Identifying **Financial** Risk Factors

with a Low-Rank Sparse Decomposition

Security Return is *Driven* by Risk Factors

Broad factors

Narrow factors

Specific factor

$$R = \underbrace{\sum_{k=1}^{K} Y_k \phi_k}_{\text{Broad factors}} + \underbrace{\sum_{k=1}^{\kappa} X_k \psi_k}_{\text{Narrow factors}} + \underbrace{\epsilon}_{\text{Specific factor}}$$

Examples of **Factors**

Broad factors
- → The market
- → Interest rates
- → Styles

Narrow factors
- → Countries
- → Sectors
- → Currencies

The Security Return Covariance Matrix Has a Low-Rank Sparse Structure

$$= YFY^\top + XGX^\top + \Delta$$
$$= L + M + \Delta$$
$$= L + S$$

Status *Quo*

Industry Standard

Human analysts determine factors

Academic Standard

PCA models, which rely on asymptotic considerations

Low-Rank *Sparse Decompositions* of Covariance Matrices

Inspired by [Candès et al., 2011] and [Chandrasekaran et al., 2012] we develop a convex optimization that uses the algorithm in [Ma et al., 2013] to provide a low-rank sparse latent factor decomposition of a sample covariance matrix:

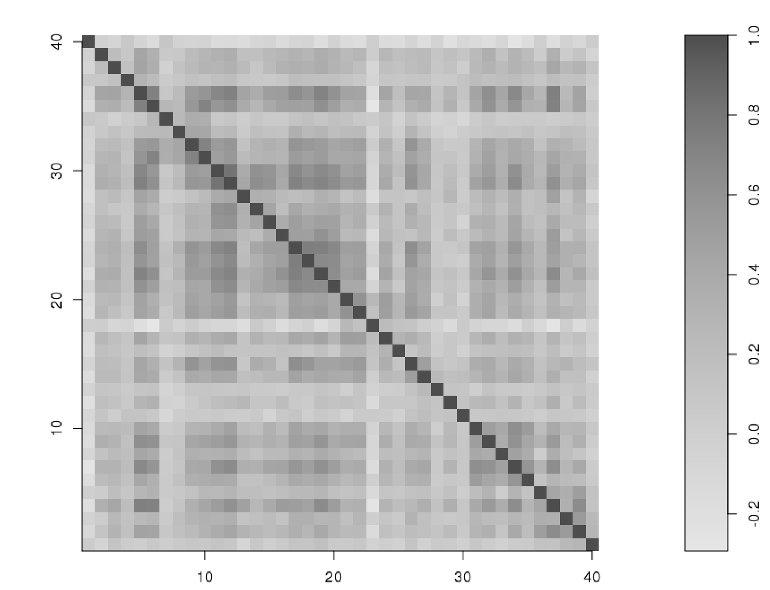
$$\hat{\Sigma} = \hat{L} + \hat{S} + \text{noise}$$

No human analysts or asymptotic considerations are required.

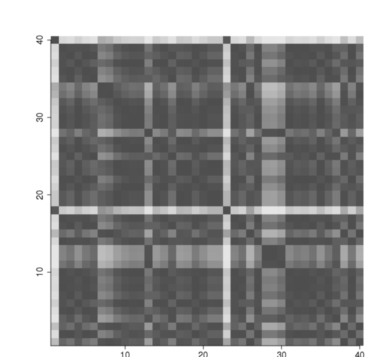
Simulation

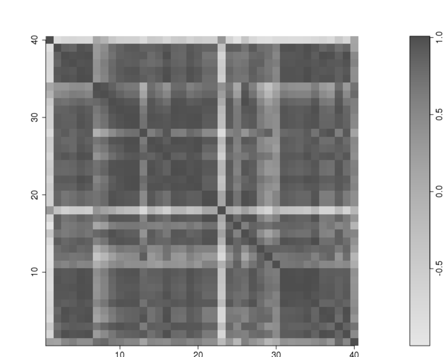
N = securities
T = 250 observations
K = 2 broad factors (the market, size)
κ = narrow factors (China, India, Saudi Arabia, Argentina, ...)

Sample *Correlation* Matrix

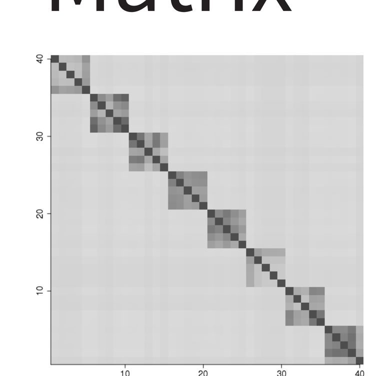


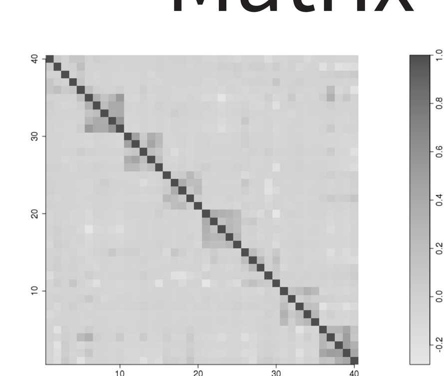
True **Low Rank** Correlation Matrix VS *Recovered* **Low Rank** Correlation Matrix



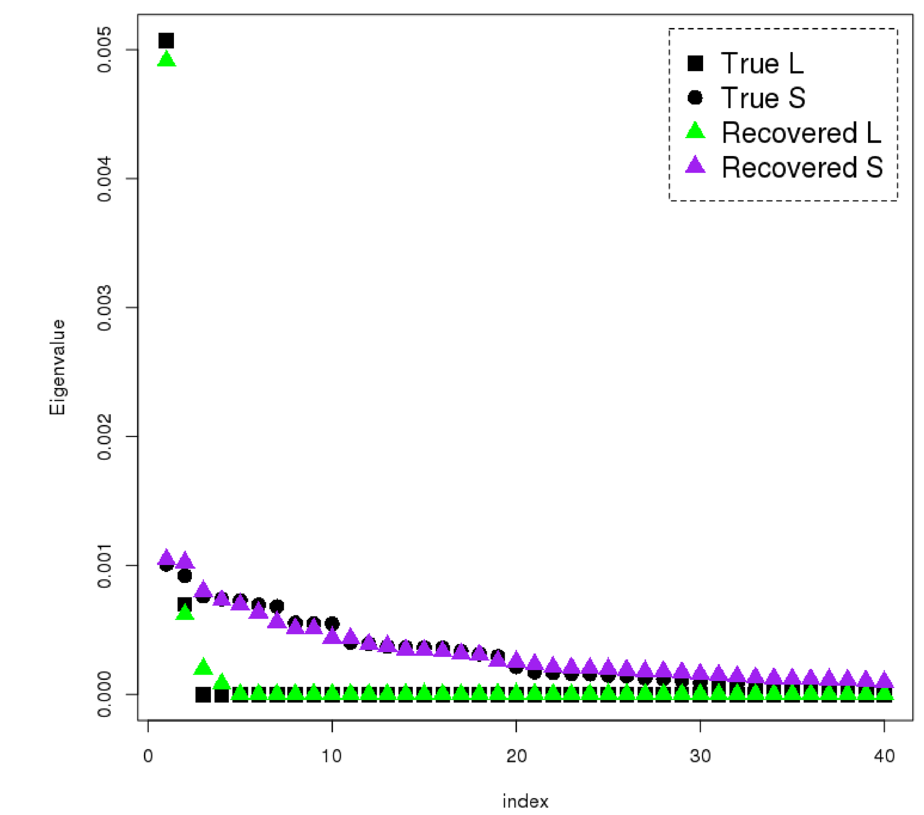


True **Sparse** Correlation Matrix VS *Recovered* **Sparse** Correlation Matrix





Eigenvalues of *Recovered* Decomposition



Summary

Modern data science techniques based on low-rank sparse decompositions may be able to automate financial factor analysis in the future

No asymptotic assumptions or principal component analysis is required

References

[Cand`es et al., 2011] Cand`es, E. J., Li, X., Ma, Y., and Wright, J. (2011).
Robust principal component analysis
Journal of the ACM, 58(3).

[Chandrasekaran et al., 2012] Chandrasekaran, V., Parillo, P. A., and Willsky, A. S. (2012).
Latent variable graphical model selection via convex optimization.
The Annals of Statistics, 40(4):1935–1967.

[Ma et al., 2013] Ma, S., Xue, L., and Zou, H. (2013).
Alternating direction methods for latent variable gaussian graphical model selection.
Neural Computation, 25:2172–2198.

Lisa Goldberg lrg@berkeley.edu Alex Shkolnik ads2@berkeley.edu