Identifying Broad and Narrow Financial Risk Factors via Convex Optimization: Part II

Alexander D. Shkolnik

ads2@berkeley.edu

MMDS Workshop. June 22, 2016.

joint with Jeffrey Bohn and Lisa Goldberg.

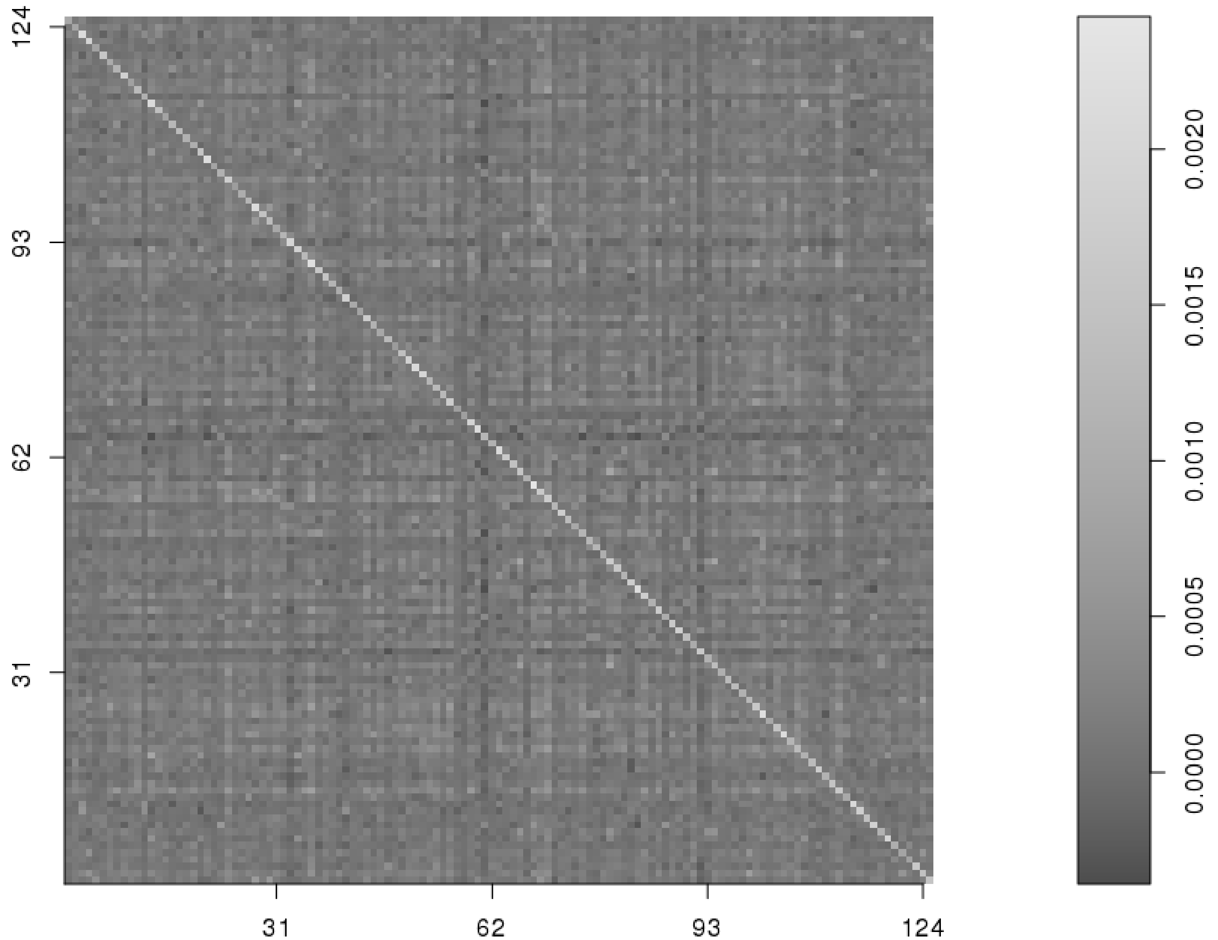
Overview

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

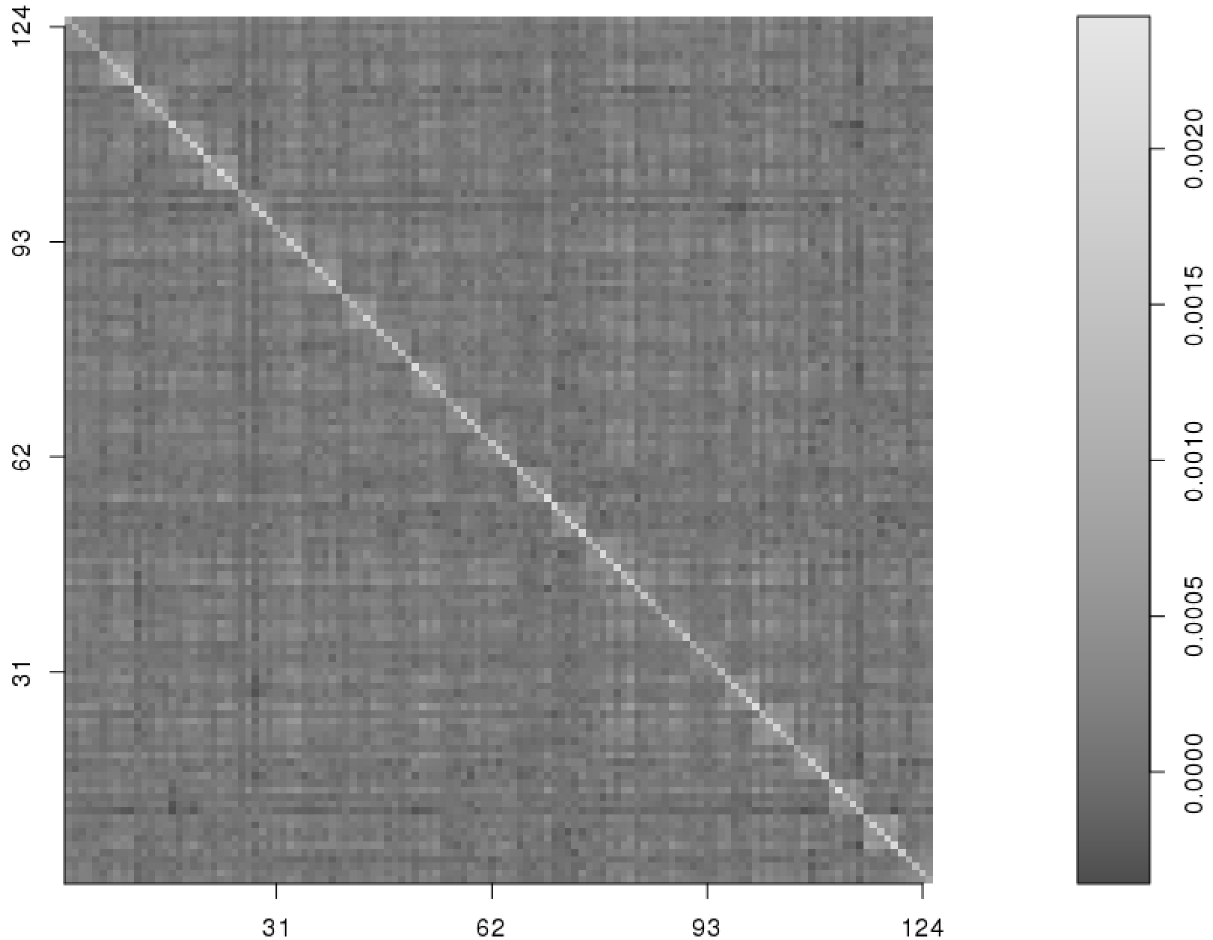
Simulations (synthetic data)

- $N = 125$ 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)),
  - λ set to to control the number of recovered broad factors.

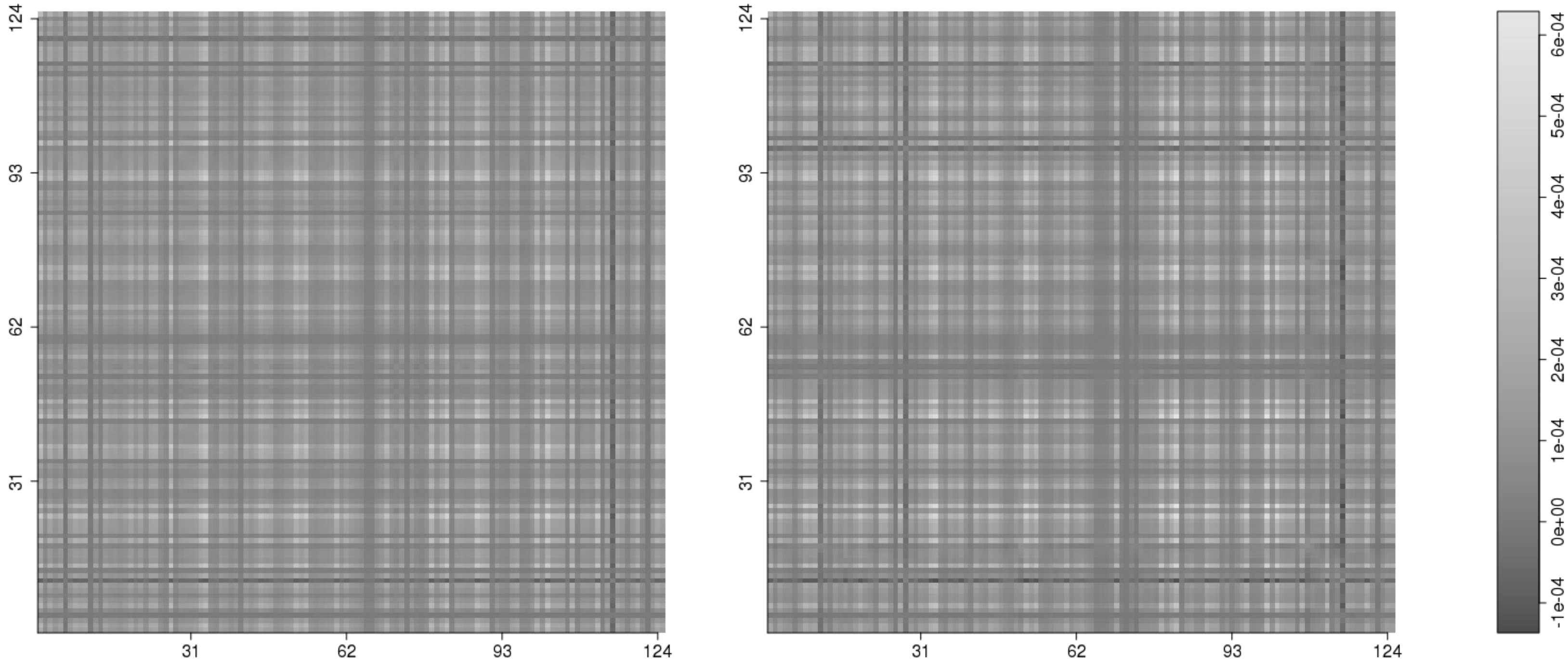
Input to algorithm ($N = 125, T = 250$)



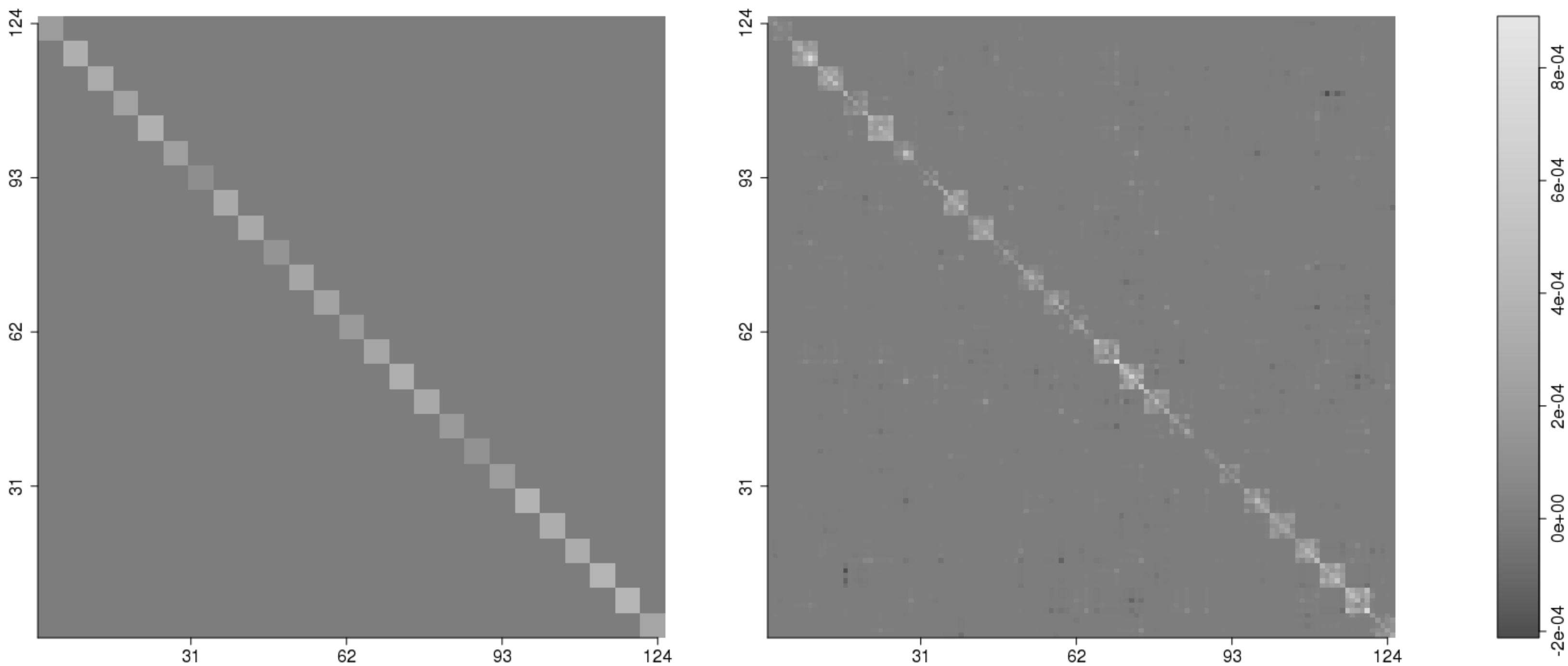
Ordered by country



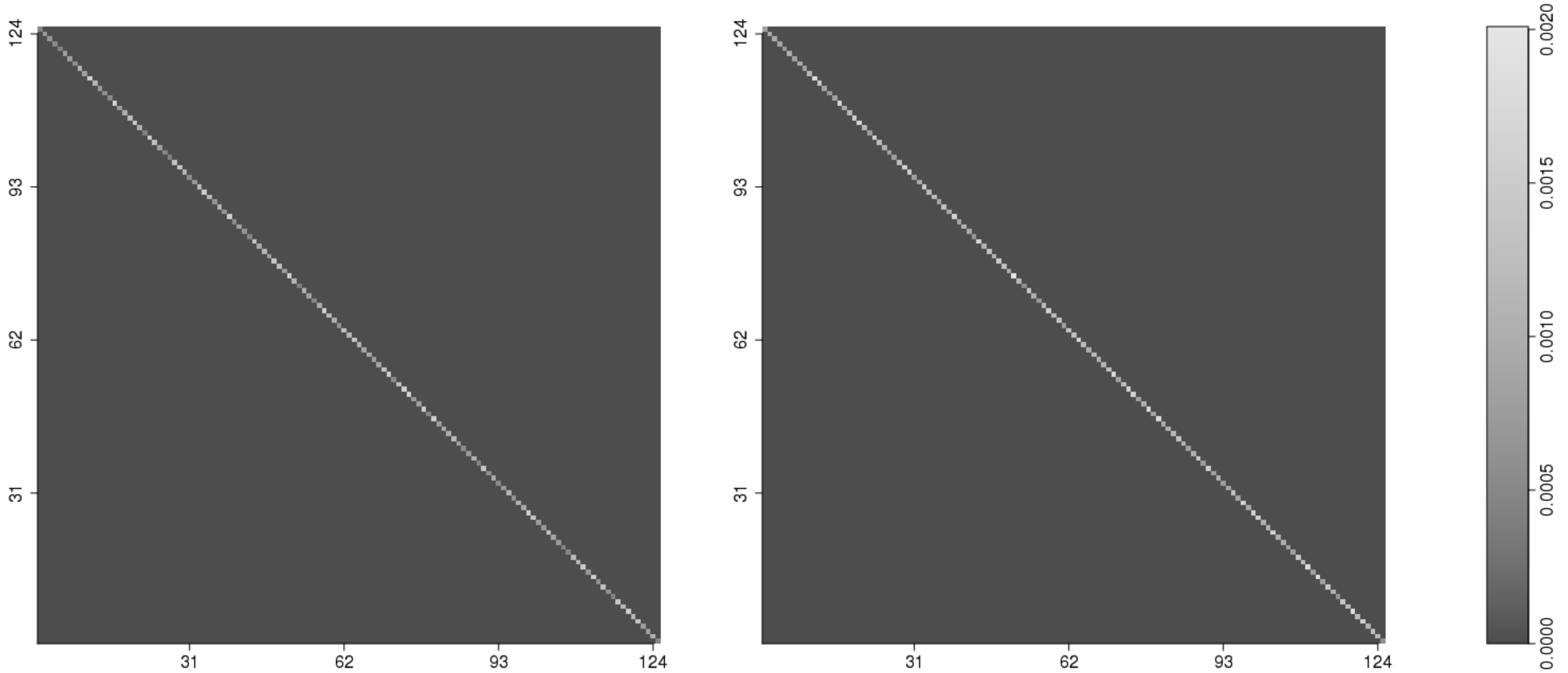
True vs recovered broad factor matrix L_K



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.

$$\mathbf{Var}_{\Sigma}(w) = w^{\top}\Sigma w \tag{1}$$

- The portfolio **risk forecasting ratio** is computed as

$$\mathscr{R}_{\Sigma}(w \mid \theta) = \frac{\mathbf{Var}_{\Sigma(\theta)}(w)}{\mathbf{Var}_{\Sigma}(w)} . \tag{2}$$

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

Test portfolios

- Let $e = (1, \dots, 1)^\top$. **Equally weighted** portfolio is given by

$$w = e/N \tag{3}$$

- **Minimum variance** (long only) portfolio is the solution $w(\theta)$ of

$$\min_{w} w^\top \Sigma(\theta) w \tag{4}$$

$$\text{subject to } w^\top e = 1\,, \tag{5}$$

$$w \geq 0\,. \tag{6}$$

- Minimum variance (long/short): solves (4) subject to $w^\top e = 1$,

$$w(\theta) = \frac{\Sigma^{-1}(\theta) e}{e^\top \Sigma^{-1}(\theta) e}\,. \tag{7}$$

Number of factors: equally weighted portfolio

N	Metric	DSLR: ($\lambda \times 10^3$) 6.4	DSLR: ($\lambda \times 10^3$) 12.8	PCA	Ground Truth K, κ
500	$K(\theta)$	8.51	2.45	2.00	2
	(stdev)	(0.83)	(0.44)	(0.00)	
	$\kappa(\theta)$	20.3	22.4	n/a	22
	(stdev)	(0.21)	(0.86)	(0.00)	

⋆ $T = 250$ (observations).

Performance: equally weighted portfolio

N	Metric	DSLR: $(\lambda \times 10^3)$ 6.4	DSLR: $(\lambda \times 10^3)$ 12.8	PCA	Ground Truth Ann. Vol.
500	$\mathscr{R}_{\Sigma}$	1.00	0.96	1.06	0.17
	(stdev)	(0.08)	(0.08)	(0.08)	
	$\mathscr{R}_{L_K}$	0.97	0.94	1.18	0.16
	(stdev)	(0.09)	(0.09)	(0.10)	
	$\mathscr{R}_{L_\kappa}$	1.05	0.68	n/a	0.05
	(stdev)	(0.09)	(0.05)	(0.00)	
	$\mathscr{R}_{\Delta}$	0.89	1.12	1.16	0.02
	(stdev)	(0.01)	(0.01)	(0.01)	

Performance: optimized portfolio

N	Metric	DSLR: ($\lambda \times 10^3$)		PCA	Ground Truth
		6.4	12.8		Ann. Vol.
500	$\mathscr{R}_{\Sigma}$	0.97	0.88	0.81	0.11
	(stdev)	(0.06)	(0.06)	(0.07)	
	$\mathscr{R}_{L_K}$	0.86	0.70	0.83	0.07
	(stdev)	(0.08)	(0.10)	(0.13)	
	$\mathscr{R}_{L_\kappa}$	0.95	0.48	n/a	0.04
	(stdev)	(0.10)	(0.07)	(0.00)	
	$\mathscr{R}_{\Delta}$	0.96	1.22	1.31	0.08
	(stdev)	(0.02)	(0.03)	(0.05)	

Summary of findings

- **SLRD** ourperforms classical **PCA** on risk forecasts.
- Accurate sparse component recovery implies:
  - accurate $\mathscr{R}_\Delta$ forecast,
  - accurate $\kappa(\theta)$ estimates.
- Innacurate $K(\theta)$ estimates **does not** imply poor risk forecast.
- $N > T$?

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

- With high probability **CPW**[a] recovers correctly the
  - rank of $\mathcal{L}$ and sparsity support of $\mathcal{S}$.
  - Rate of convergence $\sqrt{N/T}$ arises from bounds on

$$\|\mathcal{L} - \mathcal{L}(\theta)\|_2 \text{ and } \|\mathcal{S} - \mathcal{S}(\theta)\|_{\ell_\infty}. \tag{8}$$

- The error on $\|\mathcal{S} - \mathcal{S}(\theta)\|_{\ell_\infty}$ may be improved to order

$$\sqrt{\frac{\log N}{T}}. \tag{9}$$

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

[a](Chandrasekaran, Parrilo & Willsky 2012)

$K(\theta)$ estimates $(N > T)$

- With high probability **CPW**[a] recovers correctly the
  - rank of $\mathcal{L}$ and sparsity support of S.
  - Rate of convergence $\sqrt{N/T}$ arises from bounds on

$$\|\mathcal{L} - \mathcal{L}(\theta)\|_2 \tag{10}$$

- **Cannot** estimate number of broad factors accurately for $N > T$.
- Numerical experiments suggest $\|\mathcal{L} - \mathcal{L}(\theta)\|_{\ell_\infty}$ may be order

$$\sqrt{\frac{\log N}{T}}. \tag{11}$$

[a](Chandrasekaran et al. 2012)

Risk forecasts

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

$$\left|\mathscr{R}_{\Sigma}(w \mid \theta) - 1\right| \lesssim \|\mathcal{L}(\theta) - \mathcal{L}\|_{\ell_\infty} \wedge \|\mathcal{S}(\theta) - \mathcal{S}\|_{\ell_\infty} \tag{12}$$

$$= O\Big(\sqrt{\frac{\log N}{T}}\Big). \tag{13}$$

- **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 paramters $\theta = (\lambda, \gamma)$.
- Alternative convex programs guided by performance metrics.
- Scale algorithms to tens of thousands of securities.

Questions.

(ads2@berkeley.edu)

$\mathscr{R}_\Delta$ (specific risk) forecast

- **MTFA**[a] decomposes a given matrix S into a low rank $L_\kappa = XGX^\top$ component and a diagonal component Δ exactly if

$$\max_{x \in \mathscr{X}/\{0_N\}} \frac{\|x\|_\infty}{\|x\|_2} < 1/2. \tag{14}$$

where $\mathscr{X}$ is the space spanned by the narrow factor exposures in X.

- A sufficient condition for (14) is

$$x_i^2 < \sum_{i \neq j} x_j^2 \quad \text{for all } i \text{ and } x \in \mathscr{X}. \tag{15}$$

- Robust recovery for factors that are not too narrow.

[a](Saunderson, Chandrasekaran, Parrilo & Willsky 2012)

References

Candès, Emmanuel J, Xiaodong Li, Yi Ma & John Wright (2011), 'Robust principal component analysis?', *Journal of the ACM (JACM)* **58**(3), 11.

Chandrasekaran, Venkat, Pablo A Parrilo & Alan S Willsky (2012), 'Latent variable graphical model selection via convex optimization', *The Annals of Statistics (with discussion)* **40**(4), 1935–1967.

Saunderson, James, Venkat Chandrasekaran, Pablo A Parrilo & Alan S Willsky (2012), 'Diagonal and low-rank matrix decompositions, correlation matrices, and ellipsoid fitting', *SIAM Journal on Matrix Analysis and Applications* **33**(4), 1395–1416.