Endogenous Representation of Asset Returns

A new perspective for modeling and estimation

Zhipu Zhou
zhipu@ucsb.edu

UC Berkerly CDAR Seminar

Joint work with Sang-Yun Oh and Alex Shkolnik

Department of Statistics & Applied Probability
Univerisity of California, Santa Barbara

October 27, 2020

Collaborators :

Sang-Yun Oh
Assistant Professor
Department of Statistics & Applied Probability
University of California, Santa Barbara
syoh@ucbs.edu

Alexander Shkolnik
Assistant Professor
Department of Statistics & Applied Probability
University of California, Santa Barbara
shkolnik@ucbs.edu

Zhipu Zhou
Quantitative Researcher
Corporate Model Risk
Wells Fargo
zhipu@ucsb.edu

Background

The origin of asset returns modeling :

- 1950s. Markowitz's Modern Portfolio Theory.
- 1960s. Capital Asset Pricing Model (CAPM).
- 1970s. Arbitrage Pricing Theory (APT).

The origin of asset returns modeling :

- 1950s. Markowitz's Modern Portfolio Theory.
- 1960s. Capital Asset Pricing Model (CAPM).
- 1970s. Arbitrage Pricing Theory (APT).

Equity factor models :

$$Y = \mathbf{B}X + Z$$

- Market (e.g. S&P 500, MSCI World indexes).
- Fundamental (e.g. value, growth, return on equity).
- Macroeconomic (e.g. interest rates, unemployment rate).
- Statistical (e.g. PCA)

Adam Smith's Invisible Hand?

> *We define a factor as a source of common exposure or a characteristic shared by a group of securities that helps explain their risk and return behavior, both individually and collectively. ... we liken factors to many "invisible hands" at work in markets and portfolios.*

- Erianna Khusainova,
Lazzard Asset Management,
The Invisible Hand of Factors (2018).

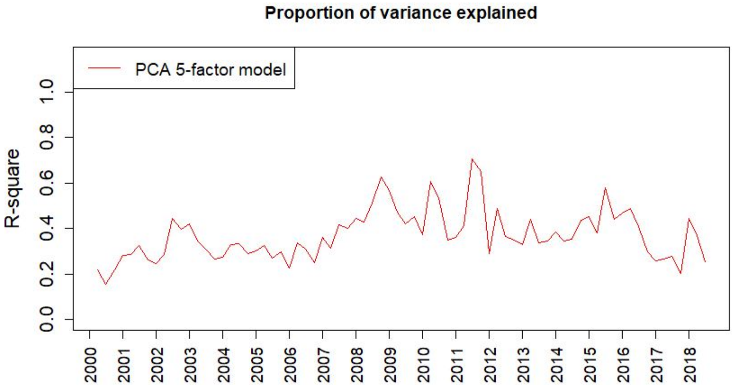
An example using statistical factors (PCA) :

- Gather $p \times n$ data matrix $\mathbf{Y}$ (p assets, n observations).
- Compute a $p \times p$ sample covariance $\mathbf{S} = \mathbf{Y}\mathbf{Y}^\top$.
- Extract 1st factor by maximizing the explained variance

$$\mathscr{s}^2 = \max_{|b|=1} b^\top \mathbf{S} b$$

$\Rightarrow$ The factor $(\mathscr{s}^2, b)$ is empirically "market-like"... it has a large variance $\mathscr{s}^2$ and mostly positive exposures (i.e. most $b_i \geq 0$).

Top 5 principal components and (out-of-sample) explained variance in the returns of the S&P 500 component stocks.



Motivation

Equity factor model (p assets, k factors, n observations) :

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{Z}$$

- $\mathbf{B}$ is a $p \times k$ factor exposure matrix.
- $\mathbf{X}$ is a $k \times n$ factor returns matrix.
- $\mathbf{Z}$ is a $p \times n$ diversifiable returns matrix.

$\Rightarrow$ Universal assumption is $\mathbf{Cov}(X, Z) = 0$.

Equity factor model (p assets, k factors, n observations) :

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{Z}$$

- Fundamental (prescribes **B** and estimates **X**)
- Macroeconomic (prescribes **X** and estimates **B**)
- PCA estimates **BX**

⇒ Factors are often regarded as “exogenous”, i.e., represent a common exposure or characteristic shared by (groups of) assets.

A “good” asset return model will

– explain the risk (variance), and
– explain behavior of return (individually and collectively).

Are factors (invisible hands) the only way to achieve this ?

Our starting point : The spatial-interaction model of Kou et. al (2011)

$$Y = \rho \mathbf{W} Y + \mathbf{B} X + Z$$

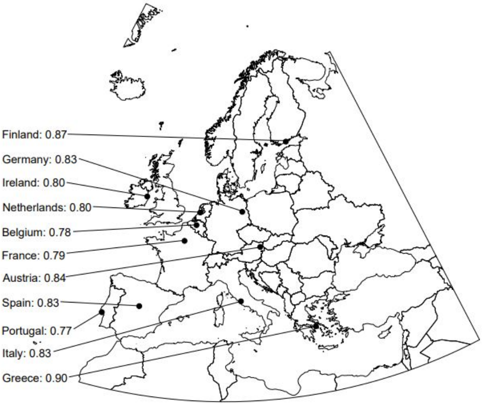
- $\mathbf{W}$ is a $p \times p$ matrix, exogenously given, using quantities related to the location (geographic distance) of assets.
- Describes the asset returns by factors plus distances of assets.
- Not applicable to US equities !

$\Rightarrow$ Ultimately, still a factor model.

Our starting point : The spatial-interaction model of Kou et. al (2011)

$$Y = \rho \mathbf{W} Y + \mathbf{B} X + Z$$

It is effective to model Eurozone stock indices.



Source : *Asset Pricing with Spatial Interaction*. Kou et al. (2011)

Can we take this to its ultimate conclusion ?

No (exogenous) factors. But, decompose asset returns as

$$\text{endogenously explained return} + \text{residual.}$$

Can we take this to its ultimate conclusion ?

No (exogenous) factors. But, decompose asset returns as

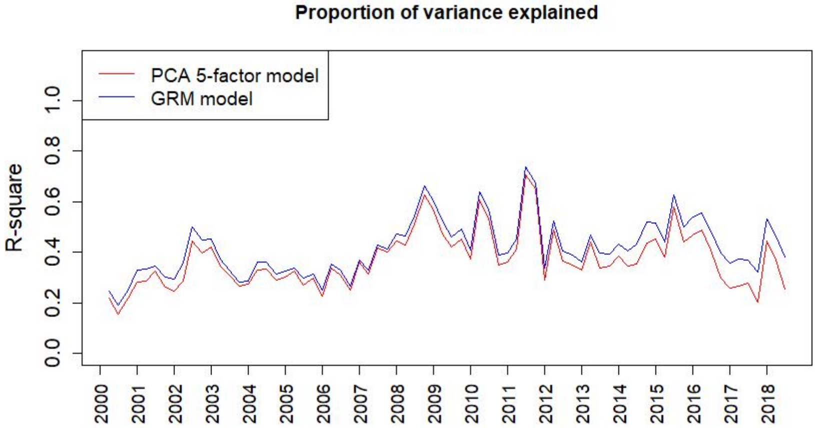
endogenously explained return + residual.

$$Y = \mathbf{A}Y + Z$$

where $\mathbf{A}$ is a sparse $p \times p$ interaction matrix.
$\Rightarrow$ Draws on literature on Graphical Modeling.
$\Rightarrow$ We call it the Graphical Representation Model (GRM).

Explained (out-of-sample) variance of S&P 500 constituents.

PCA(low-dimensional **B**) vs. GRM (sparse **A**).



Theory

Let $Y = (Y_1, Y_2, \cdots, Y_p)^\top \in \mathbb{R}^p$ denotes the returns to p assets. WLOG, assume $\mathbb{E}(Y) = 0$. For any $p \times p$ matrix $\mathscr{A}$ we may define the residual $Z = Z(\mathscr{A})$ as

$$Z = Y - \mathscr{A}Y$$

The i-th row of $\mathscr{A}$ contains coefficients (a_{ij}) such that

$$Z_i = Y_i - \sum_{j \neq i} a_{ij} Y_j \quad \text{for} \quad i = 1, 2, \cdots, p$$

Minimizing the residual is equivalent to solving p linear regression problems of the form

$$\min_{a_{ij}} \mathbb{E}\left(|Y_i - \sum_{j \neq i} a_{ij} Y_j|^2 \right)$$

Theorem

Suppose $Y \in \mathbb{R}^p$ has zero mean and positive definite covariance matrix $\mathbf{\Sigma} = \mathbf{Var}(Y)$ such that $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ exists. Then, there exists a unique minimizer $\mathbf{A}$:

$$\mathbf{A} = \arg \min_{\mathscr{A}:\mathscr{A}_{ii}=0} \mathbb{E}(|Y - \mathscr{A}Y|^2) \tag{1}$$

and the minimizer $\mathbf{A} = \mathbf{I} - \mathbf{D}\mathbf{\Omega}$ where $\mathbf{D} = \mathbf{diag}(\mathbf{\Omega})^{-1}$. That is, if $\mathbf{\Omega} = ((\omega_{ij}))$, then

$$a_{ij} = \begin{cases} -\frac{\omega_{ij}}{\omega_{ii}} & , \textit{for } i \neq j \\ 0 & , \textit{for } i = j \end{cases} \tag{2}$$

Decomposes the return as $Y = \mathbf{A}Y + Z$ where

- The term $\mathbf{A}Y$ is interpreted as the endogenous return.
- Matrix $\mathbf{A}$ encodes the interactions (i.e., between stock issuers).
- $\mathbf{Cov}(Y_i, Z_j) = 0$ for all $i \neq j$ ($\Rightarrow \mathbf{Cov}(\mathbf{A}Y, Z) = 0$).
- $\mathbf{Cov}(Z_i, Z_j) \neq 0$ in general.

Decomposes the return as $Y = \mathbf{A}Y + Z$ where

- The term $\mathbf{A}Y$ is interpreted as the endogenous return.
- Matrix $\mathbf{A}$ encodes the interactions (i.e., between stock issuers).
- $\mathbf{Cov}(Y_i, Z_j) = 0$ for all $i \neq j$ ($\Rightarrow \mathbf{Cov}(\mathbf{A}Y, Z) = 0$).
- $\mathbf{Cov}(Z_i, Z_j) \neq 0$ in general.

A variance decomposition

$$\mathbf{\Sigma} = \mathbf{A\Sigma A}^\top + (\mathbf{I} + \mathbf{A})\mathbf{D}$$

- $\mathscr{E}_Y(\mathbf{A}) = \mathbf{Var}(AY) = \mathbf{A\Sigma A}^\top$ is the endogenously explained variance.
- $(\mathbf{I} + \mathbf{A})\mathbf{D}$ is the remaining variance.
- $\mathbf{Var}(Z) = \mathbf{D\Omega D}$, the residual variance ($\neq (\mathbf{I} + \mathbf{A})\mathbf{D}$).
- $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$ is known as the precision matrix.

The decomposition of variance :

Proposition

If the precision matrix $\mathbf{\Omega} = ((\omega_{ij}))$ *of the random vector* $Y \in \mathbb{R}^p$ *is positive definite (i.e.,* $\omega_{ii} > 0$*), and* $\mathbf{A}$ *is defined in (2), then the variance of* Y_i*, denoted as* σ_{ii}*, can be decomposed as*

$$\sigma_{ii} = \mathbf{Var}(\mathbf{A}Y)_{ii} + \frac{1}{\omega_{ii}} \tag{3}$$

where $\mathscr{E}_Y(\mathbf{A}) = \mathbf{Var}(\mathbf{A}Y)$ *is the endogenous variance matrix,* $\mathbf{Var}(\mathbf{A}Y)_{ii}$ *is the endogenous variance of* Y_i *and* $1/\omega_{ii}$ *is the variance of the residual. In addition, the covariance between* Y_i *and* $Y_j (i \neq j)$*, denoted as* σ_{ij}*, can be decomposed as*

$$\sigma_{ij} = \mathbf{Var}(\mathbf{A}Y)_{ij} + \frac{-\omega_{ij}}{\omega_{ii}\omega_{jj}} \tag{4}$$

An example on $p = 2$ assets with correlation $\rho \in [-1, +1]$.

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 0 & \rho\sigma_1/\sigma_2 \\ \rho\sigma_2/\sigma_1 & 0 \end{pmatrix}$$

The endogenous variance $\mathscr{E}_Y(\mathbf{A}) = \mathbf{Var}(\mathbf{A}Y) = \rho^2\mathbf{\Sigma}$.

- When $|\rho| = 0$, the endogenous variance is zero.
- When $|\rho| = 1$, all variance may be explained endogenously.

An example on $p = 2$ assets with correlation $\rho \in [-1, +1]$.

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 0 & \rho\sigma_1/\sigma_2 \\ \rho\sigma_2/\sigma_1 & 0 \end{pmatrix}$$

The endogenous variance $\mathscr{E}_Y(\mathbf{A}) = \mathbf{Var}(\mathbf{A}Y) = \rho^2\mathbf{\Sigma}$.

- When $|\rho| = 0$, the endogenous variance is zero.
- When $|\rho| = 1$, all variance may be explained endogenously.

The variance of the residual Z is given by

$$\mathbf{Var}(Z) = (1 - \rho^2) \begin{pmatrix} \sigma_1^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

- When $|\rho| = 0$, all variance is explained by the residuals.
- When $|\rho| = 1$, the residual is zero.

The **Graphical Representation Model** (or **GRM**) for the asset returns Y

$$Y = \mathbf{A}Y + E \tag{5}$$

where $\mathbf{A} = \mathbf{I} - \mathbf{D}\boldsymbol{\Omega}, \mathbf{D} = \mathbf{diag}(\boldsymbol{\Omega})^{-1}$ and $\mathbf{A}$ is a sparse matrix.

The matrix $\mathbf{A}$ is assumed to be sparse :

- We posit that much of the variance we see in equity markets may be explained endogenously by sparsely interacting issuers.
- Partial correlations of asset returns may be sparse.
- Stabilize the estimation of model parameters.

$\Rightarrow$ The precision matrix $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$ is sparse.

Equity factor model :

$$Y = \mathbf{B}X + Z \tag{6}$$

where $\mathbf{B} \in \mathbb{R}^{p \times k}$ and $X \in \mathbb{R}^k$. Assume that $\mathbb{E}(XZ^\top) = \mathbf{0}$, then the covariance matrix and the precision matrix of Y given by the factor model are

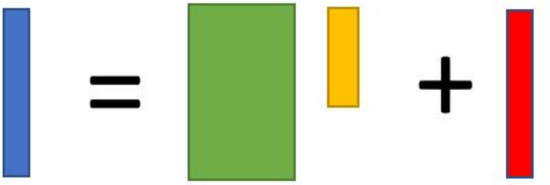
$$\mathbf{\Sigma} = \mathbf{B}\mathbf{V}\mathbf{B}^\top + \mathbf{\Delta} \quad \text{and} \quad \mathbf{\Omega} = \mathbf{\Sigma}^{-1} = \mathbf{\Delta}^{-1} + \mathbf{L} \tag{7}$$

where $\text{rank}(\mathbf{L}) = k$ and is a dense matrix. This is different from GRM model where $\mathbf{\Omega}$ can be sparse.

Model structures

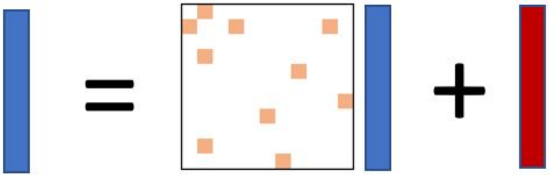
- Factor model

$$Y = \mathbf{B}X + Z$$



- GRM model

$$Y = \mathbf{A}Y + Z$$



Quantity	Graphical Model	Factor Model
Y	$\mathbf{A}Y + Z$	$\mathbf{B}X + Z$
$\mathbf{\Sigma}$	$\mathbf{A\Sigma A}^\top + (\mathbf{I} + \mathbf{A})\mathbf{D}$	$\mathbf{BVB}^\top + \mathbf{\Delta}$
Decomposition	$\mathbf{Cov}(Y, Z) = 0$	$\mathbf{Cov}(X, Z) = 0$
Regularization	sparse $\mathbf{A}$	low-rank $\mathbf{B}$

The Graphical Representation Model performs the minimization

$$\min \mathbb{E}(|Y - \mathscr{A}Y|^2)$$

over $\mathscr{A} \in \mathbb{R}^{p \times p}$ sparse with zero on the diagonal.

The Graphical Representation Model performs the minimization

$$\min \mathbb{E}(|Y - \mathscr{A}Y|^2)$$

over $\mathscr{A} \in \mathbb{R}^{p\times p}$ sparse with zero on the diagonal.

PCA-factor model performs the minimization

$$\min \mathbb{E}(|Y - \mathscr{B}\mathscr{B}^\top Y|^2)$$

over $\mathscr{B} \in \mathbb{R}^{p\times k}$ low-rank $k << p$ with orthonormal columns.

Estimation

The Graphical Representation Model performs the minimization

$$\min_{\mathscr{A}} \mathbb{E}(|Y - \mathscr{A}Y|^2)$$

over $\mathscr{A} \in \mathbb{R}^{p\times p}$ sparse with zero on the diagonal.

- the sparse estimation is actualized via l_1-penalty.
- the procedure estimates a sparse precision matrix $\mathbf{\Omega}$.
- the estimation of $\mathbf{A}$ is computed from $\mathbf{\Omega}$ estimate.

To estimate a sparse matrix $\mathbf{A}$ is equivalent to estimate a sparse matrix $\mathbf{\Omega}$. Several graphical models have been proposed to estimate a sparse $\mathbf{\Omega}$:

$$\begin{aligned} \text{Glasso}: \quad & \hat{\mathbf{\Omega}} = \arg\min_{\mathbf{\Theta} \succeq 0} \{-\log\det(\mathbf{\Theta}) + \mathrm{tr}(\mathbf{S\Theta}) + \lambda \|\mathbf{\Theta}\|_1\} \\ \text{Concord}: \quad & \hat{\mathbf{\Omega}} = \arg\min_{\mathbf{\Theta} \in \mathbb{R}^{p \times p}} \{-\log\det(\mathbf{\Theta}_D^2) + \mathrm{tr}(\mathbf{S\Theta}^2) + \lambda \|\mathbf{\Theta}_F\|_1\} \end{aligned} \tag{8}$$

where $\mathbf{\Theta}_D$ is the diagonal matrix of $\mathbf{\Theta}$, $\mathbf{\Theta}_F = \mathbf{\Theta} - \mathbf{\Theta}_D$, $\mathbf{S}$ is the sample covariance matrix of asset returns. The estimated $\mathbf{A}$ is

$$\hat{\mathbf{A}} = \mathbf{I} - \hat{\mathbf{D}}\hat{\mathbf{\Omega}} \quad \text{where} \quad \hat{\mathbf{D}} = \mathbf{diag}(\hat{\mathbf{\Omega}})^{-1}. \tag{9}$$

Empirics

Testable hypothesis #1

An estimated GRM can explain the (out-of-sample) variance in U.S. equity returns at least as much as a factor model.

Test description :

- Data source : Historical returns of S&P 500 component stocks.
- Time horizon

	Start Date	End Date
During financial crisis period		
In-Sample	2008/01/01	2008/03/31
Out-of-Sample	2008/04/01	2008/07/01
During economic expansion period		
In-Sample	2018/01/01	2018/03/31
Out-of-Sample	2018/04/01	2018/07/01

- Choose 3 factors for factor models, tune $\hat{\mathbf{A}}$ by cross validation.
- Performance measurement : Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{pn}\sum_{i=1}^{n} \|\hat{Y}^{(i)} - Y^{(i)}\|_2^2}$$

GRM explains the out-of-sample variance of equity returns at least as much as a factor model.

Model	Financial Crisis	Expansion
$Y = \mathbf{B}X + Z$ (Fama-French)	1.98	1.42
$Y = \mathbf{B}X + Z$ (PCA Factors)	1.85	1.38
$Y = \mathbf{A}Y + Z$ (Glasso)	1.73	1.26
$Y = \mathbf{A}Y + Z$ (Concord)	1.75	1.27

Testable hypothesis #2

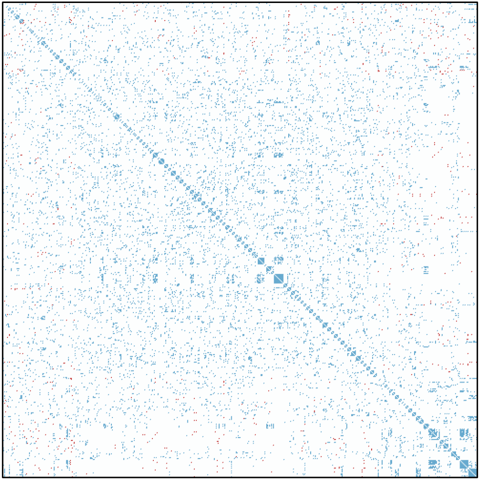
An estimated GRM will exhibit a large degree of sparsity (interpretability) in the interaction among US equity issuers.

Test description :

- Data source : Historical returns of S&P 500 component stocks.
- Time horizon : 2018/01/01 - 2018/03/31.
- Choose Glasso in the model estimation.
- Visualization for $\hat{\mathbf{A}}$.
- Community detection for equities based on sparsity of $\hat{\mathbf{A}}$.

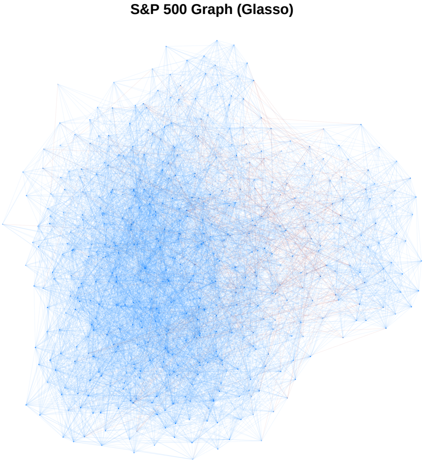
Visualization of sparsity pattern of $\hat{\mathbf{A}}$

- About 95% zero elements.
- Blue/Red : positive/negative partial correlations.



Visualization of sparsity graph of $\hat{\mathbf{A}}$

- About 95% zero elements.
- Blue/Red : positive/negative partial correlations.

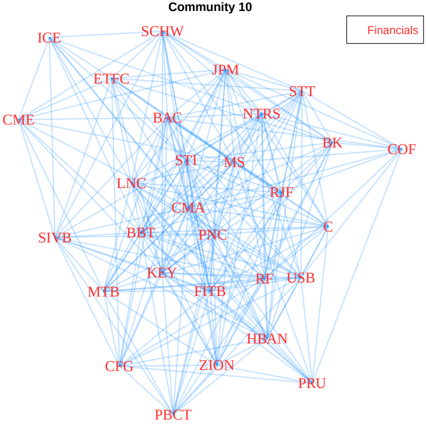


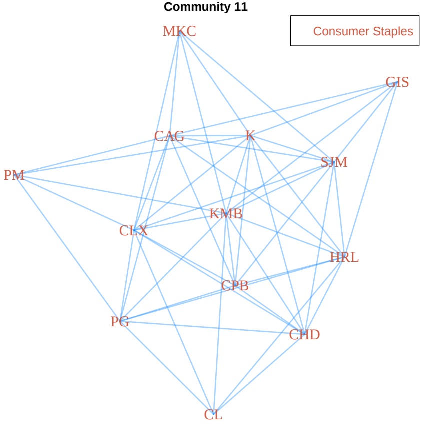
Community detection

- Detect 11 communities solely based on sparsity pattern of $\hat{\mathbf{A}}$.
- In each community, label equities by code. Check overlapping with industry sectors.

Community detection

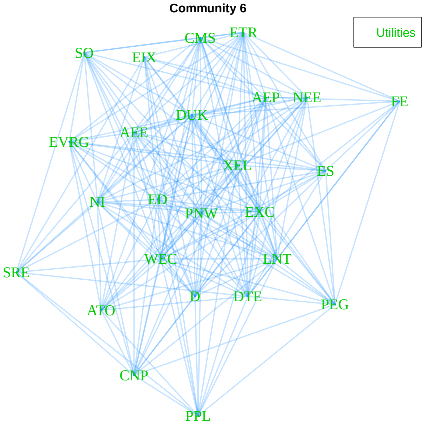
- Detect 11 communities solely based on sparsity pattern of $\hat{\mathbf{A}}$.
- In each community, label equities by code. Check overlapping with industry sectors.

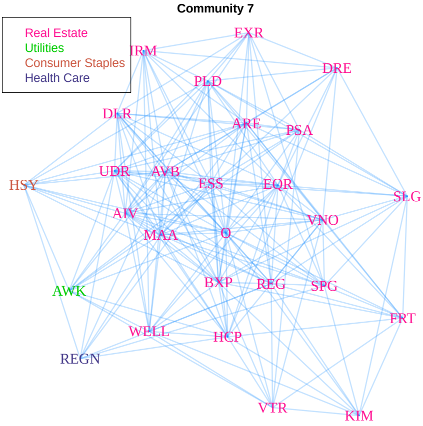




Community detection

- Detect 11 communities solely based on sparsity pattern of $\hat{\mathbf{A}}$.
- In each community, label equities by code. Check overlapping with industry sectors.

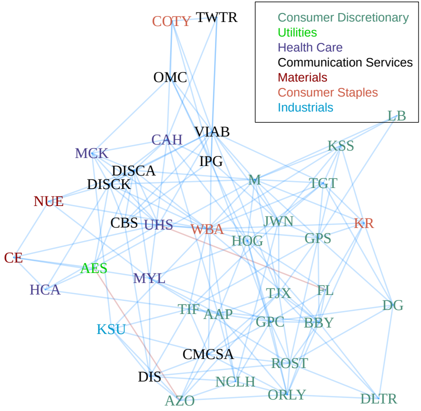




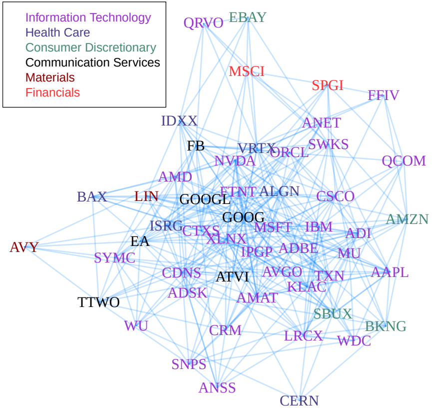
Community detection

- Detect 11 communities solely based on sparsity pattern of $\hat{\mathbf{A}}$.
- In each community, label equities by code. Check overlapping with industry sectors.

Community 3



Community 4



Testable hypothesis #3

The (out-of-sample) variance explained by estimated GRM will not increase with addition of factors.

Test description :

– Idea : test if the following model will deliver better out-of-sample performance

$$Y = \rho \mathbf{A} Y + \mathbf{B} X + Z$$

– X are Fama-French factors.

– $\mathbf{A}$ is defined as in GRM, ρ is a scalar to be estimated together with $\mathbf{B}$.

– Data source : Historical returns of S&P 500 component stocks.

GRM is sufficient to explain the out-of-sample variance of equity returns, adding additional factors does not increase explained variance for equity returns.

Model	Financial Crisis	Expansion
$Y = \rho \mathbf{A} Y + \mathbf{B} X + Z$ (Glasso, Fama-French)	1.74	1.27
$Y = \rho \mathbf{A} Y + \mathbf{B} X + Z$ (Concord, Fama-French)	1.76	1.27
$Y = \mathbf{A} Y + Z$ (Glasso)	1.73	1.26
$Y = \mathbf{A} Y + Z$ (Concord)	1.75	1.27

For more details, please refer to our manuscript on arXiv : https://arxiv.org/abs/2010.13245.

The End