NETS: Network Estimation for Time Series

Matteo Barigozzi Christian Brownlees

London School of Economics

Universitat Pompeu Fabra & Barcelona GSE

Introduction

Motivation

- Network analysis has emerged prominently in many fields of science over the last years:
  Computer Science, Social Networks, Economics, Finance, ...

- This Work:
  The literature on network analysis for multivariate time series is under construction. We propose novel network estimation techniques for the analysis of high–dim multivariate time series

Motivation

- Network analysis has emerged prominently in many fields of science over the last years:
  Computer Science, Social Networks, Economics, Finance, ...

- This Work:
  The literature on network analysis for multivariate time series is under construction. We propose novel network estimation techniques for the analysis of high–dim multivariate time series

Partial Correlation Network (Dempster, 1972; Meinshausen & Bühlmann, 2006)

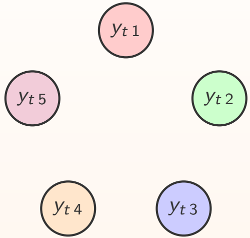
- Consider a white noise process $\mathbf{y}_t = (y_{t\,1}, y_{t\,2}, y_{t\,3}, y_{t\,4}, y_{t\,5})'$
- The network associated with the system is an undirected graph

1. the components of $\mathbf{y}_t$ denote vertices
2. the presence of an edge between i and j denotes that i and j are and the value of the partial correlation measures the strength of the link.

- It is assumed that the network is large yet sparse
- Objective: select nonzero partial correlations and estimate them.

Partial Correlation Network (Dempster, 1972; Meinshausen & Bühlmann, 2006)

- Consider a white noise process $\mathbf{y}_t = (y_{t\,1}, y_{t\,2}, y_{t\,3}, y_{t\,4}, y_{t\,5})'$
- The network associated with the system is an undirected graph

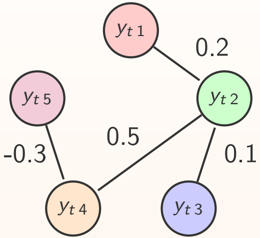


1. the components of $\mathbf{y}_t$ denote vertices
2. the presence of an edge between i and j denotes that i and j are partially correlated and the value of the partial correlation measures the strength of the link.

- It is assumed that the network is large yet sparse
- Objective: select nonzero partial correlations and estimate them.

Partial Correlation Network (Dempster, 1972; Meinshausen & Bühlmann, 2006)

- Consider a white noise process $\mathbf{y}_t = (y_{t\,1}, y_{t\,2}, y_{t\,3}, y_{t\,4}, y_{t\,5})'$
- The network associated with the system is an undirected graph

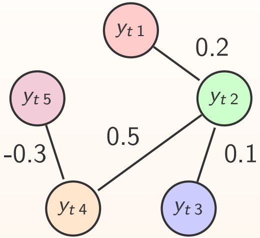


1. the components of $\mathbf{y}_t$ denote vertices
2. the presence of an edge between i and j denotes that i and j are partially correlated and the value of the partial correlation measures the strength of the link.

- It is assumed that the network is large yet sparse
- Objective: select nonzero partial correlations and estimate them.

Partial Correlation Network (Dempster, 1972; Meinshausen & Bühlmann, 2006)

- Consider a white noise process $\mathbf{y}_t = (y_{t\,1}, y_{t\,2}, y_{t\,3}, y_{t\,4}, y_{t\,5})'$
- The network associated with the system is an undirected graph

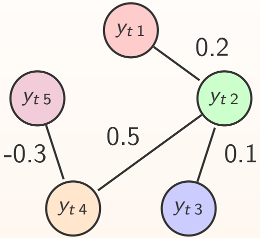


1. the components of $\mathbf{y}_t$ denote vertices
2. the presence of an edge between i and j denotes that i and j are partially correlated and the value of the partial correlation measures the strength of the link.

- It is assumed that the network is large yet sparse
- Objective: select nonzero partial correlations and estimate them.

Partial Correlation Network (Dempster, 1972; Meinshausen & Bühlmann, 2006)

- Consider a white noise process $\mathbf{y}_t = (y_{t\,1}, y_{t\,2}, y_{t\,3}, y_{t\,4}, y_{t\,5})'$
- The network associated with the system is an undirected graph



1. the components of $\mathbf{y}_t$ denote vertices
2. the presence of an edge between i and j denotes that i and j are partially correlated and the value of the partial correlation measures the strength of the link.

- It is assumed that the network is large yet sparse
- Objective: select nonzero partial correlations and estimate them.

Refresher on Partial Correlation

- **Partial Correlation** measures (cross-sect.) linear conditional dependence between $y_{t\,i}$ and $y_{t\,j}$ given on all other variables:

$$\rho^{ij} = \text{Cor}(y_{t\,i}, y_{t\,j} | \{y_{t\,k} : k \neq i, j\}).$$

- Partial Correlation is related to Linear Regression:
  For instance, consider the model

$$y_{1\,t} = c + \beta_{1\,2} y_{2\,t} + \beta_{1\,3} y_{3\,t} + \beta_{1\,4} y_{4\,t} + \beta_{1\,5} y_{5\,t} + u_{1\,t}$$

  β_{13} is different from 0 $\Leftrightarrow$ 1 and 3 are partially correlated

- Partial Correlation is related to Correlation:
  If there is exist a partial correlation path between nodes i and j, then i and j are correlated (and viceversa).

Refresher on Partial Correlation

- Partial Correlation measures (cross-sect.) linear conditional dependence between $y_{t\,i}$ and $y_{t\,j}$ given on all other variables:

$$\rho^{ij} = \mathrm{Cor}(y_{t\,i}, y_{t\,j} | \{y_{t\,k} : k \neq i, j\}).$$

- Partial Correlation is related to Linear Regression:
  For instance, consider the model

$$y_{1\,t} = c + \beta_{1\,2} y_{2\,t} + \beta_{1\,3} y_{3\,t} + \beta_{1\,4} y_{4\,t} + \beta_{1\,5} y_{5\,t} + u_{1\,t}$$

  β_{13} is different from 0 $\Leftrightarrow$ 1 and 3 are partially correlated

- Partial Correlation is related to Correlation:
  If there is exist a partial correlation path between nodes i and j, then i and j are correlated (and viceversa).

Refresher on Partial Correlation

- Partial Correlation measures (cross-sect.) linear conditional dependence between $y_{t\,i}$ and $y_{t\,j}$ given on all other variables:

$$\rho^{ij} = \text{Cor}(y_{t\,i}, y_{t\,j} | \{y_{t\,k} : k \neq i, j\}).$$

- Partial Correlation is related to Linear Regression:
  For instance, consider the model

$$y_{1\,t} = c + \beta_{1\,2} y_{2\,t} + \beta_{1\,3} y_{3\,t} + \beta_{1\,4} y_{4\,t} + \beta_{1\,5} y_{5\,t} + u_{1\,t}$$

  β_{13} is different from 0 $\Leftrightarrow$ 1 and 3 are partially correlated

- Partial Correlation is related to Correlation:
  If there is exist a partial correlation path between nodes i and j, then i and j are correlated (and viceversa).

Limitations of Partial Correlations for Time Series

- Defining the network on the basis of partial correlations is motivated by the analysis of serially uncorrelated Gaussian data.

- However, this is **not** always satisfactory for economic and financial applications where data typically exhibit serial dependence.

  (Partial correlation only captures contemporaneous dependence. However, in economic datasets it is often the case that the realization of the series A in period t might be correlated with the realization of series B in period $t-1$)

- In this work we propose a novel definition of network which overcomes these limitations.

Limitations of Partial Correlations for Time Series

- Defining the network on the basis of partial correlations is motivated by the analysis of serially uncorrelated Gaussian data.

- However, this is **not** always satisfactory for economic and financial applications where data typically exhibit serial dependence.
  (Partial correlation only captures contemporaneous dependence. However, in economic datasets it is often the case that the realization of the series A in period t might be correlated with the realization of series B in period $t-1$)

- In this work we propose a novel definition of network which overcomes these limitations.

Limitations of Partial Correlations for Time Series

- Defining the network on the basis of partial correlations is motivated by the analysis of serially uncorrelated Gaussian data.

- However, this is **not** always satisfactory for economic and financial applications where data typically exhibit serial dependence.
  (Partial correlation only captures contemporaneous dependence. However, in economic datasets it is often the case that the realization of the series A in period t might be correlated with the realization of series B in period $t-1$)

- In this work we propose a novel definition of network which overcomes these limitations.

In this Work

1. We propose **long run** partial correlation network for time series ($\Rightarrow$ partial correlation definition based on the long run covariance)
   - definition captures contemporaneous as well as lead/lag effects
   - model free – it does not hinge on a specific model
   - easy to estimate formulas
2. We propose a network estimation algorithm called NETS
   - two step LASSO regression procedure
   - allows to estimate large networks in seconds
   - we establish conditions for consistent network estimation
3. We illustrate NETS on a panel of monthly equity returns
   - The risk of each asset is decomposed in Systematic and Idiosyncratic components where the Idiosyncratic part has a Network structure

In this Work

1. We propose **long run** partial correlation network for time series ($\Rightarrow$ partial correlation definition based on the long run covariance)
   - definition captures contemporaneous as well as lead/lag effects
   - model free – it does not hinge on a specific model
   - easy to estimate formulas
2. We propose a network estimation algorithm called NETS
   - two step LASSO regression procedure
   - allows to estimate large networks in seconds
   - we establish conditions for consistent network estimation
3. We illustrate NETS on a panel of monthly equity returns
   - The risk of each asset is decomposed in Systematic and Idiosyncratic components where the Idiosyncratic part has a Network structure

In this Work

1. We propose **long run** partial correlation network for time series ($\Rightarrow$ partial correlation definition based on the long run covariance)
   - definition captures contemporaneous as well as lead/lag effects
   - model free – it does not hinge on a specific model
   - easy to estimate formulas
2. We propose a network estimation algorithm called NETS
   - two step LASSO regression procedure
   - allows to estimate large networks in seconds
   - we establish conditions for consistent network estimation
3. We illustrate NETS on a panel of monthly equity returns
   - The risk of each asset is decomposed in Systematic and Idiosyncratic components where the Idiosyncratic part has a Network structure

Related Literature

- ## On Financial Networks and Systemic Risk

  Billio, Getmanksi, Lo, Pellizzon (2012), Diebold and Yilmaz (2013), Hautsch, Schaumburg, Schienle (2011), Dungey, Luciani, Veredas (2012), Bisias, Flood, Lo, Valavanis (2012)

- ## On Graphical Models

  Dempster (1972), Lauritzen (1996), Meinshausen and Bühlmann (2006)

- ## On LASSO / VAR LASSO Estimation

  Tibshirani (1996), Fan and Peng (2004), Zou (2006), Peng, Wang, Zhou, Zhu (2009), Medeiros and Mendes (2012), Kock (2012), Kock and Callot (2012)

- ## On Robust Covariance Estimation

  White (1984), Gallant (1987), Newey and West (1991), Andrews (1991), Andrews and Monahan (1992), Den Haan and Levin (1994)

Network for Time Series

Long Run Partial Correlation Network

- Partial Correlations do not adequately capture cross-sectional dependence if the data has serial dependence

- In this work we propose to construct a measure of partial correlation on the basis of the Long Run Covariance Matrix to overcome this limitation.

- The long run covariance matrix provides a comprehensive and model free measure of cross sectional dependence for serially dependent data. It is defined as

$$\mathbf{y}_t$$

Long Run Partial Correlation Network

- Partial Correlations do not adequately capture cross-sectional dependence if the data has serial dependence

- In this work we propose to construct a measure of partial correlation on the basis of the Long Run Covariance Matrix to overcome this limitation.

- The long run covariance matrix provides a comprehensive and model free measure of cross sectional dependence for serially dependent data. It is defined as

$$\mathbf{y}_t$$

Long Run Partial Correlation Network

- Partial Correlations do not adequately capture cross-sectional dependence if the data has serial dependence

- In this work we propose to construct a measure of partial correlation on the basis of the Long Run Covariance Matrix to overcome this limitation.

- The long run covariance matrix provides a comprehensive and model free measure of cross sectional dependence for serially dependent data. It is defined as

$$\Sigma_L \equiv \lim_{M\to\infty} \frac{1}{M} \text{Var}\left(\sum_{t=1}^{M} \mathbf{y}_t \right)$$

Long Run Covariance

Consider a bivariate system with spillover effects

$$\begin{aligned} y_{t\,1} &= \epsilon_{t\,1} + \psi \epsilon_{t-1\,2} \\ y_{t\,2} &= \epsilon_{t\,2} \end{aligned} \quad \text{with} \quad \begin{aligned} \epsilon_{t\,1} &\sim \mathcal{N}(0, \sigma^2) \\ \epsilon_{t\,2} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

with $\text{Cor}(\epsilon_{t\,1}, \epsilon_{t\,2}) = 0$

Then,

- $\text{Cor}\left(y_{t\,1}, y_{t\,2}\right) = 0$
- $\text{Cor}\left(\sum_{t=1}^{12} y_{t\,1}, \sum_{t=1}^{12} y_{t\,2}\right) = \left(\frac{11}{12}\right) \frac{\psi}{\sqrt{1+\psi^2}}$
- $\lim_{M\to\infty} \text{Cor}\left(\sum_{t=1}^{M} y_{t\,1}, \sum_{t=1}^{M} y_{t\,2}\right) = \frac{\psi}{\sqrt{1+\psi^2}}$

Long Run Covariance

Consider a bivariate system with spillover effects

$$\begin{aligned} y_{t\,1} &= \epsilon_{t\,1} + \psi \epsilon_{t-1\,2} \\ y_{t\,2} &= \epsilon_{t\,2} \end{aligned} \quad \text{with} \quad \begin{aligned} \epsilon_{t\,1} &\sim \mathcal{N}(0, \sigma^2) \\ \epsilon_{t\,2} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

with $\text{Cor}(\epsilon_{t\,1}, \epsilon_{t\,2}) = 0$

Then,

- $\text{Cor}\,(y_{t\,1}, y_{t\,2}) = 0$
- $\text{Cor}\left(\sum_{t=1}^{12} y_{t\,1}, \sum_{t=1}^{12} y_{t\,2}\right) = \left(\frac{11}{12}\right) \frac{\psi}{\sqrt{1+\psi^2}}$
- $\lim_{M\to\infty} \text{Cor}\left(\sum_{t=1}^{M} y_{t\,1}, \sum_{t=1}^{M} y_{t\,2}\right) = \frac{\psi}{\sqrt{1+\psi^2}}$

Long Run Covariance

Consider a bivariate system with spillover effects

$$\begin{aligned} y_{t\,1} &= \epsilon_{t\,1} + \psi \epsilon_{t-1\,2} \\ y_{t\,2} &= \epsilon_{t\,2} \end{aligned} \quad \text{with} \quad \begin{aligned} \epsilon_{t\,1} &\sim \mathcal{N}(0, \sigma^2) \\ \epsilon_{t\,2} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

with $\text{Cor}(\epsilon_{t\,1}, \epsilon_{t\,2}) = 0$

Then,

- $\text{Cor}\left(y_{t\,1}, y_{t\,2}\right) = 0$
- $\text{Cor}\left(\sum_{t=1}^{12} y_{t\,1}, \sum_{t=1}^{12} y_{t\,2}\right) = \left(\frac{11}{12}\right) \frac{\psi}{\sqrt{1+\psi^2}}$
- $\lim_{M\to\infty} \text{Cor}\left(\sum_{t=1}^{M} y_{t\,1}, \sum_{t=1}^{M} y_{t\,2}\right) = \frac{\psi}{\sqrt{1+\psi^2}}$

Long Run Covariance

Consider a bivariate system with spillover effects

$$\begin{aligned} y_{t\,1} &= \epsilon_{t\,1} + \psi \epsilon_{t-1\,2} \\ y_{t\,2} &= \epsilon_{t\,2} \end{aligned} \qquad \text{with} \qquad \begin{aligned} \epsilon_{t\,1} &\sim \mathcal{N}(0, \sigma^2) \\ \epsilon_{t\,2} &\sim \mathcal{N}(0, \sigma^2) \end{aligned}$$

with $\text{Cor}(\epsilon_{t\,1}, \epsilon_{t\,2}) = 0$

Then,

- $\text{Cor}\,(y_{t\,1}, y_{t\,2}) = 0$
- $\text{Cor}\left(\sum_{t=1}^{12} y_{t\,1}, \sum_{t=1}^{12} y_{t\,2}\right) = \left(\frac{11}{12}\right) \frac{\psi}{\sqrt{1+\psi^2}}$
- $\lim_{M\to\infty} \text{Cor}\left(\sum_{t=1}^{M} y_{t\,1}, \sum_{t=1}^{M} y_{t\,2}\right) = \frac{\psi}{\sqrt{1+\psi^2}}$

From LR Covariance to LR Partial Correlation

- Let $\mathbf{K}_L$ denote the inverse of the long run covariance $\mathbf{\Sigma}_L$
  $\mathbf{K}_L$ is also known as the long run concentration matrix
- Let k_{ij} denote the (i, j) element of $\mathbf{K}_L$. The long run partial correlations are

$$\rho_L^{ij} = \frac{-k_{ij}}{\sqrt{k_{ii}k_{jj}}}$$

- The Long Run Partial Correlation network is defined as follows: if $\rho_L^{ij} \neq 0$ then i and j are connected by an edge

From LR Covariance to LR Partial Correlation

- Let $\mathbf{K}_L$ denote the inverse of the long run covariance $\boldsymbol{\Sigma}_L$
  $\mathbf{K}_L$ is also known as the long run concentration matrix
- Let k_{ij} denote the (i,j) element of $\mathbf{K}_L$. The long run partial correlations are

$$\rho_L^{ij} = \frac{-k_{ij}}{\sqrt{k_{ii}k_{jj}}}$$

- The Long Run Partial Correlation network is defined as follows: if $\rho_L^{ij} \neq 0$ then i and j are connected by an edge

From LR Covariance to LR Partial Correlation

- Let $\mathbf{K}_L$ denote the inverse of the long run covariance $\mathbf{\Sigma}_L$
  $\mathbf{K}_L$ is also known as the long run concentration matrix
- Let k_{ij} denote the (i,j) element of $\mathbf{K}_L$. The long run partial correlations are

$$\rho_L^{ij} = \frac{-k_{ij}}{\sqrt{k_{ii}k_{jj}}}$$

- The Long Run Partial Correlation network is defined as follows: if $\rho_L^{ij} \neq 0$ then i and j are connected by an edge

From LR Partial Correlation To LR Concentration

The Long Run Partial Correlation formula:

$$\rho_L^{ij} = \frac{-k_{ij}}{\sqrt{k_{ii}k_{jj}}}$$

where k_{ij} denotes the (i, j) element of $\mathbf{K}_L$.

- This implies that the long run partial correlation network is entirely characterized by $\mathbf{K}_L$!
  If k_{ij} is nonzero, then node i and j are connected by an edge.
- This fact has important implications for estimation:
  We can reformulate the estimation of the long run partial correlation network as the estimation of a sparse long run concentration matrix.

From LR Partial Correlation To LR Concentration

The Long Run Partial Correlation formula:

$$\rho_L^{ij} = \frac{-k_{ij}}{\sqrt{k_{ii} k_{jj}}}$$

where k_{ij} denotes the (i, j) element of $\mathbf{K}_L$.

- This implies that the long run partial correlation network is entirely characterized by $\mathbf{K}_L$!
  If k_{ij} is nonzero, then node i and j are connected by an edge.
- This fact has important implications for estimation:
  We can reformulate the estimation of the long run partial correlation network as the estimation of a sparse long run concentration matrix.

Network Estimation

VAR Approximation

- We employ an estimation strategy which builds up on the classic HAC estimation literature in Econometrics.

- We approximate the $\mathbf{y}_t$ process using a VAR

$$\mathbf{y}_t = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{y}_{t-k} + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim wn(\mathbf{0}, \boldsymbol{\Gamma}_{\epsilon})$$

- The long run concentration matrix of the VAR approximation is

$$\begin{aligned} \mathbf{K}_L &= \left(\mathbf{I} - \textstyle\sum_{k=1}^{p} \mathbf{A}_k'\right) \boldsymbol{\Gamma}_{\epsilon}^{-1} \left(\mathbf{I} - \textstyle\sum_{k=1}^{p} \mathbf{A}_k\right) \\ &= (\mathbf{I} - \mathbf{G}')\,\mathbf{C}\,(\mathbf{I} - \mathbf{G}) \end{aligned}$$

where

- $\mathbf{G} = \sum_{k=1}^{p} \mathbf{A}_k$ – as in Granger
- $\mathbf{C} = \boldsymbol{\Gamma}_{\epsilon}^{-1}$ – as in Contemporaneous

VAR Approximation

- We employ an estimation strategy which builds up on the classic HAC estimation literature in Econometrics.

- We approximate the $\mathbf{y}_t$ process using a VAR

$$\mathbf{y}_t = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{y}_{t-k} + \epsilon_t \quad \epsilon_t \sim wn(\mathbf{0}, \Gamma_\epsilon)$$

- The long run concentration matrix of the VAR approximation is

$$\begin{aligned} \mathbf{K}_L &= \left(\mathbf{I} - \sum_{k=1}^{p} \mathbf{A}'_k\right) \Gamma_\epsilon^{-1} \left(\mathbf{I} - \sum_{k=1}^{p} \mathbf{A}_k\right) \\ &= (\mathbf{I} - \mathbf{G}')\, \mathbf{C}\, (\mathbf{I} - \mathbf{G}) \end{aligned}$$

where

- $\mathbf{G} = \sum_{k=1}^{p} \mathbf{A}_k$ – as in Granger
- $\mathbf{C} = \Gamma_\epsilon^{-1}$ – as in Contemporaneous

VAR Approximation

- We employ an estimation strategy which builds up on the classic HAC estimation literature in Econometrics.

- We approximate the $\mathbf{y}_t$ process using a VAR

$$\mathbf{y}_t = \sum_{k=1}^{p} \mathbf{A}_k \mathbf{y}_{t-k} + \boldsymbol{\epsilon}_t \quad \boldsymbol{\epsilon}_t \sim wn(\mathbf{0}, \boldsymbol{\Gamma}_{\epsilon})$$

- The long run concentration matrix of the VAR approximation is

$$\begin{aligned} \mathbf{K}_L &= \left(\mathbf{I} - \textstyle\sum_{k=1}^{p} \mathbf{A}_k'\right) \boldsymbol{\Gamma}_{\epsilon}^{-1} \left(\mathbf{I} - \textstyle\sum_{k=1}^{p} \mathbf{A}_k\right) \\ &= \left(\mathbf{I} - \mathbf{G}'\right) \mathbf{C} \left(\mathbf{I} - \mathbf{G}\right) \end{aligned}$$

where

- $\mathbf{G} = \sum_{k=1}^{p} \mathbf{A}_k$ – as in Granger
- $\mathbf{C} = \boldsymbol{\Gamma}_{\epsilon}^{-1}$ – as in Contemporaneous

VAR Approximation

- The Long Run Concentration matrix implied by the VAR is

$$\mathbf{K}_L = (\mathbf{I} - \mathbf{G}')\,\mathbf{C}\,(\mathbf{I} - \mathbf{G})$$

  We work under the assumption that the VAR approximation is sparse. This, in turns, determines the sparsity of $\mathbf{G}$, $\mathbf{C}$ and $\mathbf{K}_L$

- Graphical interpretation:
  - The matrix $\mathbf{G}$ can be associated to a long run Granger network (directed) expressing long predictive relations of the system and the matrix $\mathbf{C}$ can be associated to a Contemporaneous partial correlation network of the system innovations
  - The Long Run Partial Correlation network is a (nontrivial) combination of the Granger and Contemporaneous networks

VAR Approximation

- The Long Run Concentration matrix implied by the VAR is

$$\mathbf{K}_L = (\mathbf{I} - \mathbf{G}')\,\mathbf{C}\,(\mathbf{I} - \mathbf{G})$$

  We work under the assumption that the VAR approximation is sparse. This, in turns, determines the sparsity of $\mathbf{G}$, $\mathbf{C}$ and $\mathbf{K}_L$

- Graphical interpretation:
  - The matrix $\mathbf{G}$ can be associated to a long run Granger network (directed) expressing long predictive relations of the system
  - and the matrix $\mathbf{C}$ can be associated to a Contemporaneous partial correlation network of the system innovations
  - The Long Run Partial Correlation network is a (nontrivial) combination of the Granger and Contemporaneous networks

NETS Algorithm

- Our Long Run Partial Corr領ellation Network estimator is based on the sparse estimation of **G** and **C** matrices. Sparse estimation is based on the LASSO.

- We propose an algorithm called "Network Estimator for Time Series" (NETS) to estimate sparse long run partial correlation networks

- The NETS procedure consists of estimation $\mathbf{K}_L$ using a two step LASSO regression:

  1. Estimate $\mathbf{A}$ using Adaptive LASSO (based on pre-est. $\tilde{\mathbf{A}}$) on $\mathbf{y}_t$
  2. Estimate $\mathbf{\Gamma}_\epsilon^{-1}$ using LASSO on estimated residuals $\hat{\epsilon}_t$

NETS Algorithm

- Our Long Run Partial Correllation Network estimator is based on the sparse estimation of **G** and **C** matrices. Sparse estimation is based on the LASSO.

- We propose an algorithm called
  "Network Estimator for Time Series" (NETS)
  to estimate sparse long run partial correlation networks

- The NETS procedure consists of estimation $\mathbf{K}_L$ using a two step LASSO regression:

  1. Estimate $\mathbf{A}$ using Adaptive LASSO (based on pre-est. $\tilde{\mathbf{A}}$) on $\mathbf{y}_t$
  2. Estimate $\mathbf{\Gamma}_\epsilon^{-1}$ using LASSO on estimated residuals $\hat{\epsilon}_t$

NETS Algorithm

- Our Long Run Partial Correllation Network estimator is based on the sparse estimation of **G** and **C** matrices. Sparse estimation is based on the LASSO.
- We propose an algorithm called "Network Estimator for Time Series" (NETS) to estimate sparse long run partial correlation networks
- The NETS procedure consists of estimation $\mathbf{K}_L$ using a two step LASSO regression:
  1. Estimate $\mathbf{A}$ using Adaptive LASSO (based on pre-est. $\widetilde{\mathbf{A}}$) on $\mathbf{y}_t$
  2. Estimate Γ_ϵ^{-1} using LASSO on estimated residuals $\widehat{\epsilon}_t$

NETS Algorithm

- Our Long Run Partial Correllation Network estimator is based on the sparse estimation of **G** and **C** matrices. Sparse estimation is based on the LASSO.
- We propose an algorithm called "Network Estimator for Time Series" (NETS) to estimate sparse long run partial correlation networks
- The NETS procedure consists of estimation $\mathbf{K}_L$ using a two step LASSO regression:
  1. Estimate **A** using Adaptive LASSO (based on pre-est. $\widetilde{\mathbf{A}}$) on $\mathbf{y}_t$
  2. Estimate $\mathbf{\Gamma}_\epsilon^{-1}$ using LASSO on estimated residuals $\widehat{\epsilon}_t$

NETS Algorithm

- Our Long Run Partial Correllation Network estimator is based on the sparse estimation of $\mathbf{G}$ and $\mathbf{C}$ matrices. Sparse estimation is based on the LASSO.

- We propose an algorithm called "Network Estimator for Time Series" (NETS) to estimate sparse long run partial correlation networks

- The NETS procedure consists of estimation $\mathbf{K}_L$ using a two step LASSO regression:

  1. Estimate $\mathbf{A}$ using Adaptive LASSO (based on pre-est. $\widetilde{\mathbf{A}}$) on $\mathbf{y}_t$
  2. Estimate $\boldsymbol{\Gamma}_{\epsilon}^{-1}$ using LASSO on estimated residuals $\widehat{\epsilon}_t$

NETS Steps

Network Estimator for Time Series (NETS) Algorithm: Step 1

Estimate $\mathbf{G}$ with

$$\widehat{\mathbf{G}} = \sum_{i=1}^{p} \widehat{\mathbf{A}}_i$$

where $\widehat{\mathbf{A}}_i$ $(i = 1, \ldots, p)$ are the minimizers of the objective function:

$$\mathcal{L}_T^G(\mathbf{A}_1, \ldots, \mathbf{A}_p) = \sum_{t=1}^{T} \left(\mathbf{y}_t - \sum_{i=1}^{p} \mathbf{A}_i \mathbf{y}_{t-i} \right)^2 + \lambda^G \sum_{i=1}^{p} \frac{|\mathbf{A}_i|}{|\widetilde{\mathbf{A}}_i|}$$

where $|\mathbf{A}_i|$ is equal to the sum of the absolute values of the components of $\mathbf{A}_i$.

NETS Steps

- The concentration matrix of the systems shocks can be estimated via a regression based estimator
- Consider the regression model

$$\epsilon_{t\,i} = \sum_{j=1}^{N} \theta_{ij} \epsilon_{t\,j} (1 - \delta_{ij}) + u_{t\,i}, \qquad i = 1, \ldots, N,$$

where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ii} = 1$

- The regression coefficients and residual variance of the regression is related to the entries of the concentration matrix by the following relations

$$c_{ii} = \frac{1}{\mathrm{Var}(u_{t\,i})}$$

NETS Steps

Network Estimator for Time Series (NETS) Algorithm: Step 2

Consider

$$\widehat{\epsilon}_t = \mathbf{y}_t - \widehat{\mathbf{A}}_i \mathbf{y}_{t-1}$$

and define $\widehat{\mathbf{C}}$ as the LASSO regression based estimator of the concentration matrix obtained by minimizing

$$\mathcal{L}_T^C(\rho) = \left[\sum_{t=1}^{T} \sum_{i=1}^{N} \left(\widehat{\epsilon}_{t\,i} - \sum_{j \neq i}^{N} \rho_{ij} \sqrt{\frac{\hat{c}_{ii}}{\hat{c}_{jj}}} \widehat{\epsilon}_{t\,j} \right)^2 \right] + \lambda^C \sum_{i=2}^{N} \sum_{j=1}^{i-1} |\rho_{ij}|$$

where $\hat{c}_{ii}$, $i = 1, \ldots, N$ is a pre-estimator of the reciprocal of the residual variance of component i

Large Sample Properties: Assumptions I

(Sketch of) Main Assumptions

1. Data is Weakly Dependent

   (cf. Doukan and Louhini (1999), Doukhan and Neumann (2007))

2. Data is Covariance Stationary with pd Spectral Density
3. Truncation error of VAR(∞) model decays sufficiently fast
4. Nonzero coefficients are sufficiently large
5. Pre estimators are well behaved
6. Sparsity structure of the VAR parameters

Large Sample Properties: Assumptions II

(Sketch of) Main Assumptions

1. Problem Dimension: $n_T = O(T^{\zeta_1})$ and $p_T = O(T^{\zeta_2})$ with $\zeta_1, \zeta_2 > 0$.

2. Number of Nonzero Parameters for **G** and **C** networks is at most $o\left(\sqrt{\frac{T}{\log T}}\right)$ and there is a trade–off between the two.

3. Penalties: $\frac{\lambda_T}{T}\sqrt{q_T} = o(1)$ and $\lim_{T\to\infty} \lambda_T \sqrt{\frac{T}{\log T}} = \infty$

Large Sample Properties

Proposition I: Granger Network

(a) (Consistent Selection) The probability of correctly selecting the nonzero coefficients of **G** converges to one.

(b) (Consistent Estimation) For T sufficiently large and each $\eta > 0$ there exists a κ such that with at least probability $1 - O(T^{-\eta})$

$$||\mathbf{G} - \widehat{\mathbf{G}}||_2 \leq \kappa \frac{\lambda_T}{T} \sqrt{q_T}$$

Large Sample Properties

Proposition II: Contemporaneous Network

(a) (Consistent Selection) The probability of correctly selecting the nonzero coefficients of $\mathbf{C}$ converges to one.

(b) (Consistent Estimation) For T sufficiently large and each $\eta > 0$ there exists a κ such that with at least probability $1 - O(T^{-\eta})$

$$||\mathbf{C} - \widehat{\mathbf{C}}||_2 \leq \kappa \frac{\lambda_T}{T} \sqrt{q_T}$$

Large Sample Properties

Corollary: Long Run Partial Correlation Network

(a) (Consistent Selection) The probability of correctly selecting the nonzero coefficients of $\mathbf{K}_L$ converges to one.

(b) (Consistent Estimation)

$$\widehat{\mathbf{K}}_L \xrightarrow{p} \mathbf{K}_L$$

Empirical Application

Empirical Application

- We interested in estimating the network of the idiosyncratic component in a panel of equity returns of top U.S. companies
- Monthly log returns for 41 U.S. bluechips between 1990 and 2010 (252 observations)
- Application inspired by Billio, Getmanksi, Lo, Pellizzon (2012)

Empirical Application

- We interested in estimating the network of the idiosyncratic component in a panel of equity returns of top U.S. companies
- Monthly log returns for 41 U.S. bluechips between 1990 and 2010 (252 observations)
- Application inspired by Billio, Getmanksi, Lo, Pellizzon (2012)

Empirical Application

- We interested in estimating the network of the idiosyncratic component in a panel of equity returns of top U.S. companies
- Monthly log returns for 41 U.S. bluechips between 1990 and 2010 (252 observations)
- Application inspired by Billio, Getmanksi, Lo, Pellizzon (2012)

Empirical Application

- Consider a simple one factor model for a panel equity returns

$$\mathbf{r}_{it} = \beta_i r_{m\,t} + \epsilon_{it} \quad t = 1, \ldots, T \;\; i = 1, \ldots, N$$

with $\mathrm{E}(\epsilon_{it}) = 0$ and $\mathrm{Var}(\epsilon_{it}) = \sigma_{it}$

- The model allows to decompose the risk of an asset in a systematic component that depends on the common factor $r_{m\,t}$ and idiosyncratic component ϵ_{it}

- We are going to construct the series of factor residuals $\hat{\epsilon}_{it}$ and use NETS to estimate the network of the idiosyncratic component.

Empirical Application

- Consider a simple one factor model for a panel equity returns

$$\mathbf{r}_{it} = \beta_i r_{m\,t} + \epsilon_{it} \quad t = 1, ..., T \ i = 1, ..., N$$

with $\mathrm{E}(\epsilon_{it}) = 0$ and $\mathrm{Var}(\epsilon_{it}) = \sigma_{it}$

- The model allows to decompose the risk of an asset in a systematic component that depends on the common factor $r_{m\,t}$ and idiosyncratic component ϵ_{it}

- We are going to construct the series of factor residuals $\hat{\epsilon}_{it}$ and use NETS to estimate the network of the idiosyncratic component.

Empirical Application

- Consider a simple one factor model for a panel equity returns

$$\mathbf{r}_{it} = \beta_i r_{m\,t} + \epsilon_{it} \quad t = 1, \ldots, T \; i = 1, \ldots, N$$

with $\mathrm{E}(\epsilon_{it}) = 0$ and $\mathrm{Var}(\epsilon_{it}) = \sigma_{it}$

- The model allows to decompose the risk of an asset in a systematic component that depends on the common factor $r_{m\,t}$ and idiosyncratic component ϵ_{it}

- We are going to construct the series of factor residuals $\hat{\epsilon}_{it}$ and use NETS to estimate the network of the idiosyncratic component.

Empirical Application: Results

- We run NETS to estimate the long run partial correlation network. Order p of the VAR is one. The tuning parameters λ^G and λ^C are chosen using a BIC type information criterion.

- Out of 820 possible edges, we find 57 edges. The dynamics account for 12% of the edges and the remaining 88% are contemporaneous.

- The idiosyncratic network accounts for a significant portion of the risk of an asset!
  - The market explains on average 25% of the variability of the stocks in the panel
  - The linkages identified by the idiosyncratic network account on average for an additional 15%

Empirical Application: Results

- We run NETS to estimate the long run partial correlation network. Order p of the VAR is one. The tuning parameters λ^G and λ^C are chosen using a BIC type information criterion.

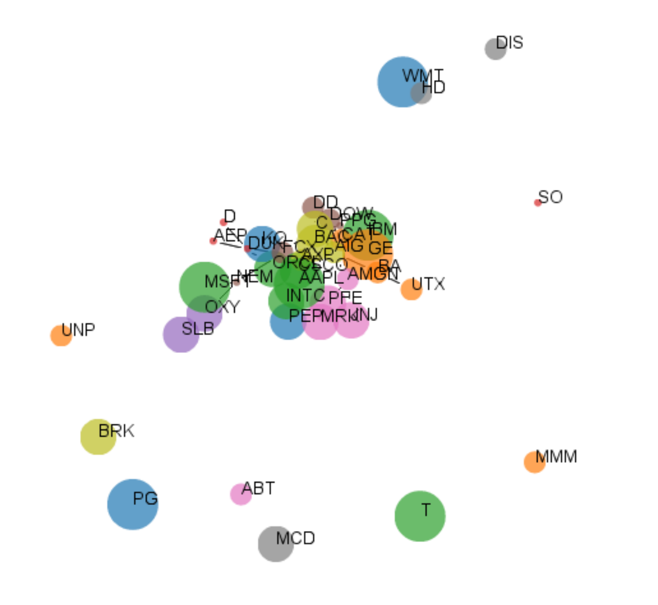
- Out of 820 possible edges, we find 57 edges. The dynamics account for 12% of the edges and the remaining 88% are contemporaneous.

- The idiosyncratic network accounts for a significant portion of the risk of an asset!
  - The market explains on average 25% of the variability of the stocks in the panel
  - The linkages identified by the idiosyncratic network account on average for an additional 15%

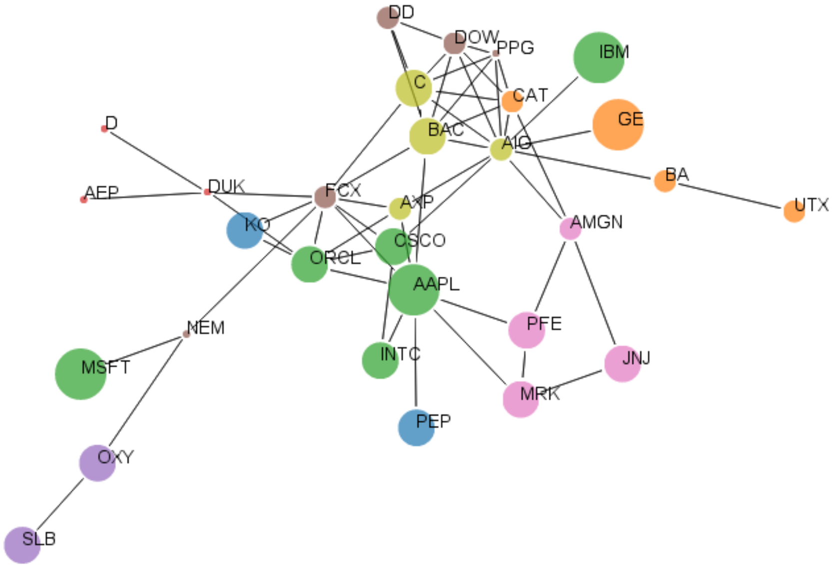
Empirical Application: Results

- We run NETS to estimate the long run partial correlation network. Order p of the VAR is one. The tuning parameters λ^G and λ^C are chosen using a BIC type information criterion.
- Out of 820 possible edges, we find 57 edges. The dynamics account for 12% of the edges and the remaining 88% are contemporaneous.
- The idiosyncratic network accounts for a significant portion of the risk of an asset!
  - The market explains on average 25% of the variability of the stocks in the panel
  - The linkages identified by the idiosyncratic network account on average for an additional 15%

Idiosyncratic Risk Network



Idiosyncratic Risk Network – Zoom



"Reading" the Idiosyncratic Risk Network

- Reading a Network can be challenging at times. There is lot of information encoded in the network.

- We are going to use some of the tools used in network analysis to summarise the information in the network. It turns out that the idiosyncratic risk network shares many of the characteristic of social networks.
  1. Similarity. Nodes that are similar are linked (industry linkages)
  2. Centrality. Financials and Technology are some of the most central sectors.
  3. Clustering. Evidence of "Small World Effects"

"Reading" the Idiosyncratic Risk Network

- Reading a Network can be challenging at times. There is lot of information encoded in the network.
- We are going to use some of the tools used in network analysis to summarise the information in the network. It turns out that the idiosyncratic risk network shares many of the characteristic of social networks.
  1. Similarity. Nodes that are similar are linked (industry linkages)
  2. Centrality. Financials and Technology are some of the most central sectors.
  3. Clustering. Evidence of "Small World Effects"

"Reading" the Idiosyncratic Risk Network

- Reading a Network can be challenging at times. There is lot of information encoded in the network.

- We are going to use some of the tools used in network analysis to summarise the information in the network. It turns out that the idiosyncratic risk network shares many of the characteristic of social networks.
  1. Similarity. Nodes that are similar are linked (industry linkages)
  2. Centrality. Financials and Technology are some of the most central sectors.
  3. Clustering. Evidence of "Small World Effects"

"Reading" the Idiosyncratic Risk Network

- Reading a Network can be challenging at times. There is lot of information encoded in the network.

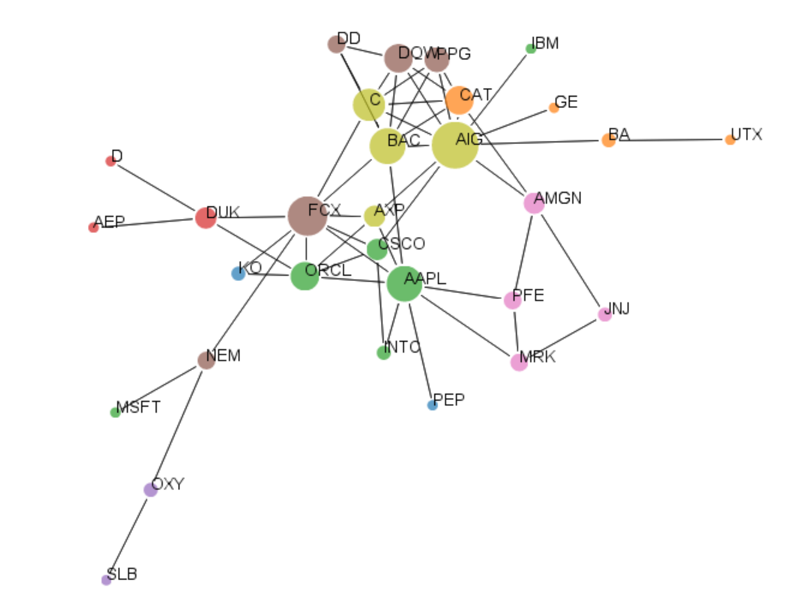
- We are going to use some of the tools used in network analysis to summarise the information in the network. It turns out that the idiosyncratic risk network shares many of the characteristic of social networks.
  1. Similarity. Nodes that are similar are linked (industry linkages)
  2. Centrality. Financials and Technology are some of the most central sectors.
  3. Clustering. Evidence of "Small World Effects"

"Reading" the Idiosyncratic Risk Network

- Reading a Network can be challenging at times. There is lot of information encoded in the network.

- We are going to use some of the tools used in network analysis to summarise the information in the network. It turns out that the idiosyncratic risk network shares many of the characteristic of social networks.
  1. Similarity. Nodes that are similar are linked (industry linkages)
  2. Centrality. Financials and Technology are some of the most central sectors.
  3. Clustering. Evidence of "Small World Effects"

Centrality


DD
DOW
PPG
IBM
C
CAT
GE
BAC
AIG
BA
UTX
D
DUK
FCX
AXP
AMGN
AEP
CSCO
KO
ORCL
AAPL
PFE
JNJ
INTC
MRK
NEM
PEP
MSFT
OXY
SLB

Conclusions

Conclusions

- We propose a network definition for multivariate time series based of long run partial correlations

- We introduce an algorithm called NETS to estimate sparse long run partial correlation networks in potentially large systems and provide large sample analysis of its properties

- We apply this methodology to study the network of idiosyncratic shocks in a panel of financial returns.
  Results show that:
  - The linkages identified by the idiosyncratic risk network explain a conspicuous part of overall variation in returns
  - The idiosyncratic risk network shares many features of social network

Conclusions

- We propose a network definition for multivariate time series based of long run partial correlations

- We introduce an algorithm called NETS to estimate sparse long run partial correlation networks in potentially large systems and provide large sample analysis of its properties

- We apply this methodology to study the network of idiosyncratic shocks in a panel of financial returns.
  Results show that:
  - The linkages identified by the idiosyncratic risk network explain a conspicuous part of overall variation in returns
  - The idiosyncratic risk network shares many features of social network

Conclusions

- We propose a network definition for multivariate time series based of long run partial correlations
- We introduce an algorithm called NETS to estimate sparse long run partial correlation networks in potentially large systems and provide large sample analysis of its properties
- We apply this methodology to study the network of idiosyncratic shocks in a panel of financial returns.
  Results show that:
  - The linkages identified by the idiosyncratic risk network explain a conspicuous part of overall variation in returns
  - The idiosyncratic risk network shares many features of social network

Questions?

Thanks!