Blind Signal Separation & Multivariate Non-Linear Dependency Measures

Peter Martey Addo
(joint with Hayette Gatfaoui & Philippe de Peretti)
Centre National De La Recherche Scientifique
Université Paris1 Panthéon-Sorbonne
NEOMA Business School

The Consortium for Systemic Risk Analytics (CSRA)
Massachusetts Institute of Technology (MIT)
Cambridge, Massachusetts

December 15, 2014

Outline: a package of (2) methodologies ☺

1. Part I: Causal dependencies in multivariate time series
   - Time series graphs & information theoretic measure
2. Part II: Tracking dependency in large multivariate financial systems.
   - Coupling & decoupling information

SYRTO

SYstemic **R**isk **TO**mography

Signals, Measurements, Transmission Channels, and Policy Interventions

Part I: Multivariate Dependency Measures

"The kiss of information theory that captures systemic risk tomography" (joint with Philippe de Peretti)

- Detection and quantification of causal dependencies in multivariate time series
- How do we create a formal measure of systemic risk that adequately captures complex linkages in the financial system? ☺
- Granger causality & transfer entropy ??

Drawbacks of Granger causality & transfer entropy ☹

- Transfer entropy is the information-theoretic analogue of Granger causality
- It reduces to Granger causality for vector auto-regressive processes
- It is advantageous for the analysis of non-linear signals where the model assumption of Granger causality might not hold.
- The transfer entropy is not uniquely determined by the interaction of the two components alone and depends on misleading effects such as autodependency and interaction with other process.
- It requires arbitrary truncation during estimation, it usually requires more samples for accurate estimation
- Can lead to false interpretation since it is not lag-specific.

Our Offer ☺

- A novel information theoretic model-free approach to understanding systemic risk
- To detect and quantify causal dependencies from multivariate time series.
- It gives similar scores to equally noisy dependencies.
- It is uniquely determined by the interaction of the two components alone and in a way autonomous of their interaction with the remaining process.
- Excludes the misleading influence of autodependency within a process
- Lag-specific. Enhance better interpretation.
- Only requires that the multivariate time series be stationary.

Methodology

Let $\mathfrak{X}$ be a multivariate time series with a set of subprocesses V at each time $t \in \mathbb{Z}$ and directional links be defined in E. Then

$$\mathcal{G} = (V \times \mathbb{Z}, E)$$

is the time series graph of $\mathfrak{X}$, where the set of nodes in the graph are made up of V.

- Like graphical models (Lauritzen 1996), TSG's are based on the concept of conditional independence.
- Note that the time-dependence in the time series is used to define directional links in the graph.

Definition (Lag-specific directed link)

We say that the nodes $\mathcal{A}_{t-\tau} \in \mathcal{G}$ and $\mathcal{B}_t \in \mathcal{G}$ are connected by a lag-specific directed link "$\mathcal{A}_t \longrightarrow^{\tau} \mathcal{B}_t$" pointing forward in time if and only if $\tau > 0$ and

$$I_{\mathcal{A} \xrightarrow{link} \mathcal{B}}(\tau) \equiv I\Big(\mathcal{A}_{t-\tau}; \mathcal{B}_t \mid \mathfrak{X}^{-} \setminus \{\mathcal{A}_{t-\tau}\}\Big) > 0. \quad (1)$$

Thus, $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$ are connected if they are not independent conditionally on the past of the whole process excluding $\{\mathcal{A}_{t-\tau}\}$ (denoted by the symbol $\setminus$) which implies a lag-specific causality with respect to $\mathfrak{X}$. $I\Big(\cdot; \cdot | \cdot\Big)$ denotes conditional mutual information.

- If $\mathcal{A} \neq \mathcal{B}$ then "$\mathcal{A}_t \longrightarrow^{\tau} \mathcal{B}_t$" represents a coupling at lag τ.
- An autodependency at lag τ corresponds to $\mathcal{A} = \mathcal{B}$.

Definition (Undirected contemporaneous link)

The nodes $\mathcal{A}_t \in \mathcal{G}$ and $\mathcal{B}_t \in \mathcal{G}$ are connected by an undirected contemporaneous link "$\mathcal{A} - \mathcal{B}$" if and only if

$$I_{\mathcal{A}\overset{link}{-}\mathcal{B}} \equiv I\Big(\mathcal{A}_t; \mathcal{B}_t \mid \mathfrak{X}_{t+1}^{-} \backslash \{\mathcal{A}_t, \mathcal{B}_t\}\Big) > 0. \quad (2)$$

Definition (Notion of *parents* and *neighbors* of subprocesses.)

Given the nodes $\mathcal{A}_t \in \mathcal{G}$ and $\mathcal{B}_t \in \mathcal{G}$, the *parents* $\mathcal{P}_{\mathcal{B}_t}$ and *neighbors* $\mathcal{N}_{\mathcal{B}_t}$ of node $\mathcal{B}_t$ are defined as

$$\mathcal{P}_{\mathcal{B}_t} \equiv \{\mathcal{A}_{t-\tau} : \mathcal{A} \in \mathfrak{X}, \tau > 0, \mathcal{A}_{t-\tau} \longrightarrow \mathcal{B}_t\} \quad (3)$$

$$\mathcal{N}_{\mathcal{B}_t} \equiv \{\mathcal{A}_t : \mathcal{A} \in \mathfrak{X}, \mathcal{A}_t - \mathcal{B}_t\} \quad (4)$$

Definition (Causal Markov Condition)

Any node $\mathcal{B}_t \in \mathcal{G}$ in the time series graph is conditionally independent of $\mathfrak{X}_t^{-} \backslash \mathcal{P}_{\mathcal{B}_t}$ given its *parents* $\mathcal{P}_{\mathcal{B}_t}$.

Definition (Momentary Information Transfer (MIT) Links)

For two subprocesses $\mathcal{A}$, $\mathcal{B}$ of a stationary multivariate discrete time process $\mathfrak{X}$ with parents $\mathcal{P}_{\mathcal{A}_t}$ and $\mathcal{P}_{\mathcal{B}_t}$ in the associated time series graph and $\tau > 0$, the general information theoretic measure between nodes $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$ is given by

$$\begin{aligned} I_{\mathcal{A} \xrightarrow{MIT} \mathcal{B}}(\tau) &\equiv I\Big(\mathcal{A}_{t-\tau}; \mathcal{B}_t \mid \mathcal{P}_{\mathcal{B}_t} \backslash \{\mathcal{A}_{t-\tau}\}, \mathcal{P}_{\mathcal{A}_{t-\tau}}\Big) > 0 \\ &= H\Big(\mathcal{B}_t \mid \mathcal{P}_{\mathcal{B}_t} \backslash \{\mathcal{A}_{t-\tau}\}, \mathcal{P}_{\mathcal{A}_{t-\tau}}\Big) - H\Big(\mathcal{B}_t \mid \mathcal{P}_{\mathcal{B}_t}\Big) \end{aligned}$$

and contemporaneous MIT defined by

$$I_{\mathcal{A} \overset{MIT}{-} \mathcal{B}} \equiv I\Big(\mathcal{A}_t; \mathcal{B}_t \mid \mathcal{P}_{\mathcal{B}_t}, \mathcal{P}_{\mathcal{A}_t}, \mathcal{N}_{\mathcal{A}_t} \backslash \{\mathcal{B}_t\}, \mathcal{N}_{\mathcal{B}_t} \backslash \{\mathcal{A}_t\}, \mathcal{P}_{\mathcal{N}_{\mathcal{A}_t} \backslash \{\mathcal{B}_t\}}, \mathcal{P}_{\mathcal{N}_{\mathcal{B}_t} \backslash \{\mathcal{A}_t\}}\Big). \quad (5)$$

where $H(\mathcal{X})$ is Shannon's entropy and $H(\mathcal{X}|\mathcal{Y})$ denotes conditional Shannon's entropy.

The *parents* of all subprocesses in $\mathfrak{X}$ together with the contemporaneous links forms the time series graph.

Definition (Coupling Strength)

The partial correlation MIT measure, denoted ρ^{MIT}, associated with equation (5), for the strength of coupling mechanism between $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$ is given by

$$\rho_{\mathcal{A}\xrightarrow{MIT}\mathcal{B}}(\tau) \equiv \rho\Big(\mathcal{A}_{t-\tau}; \mathcal{B}_t \mid \mathcal{P}_{\mathcal{B}_t} \backslash \{\mathcal{A}_{t-\tau}\}, \mathcal{P}_{\mathcal{A}_{t-\tau}}\Big). \quad (6)$$

- The measure ρ^{MIT} quantifies how much the variability in $\mathcal{A}$ at the exact lag τ directly influences $\mathcal{B}$, irrespective of the past of $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$.
- ρ^{MIT} is the cross-correlation of the residual after $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$ have been regressed on both the parents of $\mathcal{A}_{t-\tau}$ and $\mathcal{B}_t$.
- The contemporaneous MIT in the linear case is equivalent to the partial correlation of the errors after regressing each process on its *parents*.
- Unlike classical statistics, interactions in the framework of information theory are viewed as transfers of information and thus this approach is model-free.

Estimation

Summary

- In the first Step, PC algorithm (Spirtes et al [SSCR 1991]) is used to estimate the parents of each process, i.e., as a variable selection method.
  - Unlike graphical models, only undirected links are inferred and the second step of PC algorithm is omitted.
  - This first step determines the existence or absence of a link, which also provide useful information on the causality between lagged components of the multivariate process.
- In the second step, MIT is used and all possible links are tested again.
  - In this step, the problem of serial dependencies is drastically reduced using MIT (Runge et al[Physical Review E, 2012]).

Simulation Exercise

Consider a simulated 1000 points of a stationary multivariate autoregressive process made up of four subprocesses $\{X_t, Y_t, Z_t, W_t\}'$ defined by

$$X_t = aX_{t-1} + cZ_{t-4} + \varepsilon_x \quad (7)$$
$$Y_t = kX_{t-1} + hY_{t-1} + \varepsilon_y \quad (8)$$
$$Z_t = dY_{t-2} + bZ_{t-1} + fW_{t-1} + \varepsilon_z \quad (9)$$
$$W_t = eY_{t-3} + gW_{t-1} + \varepsilon_w \quad (10)$$

and the innovation covariance matrix given by $\Sigma_\varepsilon = \begin{pmatrix} 1 & 0 & d & 0 \\ 0 & 1 & 0 & d \\ d & 0 & 1 & 0 \\ 0 & d & 0 & 1 \end{pmatrix}$, where $a = 0.6, b = 0.4, c = 0.3, d = -0.3, e = -0.6, f = 0.2, g = 0.4, k = 0.3$, and $h = 0.6$.

Notice that the lagged causal chain for this process is $X \longrightarrow^1 Y \longrightarrow^2 Z$ with feedback $Z \longrightarrow^4 X$, and $Y \longrightarrow^3 W \longrightarrow^1 Z$, plus contemporaneous links $X - Z$ and $Y - W$.

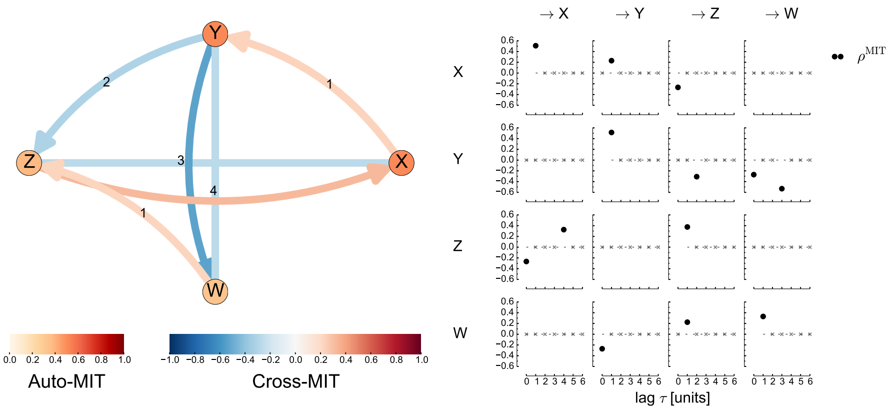

Figure : *MIT plot* for the simulated process. The plot of significant lags for simulated process associated with the *MIT plot*.

The results indicates a lagged causal chain as $X \longrightarrow^1 Y \longrightarrow^2 Z$ with feedback $Z \longrightarrow^4 X$, and $Y \longrightarrow^3 W \longrightarrow^1 Z$, plus contemporaneous links $X - Z$ and $Y - W$.

Research-in-progress

- Develop nonlinear equivalent measure of the coupling strength.
- Empirical application: Causal dependencies in financial institutions
- Further Application: Sovereign CDS, Credit Risk etc

Part II: Blind Signal Separation

Tracking dependency in large multivariate financial systems.

- Track dependency in large multivariate financial systems.
    - Study the time-varying information coupling and decoupling
    - Deduce measures of excess dependency, frailty of the market or multivariate financial risk
    - Deduce Early Warning Indicators.
- Build dependency measures for multivariate systems that exhibit :
    - Rapid changing dynamics, i.e. exhibit a high degree of non-linearity
    - Non-Gaussianity
    - Non-stationarity.

Independent Component Analysis (ICA) ☺

- Assume that the dynamics of the multivariate signal is driven by latent independent non-gaussian signals

- Let $\{\mathfrak{X}_t\}_{t=1}^{T}$ be a multivariate process with $d \in \mathbb{Z}^+$ components such that

$$\mathfrak{X}_t = \mathcal{A}\mathcal{S}_t \tag{11}$$

  where $\mathfrak{X}_t$ is the observed process, $\mathcal{S}_t$ are the unobserved independent signals. $\widehat{\mathcal{S}}_t$ signals, as well as the unmixing matrix $\mathcal{W}$:

$$\widehat{\mathcal{S}}_t = \mathcal{W}\mathfrak{X}_t$$

- Thus $\mathcal{W} = \mathcal{A}^{-1}$ will contain all relevant information about the dependency structure of the system, revealed by off-diagonal elements.

- The Idea is to build information coupling measures entirely based on $\mathcal{W}$.
- Here we have full information decoupling if $\mathcal{W}$ is the identity matrix, and information coupling between some components if some elements in rows are non-zero.
- Then given this measures, use rolling windows to study how the coupling information evolves over time. Uses it to develep new risk measures or early warnings.
- The mixing process $\mathcal{A}$ might change with time - thus a dynamic mixing process can arise due to regime changes or structural changes.
- Note that mixing of the latent signals does not need to be linear.

Research continues

"If you're not prepared to be wrong, you'll never come up with anything original."
(Sir Ken Robinson at TED 2006)

EU's Seventh Framework Programme (SYRTO Project)

This project has received funding from the European Union's Seventh Framework Programme (FP7-SSH/2007-2013) for research, technological development and demonstration under grant agreement *n*°320270 (SYRTO)

SYRTO

SYstemic **R**isk **TO**mography

Signals, Measurements, Transmission Channels, and Policy Interventions

Thank you for your attention ☺