((N > T)\)

- With high probability \(^\text{a}\) CPW recovers correctly the
  - rank of \(\mathcal{L}\) and sparsity support of \(S\).
  - Rate of convergence \(\sqrt{N/T}\) arises from bounds on
    \[
    \|\mathcal{L} - \mathcal{L}(\theta)\|_2 \text{ and } \|S - S(\theta)\|_{\ell_\infty}.
    \] (8)

- The error on \(\|S - S(\theta)\|_{\ell_\infty}\) may be improved to order
  \[
  \sqrt{\frac{\log N}{T}}.
  \] (9)

- When sparsity pattern is correct, \(\kappa (\theta)\) is correct.

\(^{\text{a}}\) (Chandrasekaran, Parrilo & Willsky 2012)
$K(\theta)$ estimates ($N > T$)

- With high probability $\text{CPW}^a$ recovers correctly the
  - rank of $\mathcal{L}$ and sparsity support of $S$.
  - Rate of convergence $\sqrt{N/T}$ arises from bounds on
    \[ \| \mathcal{L} - \mathcal{L}(\theta) \|_2 \]  
\[ (10) \]

- **Cannot** estimate number of broad factors accurately for $N > T$.

- Numerical experiments suggest $\| \mathcal{L} - \mathcal{L}(\theta) \|_{\ell\infty}$ may be order
  \[ \sqrt{\log N} \frac{\sqrt{\log N}}{T}. \]  
\[ (11) \]

---

$^a$(Chandrasekaran et al. 2012)
Risk forecasts

• **Equally weighted:** accuracy of portfolio risk forecast depends on

\[ \left| \mathcal{R}_\Sigma (w | \theta) - 1 \right| \lesssim \| \mathcal{L} (\theta) - \mathcal{L} \|_{\ell_\infty} \wedge \| S (\theta) - S \|_{\ell_\infty} \]  

\[ = O \left( \sqrt{\frac{\log N}{T}} \right). \]  

(12)

(13)

• **Optimized:** same bound holds with different constants.

• Similar for factor and specific risk forecasts.

• Support for accuracy of risk forecasts in \( \log (N) / T \downarrow 0 \) regime.
Conclusions

• The approach shows promise in extracting narrow factors that highlight country/industry relationships in data.

• Financial applications require a reformulation of available low rank and sparse decomposition methods.
  – low-rank + sparse + diagonal decomposition,
  – sparse eigenvectors (narrow factors).

• Finance oriented performance metrics may lead to an alternative analysis of estimator consistency.
Ongoing & future work

- Theoretical performance guarantees for risk forecasting ratios.
- Data-driven methods of selecting optimal parameters $\theta = (\lambda, \gamma)$.
- Alternative convex programs guided by performance metrics.
- Scale algorithms to tens of thousands of securities.
Questions.
(ads2@berkeley.edu)
\( R_\Delta \) (specific risk) forecast

- **MTFA**\(^a\) decomposes a given matrix \( S \) into a low rank \( L_\kappa = XGX^T \) component and a diagonal component \( \Delta \) exactly if

\[
\max_{x \in \mathcal{X}/ \{0_N \}} \frac{\|x\|_\infty}{\|x\|_2} < \frac{1}{2}.
\]

where \( \mathcal{X} \) is the space spanned by the narrow factor exposures in \( X \).

- A sufficient condition for (14) is

\[
x_i^2 < \sum_{i \neq j} x_j^2 \text{ for all } i \text{ and } x \in \mathcal{X}.
\]

- Robust recovery for factors that are not too narrow.

\(^a\)(Saunderson, Chandrasekaran, Parrilo & Willsky 2012)
References

