LMS Durham Symposium
Mathematical Aspects of Graphical Models

Learning Causal Structures in Multivariate Time Series

Michael Eichler

Department of Quantitative Economics
University of Maastricht

7 July 2008

Outline

- Graphical representations of time series
  - undirected graphs
  - mixed graphs
- Markov properties
- Representations of systems affected by latent variables
- Identification of causal structure
- Modelling systems with latent variables
- Summary

Multivariate Time Series

Consider multivariate time series

$$X_V = \big(X_V(t)\big)_{t\in\mathbb{Z}}, \qquad X_V(t) = \big(X_v(t)\big)_{v\in V}.$$

Assumptions:

- X_V is Gaussian process
- $\mathbb{E}\big(X_V(t)\big) = 0$
- $\Gamma(u) = \mathbb{E}\big(X_V(t)X_V(t)^{\mathsf{T}}\big)$

$\Big\}$ stationary process

Thus our model is

$$X_V \sim \mathcal{N}(0, \Gamma), \qquad \Gamma = \big(\Gamma(u-v)\big)_{u,v\in\mathbb{Z}}.$$

Undirected Graphs

Let $X_V \sim \mathcal{N}(0, \Gamma)$ with $\Gamma = \big(\Gamma(u-v)\big)_{u,v\in\mathbb{Z}}$.

Covariance graphs:

$$\Gamma_{ij}(u) = 0 \qquad \Rightarrow \qquad X_i(t) \perp\!\!\!\perp X_j(t-u) \quad \forall t \in \mathbb{Z}$$

Undirected Graphs

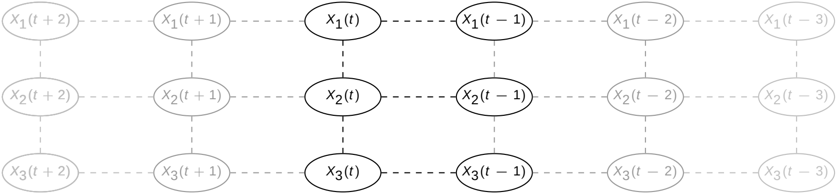
Let $X_V \sim \mathcal{N}(0, \Gamma)$ with $\Gamma = \big(\Gamma(u-v)\big)_{u,v\in\mathbb{Z}}$.

Covariance graphs:

$$\Gamma_{ij}(u) = 0 \qquad \Rightarrow \qquad X_i(t) \perp\!\!\!\perp X_j(t-u) \quad \forall t \in \mathbb{Z}$$

define covariance graph $G = (V \times \mathbb{Z}, E)$

$$(t,i) \text{ --- } (s,j) \notin E \quad \Leftrightarrow \quad \Gamma_{ij}(t-s) = 0$$



Undirected Graphs

Let $X_V \sim \mathcal{N}(0, \Gamma)$ with $\Gamma = \big(\Gamma(u-v)\big)_{u,v\in\mathbb{Z}}$.

Covariance selection graphs:

let $\Gamma^{(\mathrm{i})} = \Gamma^{-1}$ be the inverse covariance matrix

$$\Gamma^{(\mathrm{i})}_{ij}(u) = 0 \qquad \Rightarrow \qquad X_i(t) \perp\!\!\!\perp X_j(t-u) \,|\, X_V \backslash \{X_i(t), X_j(t-u)\} \quad \forall t \in \mathbb{Z}$$

Undirected Graphs

Let $X_V \sim \mathcal{N}(0, \Gamma)$ with $\Gamma = \big(\Gamma(u-v)\big)_{u,v \in \mathbb{Z}}$.

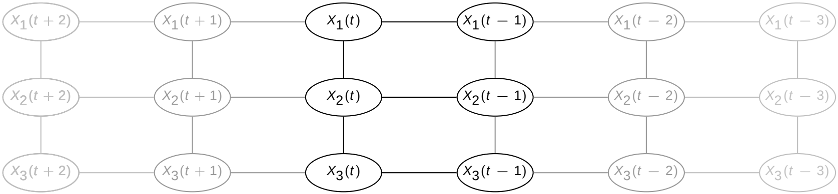
Covariance selection graphs:

let $\Gamma^{(i)} = \Gamma^{-1}$ be the inverse covariance matrix

$$\Gamma^{(i)}_{ij}(u) = 0 \qquad \Rightarrow \qquad X_i(t) \perp\!\!\!\perp X_j(t-u) \,|\, X_V \backslash \{X_i(t), X_j(t-u)\} \quad \forall t \in \mathbb{Z}$$

define concentration graph $G = (V \times \mathbb{Z}, E)$ by

$$(t,i) \text{ --- } (s,j) \notin E \quad \Leftrightarrow \quad \Gamma^{(i)}_{ij}(t-s) = 0$$



A stationary Gaussian process has a spectral representation

$$X_V(t) = \int_\Pi e^{i\lambda t}\, dZ_{X_V}(\lambda)$$

where $Z_{X_V}(\lambda)$ is a complex-valued Gaussian process on $\Pi = [-\pi, \pi)$ with

- $\mathbb{E}\big(Z_{X_V}(\lambda)\big) = 0$,
- $\text{var}\,\big(Z_{X_V}(\lambda)\big) = \int_{-\pi}^{\lambda} f(\lambda)\, d\lambda$,
- $\mathbb{E}\Big(\big[Z_{X_V}(\lambda_2) - Z_{X_V}(\lambda_1)\big]\,\big[\overline{Z_{X_V}(\mu_2) - Z_{X_V}(\mu_1)}\big]\Big) = 0$ if $(\lambda_1, \lambda_2] \cap (\mu_1, \mu_2] = \varnothing$

$f(\lambda)$ is called the spectral matrix of X_V.

A stationary Gaussian process has a spectral representation

$$X_V(t) = \int_\Pi e^{i\lambda t}\, dZ_{X_V}(\lambda)$$

where $Z_{X_V}(\lambda)$ is a complex-valued Gaussian process on $\Pi = [-\pi, \pi)$ with

- $\mathbb{E}(Z_{X_V}(\lambda)) = 0$,
- $\operatorname{var}(Z_{X_V}(\lambda)) = \int_{-\pi}^{\lambda} f(\lambda)\, d\lambda$,
- $\mathbb{E}\big([Z_{X_V}(\lambda_2) - Z_{X_V}(\lambda_1)]\,[\overline{Z_{X_V}(\mu_2) - Z_{X_V}(\mu_1)}]\big) = 0$ if $(\lambda_1, \lambda_2] \cap (\mu_1, \mu_2] = \varnothing$

$f(\lambda)$ is called the spectral matrix of X_V.

Interpretation:

- X_V can be decomposed into frequency components (sines and cosines)
- $dZ_{X_V}(\lambda)$ is the amplitude of the component for frequency λ
- $f(\lambda)$ is the covariance matrix of $dZ_{X_V}(\lambda)$

- *Covariance graphs:* we have

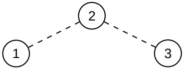
$$\Gamma(u) = \int_{\Pi} e^{i\lambda u} f(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned}\Gamma_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\ &\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \quad \forall \lambda \in \Pi \\ &\Leftrightarrow X_a \perp\!\!\!\perp X_b\end{aligned}$$

Undirected Graphs
Frequency domain

- *Covariance graphs:* we have

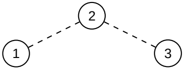


$$\Gamma(u) = \int_{\Pi} e^{\mathrm{i}\lambda u} f(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned}
\Gamma_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\
&\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \quad \forall \lambda \in \Pi \\
&\Leftrightarrow X_a \perp\!\!\!\perp X_b
\end{aligned}$$

- *Covariance graphs:* we have



$$\Gamma(u) = \int_{\Pi} e^{\mathrm{i}\lambda u} f(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned}\Gamma_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\ &\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \quad \forall \lambda \in \Pi \\ &\Leftrightarrow X_a \perp\!\!\!\perp X_b\end{aligned}$$

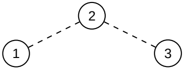
- *Covariance selection graphs:* we have

$$\Gamma^{(\mathrm{i})}(u) = \frac{1}{4\pi^2} \int_{\Pi} e^{\mathrm{i}\lambda u} f^{-1}(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned}\Gamma^{(\mathrm{i})}_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f^{-1}_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\ &\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \,|\, dZ_{X_{V\setminus\{a,b\}}}(\lambda) \quad \forall \lambda \in \Pi \\ &\Leftrightarrow X_a \perp\!\!\!\perp X_b \,|\, X_{V\setminus\{a,b\}}\end{aligned}$$

- *Covariance graphs:* we have

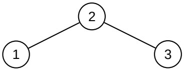


$$\Gamma(u) = \int_{\Pi} e^{\mathrm{i}\lambda u} f(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned} \Gamma_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\ &\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \quad \forall \lambda \in \Pi \\ &\Leftrightarrow X_a \perp\!\!\!\perp X_b \end{aligned}$$

- *Covariance selection graphs:* we have



$$\Gamma^{(\mathrm{i})}(u) = \frac{1}{4\pi^2} \int_{\Pi} e^{\mathrm{i}\lambda u} f^{-1}(\lambda)\, d\lambda$$

It follows that

$$\begin{aligned} \Gamma^{(\mathrm{i})}_{ab}(u) = 0 \quad \forall u \in \mathbb{Z} &\Leftrightarrow f^{-1}_{ab}(\lambda) = 0 \quad \forall \lambda \in \Pi \\ &\Leftrightarrow dZ_{X_a}(\lambda) \perp\!\!\!\perp dZ_{X_a}(\lambda) \,|\, dZ_{X_{V\setminus\{a,b\}}}(\lambda) \quad \forall \lambda \in \Pi \\ &\Leftrightarrow X_a \perp\!\!\!\perp X_b \,|\, X_{V\setminus\{a,b\}} \end{aligned}$$

Relation to common time series models:

- $\Gamma(u) = 0 \quad \forall |u| > q$ then X_V is a VMA(q) process:

$$X_V(t) = \sum_{h=0}^{q} \Psi(h)\,\varepsilon_V(t-h), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

- $\Gamma^{(\mathrm{i})}(u) = 0 \quad \forall |u| > p$ then X_V is a VAR(p) process:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h)\,X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Note: graphical constraints on the parameters in each model are nonlinear

Modelling dynamics

Consider VAR(p) process - vector autoregressive process of order p:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h) X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Let $\overline{X}_V(t) = \{X_V(s), s \leq t\}$:

Modelling dynamics

Consider VAR(p) process - vector autoregressive process of order p:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h)\, X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Let $\overline{X}_V(t) = \{X_V(s), s \leq t\}$:

- $\Phi_{ij}(h) = 0 \;\Rightarrow\; X_i(t) \perp\!\!\!\perp X_j(t-h) \,|\, \overline{X}_V(t-1) \backslash \{X_j(t-h)\}$

Modelling dynamics

Consider VAR(p) process - vector autoregressive process of order p:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h)\, X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Let $\overline{X}_V(t) = \{X_V(s), s \leq t\}$:

- $\Phi_{ij}(h) = 0 \;\Rightarrow\; X_i(t) \perp\!\!\!\perp X_j(t-h) \,|\, \overline{X}_V(t-1) \backslash \{X_j(t-h)\}$
- $\Sigma_{ij} = 0 \;\Rightarrow\; X_i(t) \perp\!\!\!\perp X_j(t) \,|\, \overline{X}_V(t-1)$

Modelling dynamics

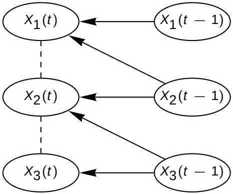
Consider VAR(p) process - vector autoregressive process of order p:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h)\, X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Let $\overline{X}_V(t) = \{X_V(s), s \leq t\}$:

- $\Phi_{ij}(h) = 0 \Rightarrow X_i(t) \perp\!\!\!\perp X_j(t-h) \,|\, \overline{X}_V(t-1) \backslash \{X_j(t-h)\}$
- $\Sigma_{ij} = 0 \Rightarrow X_i(t) \perp\!\!\!\perp X_j(t) \,|\, \overline{X}_V(t-1)$

Mixed graph for conditional distribution of $X_V(t)$ given $\overline{X}_V(t-1)$:



Modelling dynamics

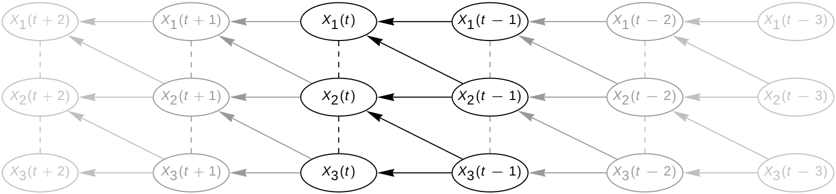
Consider VAR(p) process - vector autoregressive process of order p:

$$X_V(t) = \sum_{h=1}^{p} \Phi(h)\, X_V(t-h) + \varepsilon_V(t), \qquad \varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

Let $\overline{X}_V(t) = \{X_V(s), s \leq t\}$:

- $\Phi_{ij}(h) = 0 \;\Rightarrow\; X_i(t) \perp\!\!\!\perp X_j(t-h) \,|\, \overline{X}_V(t-1) \backslash \{X_j(t-h)\}$
- $\Sigma_{ij} = 0 \;\Rightarrow\; X_i(t) \perp\!\!\!\perp X_j(t) \,|\, \overline{X}_V(t-1)$

Mixed graph for conditional distribution of $X_V(t)$ given $\overline{X}_V(t-1)$:



The spectral representation of $X_V(t)$ yields the linear equation system

$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

where $\Phi(\lambda) = \Phi(1)\, e^{i\lambda} + \ldots + \Phi(p)\, e^{i\lambda p}$.

Such linear (structural) equation systems have been visualized by path diagrams (eg Wright 1923, Koster 1999).

The spectral representation of $X_V(t)$ yields the linear equation system

$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

where $\Phi(\lambda) = \Phi(1)\, e^{\mathrm{i}\lambda} + \ldots + \Phi(p)\, e^{\mathrm{i}\lambda p}$.

Such linear (structural) equation systems have been visualized by path diagrams (eg Wright 1923, Koster 1999).

The path diagram $G = (V, E)$ associated with the above equation system satisfies

- $i \longrightarrow j \notin E \quad \Leftrightarrow \quad \Phi_{ji}(\lambda) \equiv 0$
- $i \dashrightarrow j \notin E \quad \Leftrightarrow \quad \Sigma_{ij} = 0$

The spectral representation of $X_V(t)$ yields the linear equation system

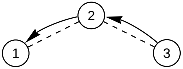
$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

where $\Phi(\lambda) = \Phi(1)\, e^{\mathrm{i}\lambda} + \ldots + \Phi(p)\, e^{\mathrm{i}\lambda p}$.

Such linear (structural) equation systems have been visualized by path diagrams (eg Wright 1923, Koster 1999).

The path diagram $G = (V, E)$ associated with the above equation system satisfies

- $i \longrightarrow j \notin E \quad \Leftrightarrow \quad \Phi_{ji}(\lambda) \equiv 0$
- $i \dashrightarrow j \notin E \quad \Leftrightarrow \quad \Sigma_{ij} = 0$



Let G be the path diagram associated with linear equation system

$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

Let G be the path diagram associated with linear equation system

$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

- directed edges:

$$j \longrightarrow i \notin G \quad \Rightarrow \quad \Phi_{ij}(\lambda) \equiv 0 \quad \Rightarrow \quad \Phi_{ij}(1) = \ldots = \Phi_{ij}(p) = 0$$
$$\Rightarrow \quad X_i(t) \perp\!\!\!\perp \overline{X}_j(t-1) \,|\, \overline{X}_{V\setminus\{j\}}(t-1)$$

$\leadsto$ X_j is *Granger-noncausal* for X_i with respect to X_V ($X_j \nrightarrow X_i \;\; [X_V]$)

Let G be the path diagram associated with linear equation system

$$dZ_{X_V}(\lambda) = \Phi(\lambda)\, dZ_{X_V}(\lambda) + dZ_{\varepsilon_V}(\lambda)$$

- directed edges:

$$j \longrightarrow i \notin G \quad \Rightarrow \quad \Phi_{ij}(\lambda) \equiv 0 \quad \Rightarrow \quad \Phi_{ij}(1) = \ldots = \Phi_{ij}(p) = 0$$
$$\Rightarrow \quad X_i(t) \perp\!\!\!\perp \overline{X}_j(t-1) \,|\, \overline{X}_{V\setminus\{j\}}(t-1)$$

$\rightsquigarrow$ X_j is *Granger-noncausal* for X_i with respect to X_V ($X_j \nrightarrow X_i \ \ [X_V]$)

- undirected edges:

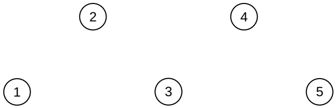
$$i \text{ - - - } j \notin G \quad \Rightarrow \quad \Sigma_{ij} = 0 \quad \Rightarrow \quad \varepsilon_i(t) \perp\!\!\!\perp \varepsilon_j(t)$$
$$\Rightarrow \quad X_i(t) \perp\!\!\!\perp X_j(t) \,|\, \overline{X}_V(t-1)$$

X_i and X_j are *contemporaneously independent* w.r.t. X_V ($X_j \nsim X_i \ \ [X_V]$)

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi X_V(t-1) + \varepsilon_V(t)$$

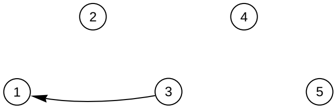
$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$



Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi\, X_V(t-1) + \varepsilon_V(t)$$

$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$



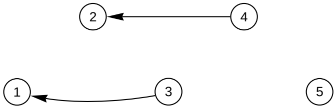
with

- $X_1(t) = \Phi_{11}\, X_1(t-1) + \Phi_{13}\, X_3(t-1) + \varepsilon_1(t)$

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi\, X_V(t-1) + \varepsilon_V(t)$$

$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

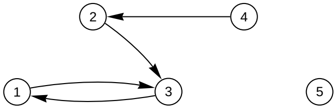


with

- $X_1(t) = \Phi_{11}\, X_1(t-1) + \Phi_{13}\, X_3(t-1) + \varepsilon_1(t)$
- $X_2(t) = \Phi_{22}\, X_2(t-1) + \Phi_{24}\, X_4(t-1) + \varepsilon_2(t)$

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi X_V(t-1) + \varepsilon_V(t)$$
$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

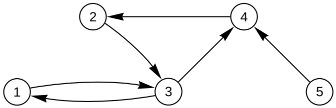


with

- $X_1(t) = \Phi_{11} X_1(t-1) + \Phi_{13} X_3(t-1) + \varepsilon_1(t)$
- $X_2(t) = \Phi_{22} X_2(t-1) + \Phi_{24} X_4(t-1) + \varepsilon_2(t)$
- $X_3(t) = \Phi_{31} X_1(t-1) + \Phi_{32} X_2(t-1) + \Phi_{33} X_3(t-1) + \varepsilon_3(t)$

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi\, X_V(t-1) + \varepsilon_V(t)$$
$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

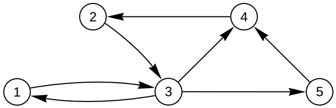


with

- $X_1(t) = \Phi_{11}\, X_1(t-1) + \Phi_{13}\, X_3(t-1) + \varepsilon_1(t)$
- $X_2(t) = \Phi_{22}\, X_2(t-1) + \Phi_{24}\, X_4(t-1) + \varepsilon_2(t)$
- $X_3(t) = \Phi_{31}\, X_1(t-1) + \Phi_{32}\, X_2(t-1) + \Phi_{33}\, X_3(t-1) + \varepsilon_3(t)$
- $X_4(t) = \Phi_{43}\, X_3(t-1) + \Phi_{44}\, X_4(t-1) + \Phi_{45}\, X_5(t-1) + \varepsilon_4(t)$

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi X_V(t-1) + \varepsilon_V(t)$$
$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$

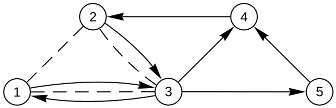


with

- $X_1(t) = \Phi_{11} X_1(t-1) + \Phi_{13} X_3(t-1) + \varepsilon_1(t)$
- $X_2(t) = \Phi_{22} X_2(t-1) + \Phi_{24} X_4(t-1) + \varepsilon_2(t)$
- $X_3(t) = \Phi_{31} X_1(t-1) + \Phi_{32} X_2(t-1) + \Phi_{33} X_3(t-1) + \varepsilon_3(t)$
- $X_4(t) = \Phi_{43} X_3(t-1) + \Phi_{44} X_4(t-1) + \Phi_{45} X_5(t-1) + \varepsilon_4(t)$
- $X_5(t) = \Phi_{53} X_3(t-1) + \Phi_{55} X_5(t-1) + \varepsilon_5(t)$

Example: consider five-dimensional VAR(1) process

$$X_V(t) = \Phi X_V(t-1) + \varepsilon_V(t)$$
$$\varepsilon_V(t) \overset{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$$



with

- $X_1(t) = \Phi_{11} X_1(t-1) + \Phi_{13} X_3(t-1) + \varepsilon_1(t)$
- $X_2(t) = \Phi_{22} X_2(t-1) + \Phi_{24} X_4(t-1) + \varepsilon_2(t)$
- $X_3(t) = \Phi_{31} X_1(t-1) + \Phi_{32} X_2(t-1) + \Phi_{33} X_3(t-1) + \varepsilon_3(t)$
- $X_4(t) = \Phi_{43} X_3(t-1) + \Phi_{44} X_4(t-1) + \Phi_{45} X_5(t-1) + \varepsilon_4(t)$
- $X_5(t) = \Phi_{53} X_3(t-1) + \Phi_{55} X_5(t-1) + \varepsilon_5(t)$
- $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & 0 & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{44} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{55} \end{pmatrix}$

Objective: derive dynamic structure for subprocess X_S, $S \subseteq V$

Objective: derive dynamic structure for subprocess X_S, $S \subseteq V$

Idea: characterize pathways that induce associations

Objective: derive dynamic structure for subprocess X_S, $S \subseteq V$

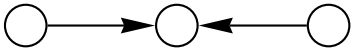
Idea: characterize pathways that induce associations

Tool: concepts of separation in graphs

- DAGs: d-separation (Pearl 1988)
- mixed graphs: d-separation (Spirtes et al. 1998, Koster 1999) or m-separation (Richardson 2003)

Definitions:

- *Collider*



Objective: derive dynamic structure for subprocess X_S, $S \subseteq V$

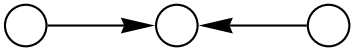
Idea: characterize pathways that induce associations

Tool: concepts of separation in graphs

- DAGs: d-separation (Pearl 1988)
- mixed graphs: d-separation (Spirtes et al. 1998, Koster 1999) or m-separation (Richardson 2003)

Definitions:

- *Collider*



- *Non-collider*

Objective: derive dynamic structure for subprocess X_S, $S \subseteq V$

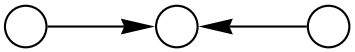
Idea: characterize pathways that induce associations

Tool: concepts of separation in graphs

- DAGs: d-separation (Pearl 1988)
- mixed graphs: d-separation (Spirtes et al. 1998, Koster 1999) or m-separation (Richardson 2003)

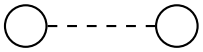
Definitions:

- *Collider*



- *Non-collider*

- *Dashed undirected edges*

 corresponds to 

A path π (sequence of adjacent edges) between *a* and *b* is said to be *connecting* given *S* if

- every non-collider on π is not in S;
- every collider on π is in S.

Otherwise the path π is blocked given S.

Example: On the path $1 \longrightarrow 2 \longleftarrow 3 \dashrightarrow\!\!\!\!\!- 4 \longleftarrow 5 \longrightarrow 6$

- vertices 2 and 4 are colliders;
- vertices 3 and 5 are non-colliders;
- the endpoints 1 and 6 are neither colliders nor non-colliders.

Consequently, the path is connecting only given $S = \{2, 4\}$; otherwise it is blocked.

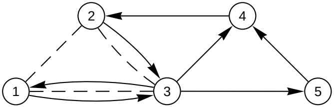
Global Granger-causal Markov property

Results: A, B, C disjoint subsets of V

- Global Markov property:

$$A \bowtie_m B \,|\, C \quad \Rightarrow \quad X_A \perp\!\!\!\perp X_B \,|\, X_C.$$

Example:



$$1 \bowtie_m 5 \,|\, \{3,4\} \quad \Rightarrow \quad X_1 \perp\!\!\!\perp X_5 \,|\, X_{\{3,4\}}$$

Global Granger-causal Markov property

Results: A, B, C disjoint subsets of V

- Global Markov property:

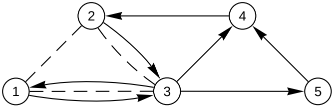
$$A \bowtie_m B \,|\, C \quad \Rightarrow \quad X_A \perp\!\!\!\perp X_B \,|\, X_C.$$

- Global Granger-causal Markov property:
  If all paths π from A into B are blocked by $C \cup B$ then

$$X_A \nrightarrow X_B \;\; [X_{A\cup B\cup C}].$$

path into B:
arrowhead at B

Example:



Global Granger-causal Markov property

Results: A, B, C disjoint subsets of V

- Global Markov property:

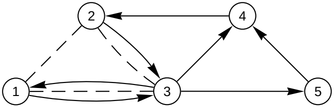
$$A \bowtie_m B \,|\, C \quad \Rightarrow \quad X_A \perp\!\!\!\perp X_B \,|\, X_C.$$

- Global Granger-causal Markov property:
  If all paths π from A into B are blocked by $C \cup B$ then

$$X_A \nrightarrow X_B \;\; [X_{A \cup B \cup C}].$$

path into B:
arrowhead at B

Example:



all paths from 1 into 4 are blocked by $\{3, 4\}$

$\Rightarrow$ X_1 is Granger-noncausal for X_4 w.r.t. $X_{\{1,3,4\}}$

Global Granger-causal Markov property

Results: A, B, C disjoint subsets of V

- Global Markov property:

$$A \bowtie_m B \mid C \quad \Rightarrow \quad X_A \perp\!\!\!\perp X_B \mid X_C.$$

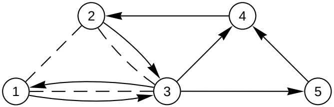
- Global Granger-causal Markov property:
  If all paths π from A into B are blocked by $C \cup B$ then

$$X_A \nrightarrow X_B \quad [X_{A \cup B \cup C}].$$

- Similar result for contemporaneous independence.

path into B:
arrowhead at B

Example:



all paths from 1 into 4 are blocked by $\{3, 4\}$

$\Rightarrow$ X_1 is Granger-noncausal for X_4 w.r.t. $X_{\{1,3,4\}}$

Systems with latent variables

Now suppose that X_V is potentially affected by latent variables.

Objective: find graphical representation of Granger-causal relationships

Problem: the class of Granger-causality graphs just introduced is not sufficient

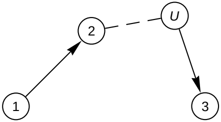
Systems with latent variables

Now suppose that X_V is potentially affected by latent variables.

Objective: find graphical representation of Granger-causal relationships

Problem: the class of Granger-causality graphs just introduced is not sufficient

Example:



all variables

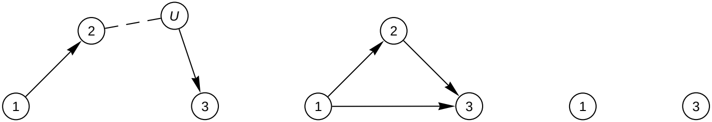
Systems with latent variables

Now suppose that X_V is potentially affected by latent variables.

Objective: find graphical representation of Granger-causal relationships

Problem: the class of Granger-causality graphs just introduced is not sufficient

Example:



all variables — three variables — two variables

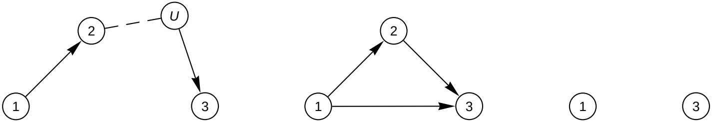
Systems with latent variables

Now suppose that X_V is potentially affected by latent variables.

Objective: find graphical representation of Granger-causal relationships

Problem: the class of Granger-causality graphs just introduced is not sufficient

Example:



all variables three variables two variables

spurious causality of type I

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$$a \longleftarrow u \longrightarrow b \quad \text{leads to} \quad a \dashrightarrow b$$

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$$a \longleftarrow u \longrightarrow b \quad \text{leads to} \quad a \dashrightarrow b \quad \text{and/or} \quad a \dashleftarrow b$$
$$\text{and/or} \quad a \dashline b$$

The combination of $a \dashrightarrow b$, $a \dashleftarrow b$, and $a \dashline b$ is abbreviated by $a \dashleftarrow\!\!\dashrightarrow b$.

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$$a \longleftarrow u \longrightarrow b \quad \text{leads to} \quad a \dashrightarrow b \quad \text{and/or} \quad a \dashleftarrow b$$

$$\text{and/or} \quad a \dashline b$$

The combination of $a \dashrightarrow b$, $a \dashleftarrow b$, and $a \dashline b$ is abbreviated by $a \dashleftrightarrow b$.

Extended definition of m-separation:

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$a \longleftarrow u \longrightarrow b$ leads to $a \dashrightarrow b$ and/or $a \dashleftarrow b$

and/or $a \dashline b$

The combination of $a \dashrightarrow b$, $a \dashleftarrow b$, and $a \dashline b$ is abbreviated by $a \dashleftrightarrow b$.

Extended definition of m-separation:

- non-collider:
  $\longrightarrow \circ \longrightarrow$ or $\longleftarrow \circ \longrightarrow$

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$$a \longleftarrow u \longrightarrow b \quad \text{leads to} \quad a \dashrightarrow b \quad \text{and/or} \quad a \dashleftarrow b$$

$$\text{and/or} \quad a \mathrel{-\,-\,-} b$$

The combination of $a \dashrightarrow b$, $a \dashleftarrow b$, and $a \mathrel{-\,-\,-} b$ is abbreviated by $a \dashleftarrow\!\!\dashrightarrow b$.

Extended definition of m-separation:

- non-collider:
  $\longrightarrow \circ \longrightarrow$ or $\longleftarrow \circ \longrightarrow$
- collider:
  $\longrightarrow \circ \longleftarrow$
- dashed edges:
  $\mathrel{-\,-\,-}$ and $\dashrightarrow$ correspond to $\longleftrightarrow$

Systems with latent variables

Graphical representation

Idea: Use new type of edges $a \dashrightarrow b$ to represent spurious causality:

$$a \longleftarrow u \longrightarrow b \quad \text{leads to} \quad a \dashrightarrow b \quad \text{and/or} \quad a \dashleftarrow b$$

$$\text{and/or} \quad a \;\text{---}\; b$$

The combination of $a \dashrightarrow b$, $a \dashleftarrow b$, and $a \;\text{---}\; b$ is abbreviated by $a \leftarrow\!\!\!\dashrightarrow b$.

Extended definition of m-separation:

- non-collider:
  $\longrightarrow \circ \longrightarrow$ or $\longleftarrow \circ \longrightarrow$
- collider:
  $\longrightarrow \circ \longleftarrow$
- dashed edges:
  --- and $\dashrightarrow$ correspond to $\longleftrightarrow$

Example:

- Collider $\dashrightarrow c \longleftarrow$, $\dashrightarrow c \;\text{---}$, $\text{---}\; c \;\text{---}$, $\dashleftarrow c \dashrightarrow$
- Non-collider $\longleftarrow n \longrightarrow$, $\longrightarrow n \longrightarrow$, $\dashrightarrow n \longrightarrow$

Graphical representation for X_S with $S \subseteq V$

For $S = V\backslash\{c\}$ we obtain the new Graph $G^{(c)}$ with vertex set $V\backslash\{c\}$ by

- removing vertex c;
- removing all edges adjacent to c;
- adding edges $e = a \cdots b$ according to the following table

	$c \cdots b$				
$a \cdots c$	$\longrightarrow$	$\longleftarrow$	$\dashrightarrow$	$\dashleftarrow$	$---$
$\longrightarrow$	$\longrightarrow$				
$\longleftarrow$	$\dashleftarrow\!\!\dashrightarrow$	$\longleftarrow$	$\dashleftarrow\!\!\dashrightarrow$	$\dashleftarrow$	$\longleftarrow$
$\dashrightarrow$	$\dashrightarrow$				
$\dashleftarrow$	$\dashleftarrow\!\!\dashrightarrow$				
$---$	$\dashrightarrow$				

For $C = \{c_1, \ldots, c_n\}$ we define recursively $G^{(C)} = (G^{(\{c_1,\ldots,c_{n-1}\})})^{(c_n)}$.

Result: If X_V is Granger-causal Markov w.r.t. G then X_S is Granger-causal Markov w.r.t. $G^{(V\backslash S)}$.

Result: Suppose

- X_V satisfies the Granger-causal Markov property w.r.t. G
- A, B, D are disjoint subsets of S

Result: Suppose

- X_V satisfies the Granger-causal Markov property w.r.t. G
- A, B, D are disjoint subsets of S

Then the following two statements are equivalent:

- in G all paths from A into B are blocked given D;
- in $G^{(V \setminus S)}$ all paths from A into B are blocked given D.

Result: Suppose

- X_V satisfies the Granger-causal Markov property w.r.t. G
- A, B, D are disjoint subsets of S

Then the following two statements are equivalent:

- in G all paths from A into B are blocked given D;
- in $G^{(V \setminus S)}$ all paths from A into B are blocked given D.

This means that the original graph G and the marginal graph $G^{(V \setminus S)}$ encode the same Granger-causal relationships for X_S
$\rightsquigarrow$ no information about structure of X_S is lost.

Similar result holds for contemporaneous independence.

Markov equivalence

Let X_V be a multivariate time series possibly affected by latent variables.

Objective: Find graphical representation encoding the Granger-causality relationships satisfied by X_V.

Markov equivalence

Let X_V be a multivariate time series possibly affected by latent variables.

Objective: Find graphical representation encoding the Granger-causality relationships satisfied by X_V.

Note: We do not know the full system - including latent variables.

Markov equivalence

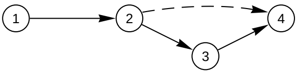
Let X_V be a multivariate time series possibly affected by latent variables.

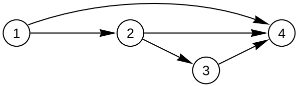
Objective: Find graphical representation encoding the Granger-causality relationships satisfied by X_V.

Note: We do not know the full system - including latent variables.

Problem: Graphical representation is not unique.

Example:





Markov equivalence

Let X_V be a multivariate time series possibly affected by latent variables.

Objective: Find graphical representation encoding the Granger-causality relationships satisfied by X_V.

Note: We do not know the full system - including latent variables.

Problem: Graphical representation is not unique.

Example:

Graphs that encode the same set of relationships are said to be *Markov equivalent*.

Dynamic maximal ancestral graphs

Suitable subclass: *dynamic Maximal Ancestral graphs (dMAGs)*

- if $a \longrightarrow \cdots \longrightarrow b$ in G then $a \dashrightarrow b$ not in G
- if $a \longrightarrow b$ and $a \dashrightarrow b$ not in G then there exists $S \subseteq V$ such that

$$X_a \not\rightarrow X_b \ \ [X_S]$$

or there exist disjoint subsets S_1, S_2 with $b \in S_1$ and $a \notin S_1 \cup S_2$ such that

$$X_a(t-k) \perp\!\!\!\perp X_b(t+1) \,|\, \overline{X}_{S_1}(t), \overline{X}_{S_2}(t-k), \overline{X}_a(t-k-1).$$

for all $k \in \mathbb{N}$.

Dynamic maximal ancestral graphs

Suitable subclass: *dynamic Maximal Ancestral graphs (dMAGs)*

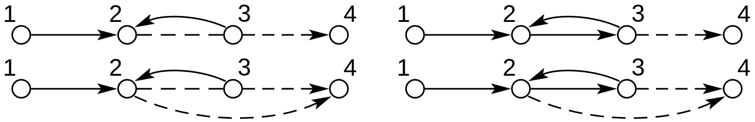
- if $a \longrightarrow \cdots \longrightarrow b$ in G then $a \dashrightarrow b$ not in G
- if $a \longrightarrow b$ and $a \dashrightarrow b$ not in G then there exists $S \subseteq V$ such that

$$X_a \not\rightarrow X_b \quad [X_S]$$

or there exist disjoint subsets S_1, S_2 with $b \in S_1$ and $a \notin S_1 \cup S_2$ such that

$$X_a(t-k) \perp\!\!\!\perp X_b(t+1) \,|\, \overline{X}_{S_1}(t), \overline{X}_{S_2}(t-k), \overline{X}_a(t-k-1).$$

for all $k \in \mathbb{N}$.



Dynamic maximal ancestral graphs

Suitable subclass: *dynamic Maximal Ancestral graphs (dMAGs)*

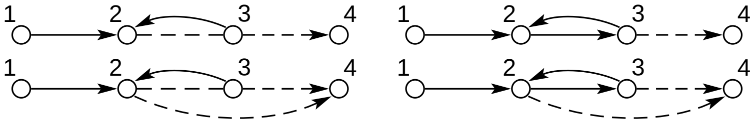
- if $a \longrightarrow \cdots \longrightarrow b$ in G then $a \dashrightarrow b$ not in G
- if $a \longrightarrow b$ and $a \dashrightarrow b$ not in G then there exists $S \subseteq V$ such that

$$X_a \nrightarrow X_b \quad [X_S]$$

or there exist disjoint subsets S_1, S_2 with $b \in S_1$ and $a \notin S_1 \cup S_2$ such that

$$X_a(t-k) \perp\!\!\!\perp X_b(t+1) \,|\, \overline{X}_{S_1}(t), \overline{X}_{S_2}(t-k), \overline{X}_a(t-k-1).$$

for all $k \in \mathbb{N}$.



Properties:

- smaller Markov equivalence classes (still not unique)
- advantageous for inference

Identification of causal structure

Objective:

Given a set of Granger-causal relationships for X_V identify the corresponding Markov equivalence class of dMAGs.

Algorithm:

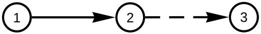
- *Adjacencies:*
  - insert $a \longrightarrow b$ whenever X_a and X_b are not contemporaneously independent
  - insert $a \dashrightarrow b$ whenever X_b Granger-causes X_a w.r.t. X_S for all $S \subseteq V$ with $a, b \in S$ and

$$X_a(t-k) \not\perp\!\!\!\perp X_b(t+1) \,|\, \overline{X}_{S_1}(t), \overline{X}_{S_2}(t-k), \overline{X}_a(t-k-1)$$

    for some $k \in \mathbb{N}$ and all S_1, S_2 with $b \in S_1$ and $a \notin S_1 \cup S_2$.
- *Identification of tails:*
  - colliders
  - non-colliders
  - ancestors
  - discriminating paths

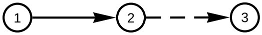
Identification of causal structure

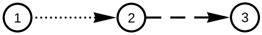
Examples:



Identification of causal structure

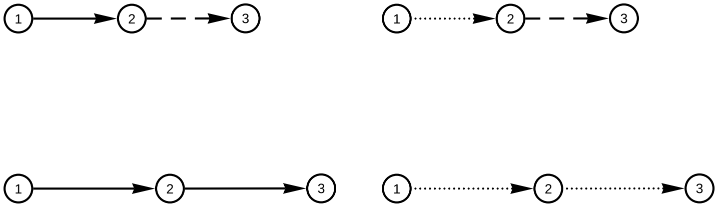
Examples:





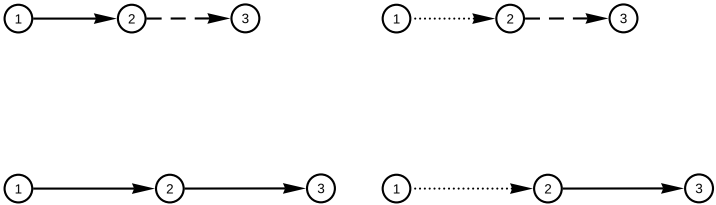
Identification of causal structure

Examples:



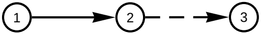
Identification of causal structure

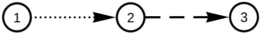
Examples:

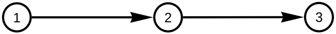


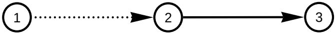
Identification of causal structure

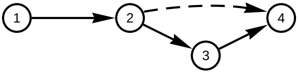
Examples:

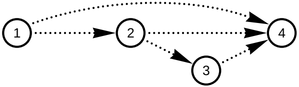






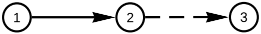


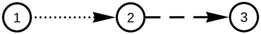


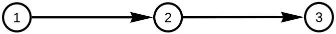


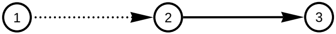
Identification of causal structure

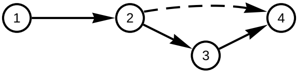
Examples:

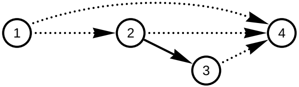












Interpretation

Question: What can we learn from the graphical analysis about the true causal structure?

Interpretation

Question: What can we learn from the graphical analysis about the true causal structure?

- *Causal Markov assumption:*
  all observed dependencies are due to causal influences; key assumption underlying all approaches to causal inference from observational data

Interpretation

Question: What can we learn from the graphical analysis about the true causal structure?

- *Causal Markov assumption:*

  all observed dependencies are due to causal influences; key assumption underlying all approaches to causal inference from observational data

- *Faithfulness assumption:*

  the independencies observed are structural and not due to several influences exactly cancelling out ($\rightsquigarrow$ spurious causality of type I)

Interpretation

Question: What can we learn from the graphical analysis about the true causal structure?

- *Causal Markov assumption:*

  all observed dependencies are due to causal influences; key assumption underlying all approaches to causal inference from observational data
- *Faithfulness assumption:*

  the independencies observed are structural and not due to several influences exactly cancelling out ($\rightsquigarrow$ spurious causality of type I)
- *Invariance of edges:*

  only edges that are invariant in the Markov equivalence class are identified as
  - causes if the edge is $\longrightarrow$;
  - spurious causes if the edge is $\dashrightarrow$.

Interpretation

Question: What can we learn from the graphical analysis about the true causal structure?

- *Causal Markov assumption:*

  all observed dependencies are due to causal influences; key assumption underlying all approaches to causal inference from observational data
- *Faithfulness assumption:*

  the independencies observed are structural and not due to several influences exactly cancelling out ($\rightsquigarrow$ spurious causality of type I)
- *Invariance of edges:*

  only edges that are invariant in the Markov equivalence class are identified as
  - causes if the edge is $\longrightarrow$;
  - spurious causes if the edge is $\dashrightarrow$.
- *Ancestrality of graphs:*

  Directed edges cannot be interpreted as direct causes, but signify a causal link.

Modelling latent variable structures

Extension of the VAR(p) model:

- $X_V(t)$ stationary process given by

$$X_V(t) = \sum_{k=1}^{p} \Phi(k)\, X_V(t-k) + \varepsilon_V(t)$$

- $\varepsilon_V(t)$ stationary Gaussian process with $\mathbb{E}\big(\varepsilon_V(t)\big) = 0$ and

$$\operatorname{cov}\big(\varepsilon_V(t+k), \varepsilon_V(t)\big) = \begin{cases} \Omega(k) & \text{if } |k| \leq q \\ 0 & \text{otherwise} \end{cases}$$

- *Model parameters:*

$$\phi = \mathrm{vec}(\Phi(1), \ldots, \Phi(k)) \qquad \text{and} \qquad \omega = \big(\Omega_{ij}(u), j \leq i, |u| \leq q\big)$$

Modelling latent variable structures

Extension of the VAR(p) model:

- $X_V(t)$ stationary process given by

$$X_V(t) = \sum_{k=1}^{p} \Phi(k)\, X_V(t-k) + \varepsilon_V(t)$$

- $\varepsilon_V(t)$ stationary Gaussian process with $\mathbb{E}\big(\varepsilon_V(t)\big) = 0$ and

$$\operatorname{cov}\big(\varepsilon_V(t+k), \varepsilon_V(t)\big) = \begin{cases} \Omega(k) & \text{if } |k| \leq q \\ 0 & \text{otherwise} \end{cases}$$

- *Model parameters:*

$$\phi = \operatorname{vec}(\Phi(1), \ldots, \Phi(k)) \qquad \text{and} \qquad \omega = \big(\Omega_{ij}(u), j \leq i, |u| \leq q\big)$$

Relation to VARMA models:

This is equivalent to a VARMA(p,q) model with parameter constraints encoded by G.

Modelling latent variable structures

Graphical restrictions:

$G = (V, E)$ mixed graph (edge types $\longrightarrow$, $---$, $--\rightarrow$)

- $j \longrightarrow i \notin E \;\Rightarrow\; \Phi_{ij}(k) = 0,\, k = 1, \ldots, p$
- $i --- j \notin E \;\Rightarrow\; \Omega_{ij}(0) = 0$
- $i --\rightarrow j \notin E \;\Rightarrow\; \Omega_{ij}(k) = 0,\, k = 1, \ldots, p$
- $i \leftarrow-- j \notin E \;\Rightarrow\; \Omega_{ij}(k) = 0,\, k = -1, \ldots, -p$

Then X_V satisfies the global Granger-causal Markov property with respect to G.

Modelling latent variable structures

Graphical restrictions:

$G = (V, E)$ mixed graph (edge types $\longrightarrow$, $\relbar\relbar\relbar$, $\dashrightarrow$)

- $j \longrightarrow i \notin E \;\Rightarrow\; \Phi_{ij}(k) = 0,\ k = 1, \ldots, p$
- $i \relbar\relbar\relbar j \notin E \;\Rightarrow\; \Omega_{ij}(0) = 0$
- $i \dashrightarrow j \notin E \;\Rightarrow\; \Omega_{ij}(k) = 0,\ k = 1, \ldots, p$
- $i \dashleftarrow j \notin E \;\Rightarrow\; \Omega_{ij}(k) = 0,\ k = -1, \ldots, -p$

Then X_V satisfies the global Granger-causal Markov property with respect to G.

Identifiability:

The parameters are identifiable if the graph G is ancestral, that is,

$$i \dashrightarrow j \notin E \text{ whenever } i \longrightarrow \ldots \longrightarrow j \text{ in } G$$

Likelihood inference

- Exact likelihood function: problems due to autoregressive part

Inference

Likelihood inference

- Exact likelihood function: problems due to autoregressive part
- Conditional likelihood function: problems due to dependence of $\varepsilon(t)$

Inference

Likelihood inference

- Exact likelihood function: problems due to autoregressive part
- Conditional likelihood function: problems due to dependence of $\varepsilon(t)$
- Whittle likelihood: Ignores edge effects $\rightsquigarrow$ no problems

Likelihood inference

- Exact likelihood function: problems due to autoregressive part
- Conditional likelihood function: problems due to dependence of $\varepsilon(t)$
- Whittle likelihood: Ignores edge effects $\rightsquigarrow$ no problems

Whittle likelihood:

$$\mathscr{L}^{(T)}(\phi,\omega) = \frac{1}{4\pi}\int_{\Pi}\Big\{\log\det g_\omega(\lambda) + \operatorname{tr}\big[\Phi(e^{-i\lambda})I_{XX}^{(T)}(\lambda)\Phi(e^{i\lambda})' g_\omega(\lambda)^{-1}\big]\Big\}d\lambda$$

where

$$\Phi(z) = \mathbb{1} - \Phi(1)\,z - \ldots - \Phi(p)\,z^p$$

$$g_\omega(\lambda) = \frac{1}{2\pi}\sum_{|u|\le p}\Omega(u)e^{-i\lambda u}$$

$$I_{XX}^{(T)}(\lambda) = \big(2\pi T\big)^{-1} d_X^{(T)}(\lambda) d_X^{(T)}(-\lambda)'$$

$$d_X^{(T)}(\lambda) = \sum_{t=1}^{T} X(t)\,e^{-i\lambda t}$$

Inference

Likelihood equations:

- $i, j \in V, j \longrightarrow i \in E, h = 1, \ldots, p$

$$\int_{\Pi} \left[g_\omega(\lambda)^{-1} \Phi(e^{-\mathrm{i}\lambda}) I_{XX}^{(T)}(\lambda) \right]_{ij} e^{\mathrm{i}\lambda h} \, d\lambda = 0$$

- $i, j \in V, i \dashrightarrow\!\!\!\!\!- j \in E$

$$\int_{\Pi} \left[g_\omega(\lambda)^{-1} - g_\omega(\lambda)^{-1} \Phi(e^{-\mathrm{i}\lambda}) I_{XX}^{(T)}(\lambda) \Phi(e^{\mathrm{i}\lambda})' g_\omega(\lambda)^{-1} \right]_{ij} d\lambda = 0$$

- $i, j \in V, j \dashrightarrow i \in E, h = 1, \ldots, p$

$$\int_{\Pi} \left[g_\omega(\lambda)^{-1} - g_\omega(\lambda)^{-1} \Phi(e^{-\mathrm{i}\lambda}) I_{XX}^{(T)}(\lambda) \Phi(e^{\mathrm{i}\lambda})' g_\omega(\lambda)^{-1} \right]_{ij} e^{\mathrm{i}\lambda h} \, d\lambda = 0$$

Solution to the first equations (for ω fixed):

$$\pi_\phi \operatorname{vec} \int_\Pi e' \otimes g_\omega(\lambda)^{-1} I_{XX}^{(T)}(\lambda) d\lambda = \pi_\phi \int_\Pi E \otimes I_{XX}^{(T)}(\lambda)' \otimes g_\omega(\lambda)^{-1} d\lambda \, \pi_\phi \cdot \phi$$

where

- π_ϕ is the projection onto the restricted parameter space for ϕ
- $e = (e^{\mathrm{i}\lambda}, \ldots, e^{\mathrm{i}\lambda p})'$
- $E = e \cdot \bar{e}'$

Inference

Solution to the first equations (for ω fixed):

$$\pi_\phi \,\text{vec} \int_\Pi e' \otimes g_\omega(\lambda)^{-1} I_{XX}^{(T)}(\lambda) d\lambda = \pi_\phi \int_\Pi E \otimes I_{XX}^{(T)}(\lambda)' \otimes g_\omega(\lambda)^{-1} d\lambda \, \pi_\phi \cdot \phi$$

where

- π_ϕ is the projection onto the restricted parameter space for ϕ
- $e = (e^{\mathrm{i}\lambda}, \ldots, e^{\mathrm{i}\lambda p})'$
- $E = e \cdot \bar{e}'$

Solution to the remaining equations (for ϕ fixed):
Anderson's method (T.W. Anderson, 1973): Iterate

$$\omega^{(r+1)} = \left[\pi_\omega \cdot \frac{1}{2\pi} \int_\Pi \left(E \otimes g_{\omega^{(r)}}(\lambda)^{-1} \otimes g_{\omega^{(r)}}(\lambda)^{-1} \right) d\lambda \cdot \pi_\omega \right]^{-1}$$
$$\cdot \, \pi_\omega \,\text{vec} \left[\int_\Pi e' \otimes g_{\omega^{(r)}}(\lambda)^{-1} \Phi(e^{-\mathrm{i}\lambda}) I_{XX}^{(T)}(\lambda) \Phi(e^{\mathrm{i}\lambda})' g_{\omega^{(r)}}(\lambda)^{-1} d\lambda \right].$$

Summary

Granger causality graphs

- encode (global) Granger noncausality relations
- convenient framework for discussing spurious causality in higher-dimensional time series

Latent variable structures

- cannot be encoded completely by simple Granger causality graphs
- can be represented by an extended version of these graphs (with edges $\dashrightarrow$)
- identification only possible up to Markov equivalence

Extension of VAR(p) models

- model latent variable structures by lagged covariances of error process
- correspond to VARMA models with nonlinear constraints on parameters
- can be fitted iteratively based on Whittle likelihood

References

- Eichler, M. (2005), A graphical approach for evaluating effective connectivity in neural systems, *Philosophical Transactions of The Royal Society B* 360, 953-967.
- Eichler, M. (2006), Graphical modelling of dynamic relationships in multivariate time series. In: M. Winterhalder, B. Schelter, J. Timmer (eds), *Handbook of Time Series Analysis*, Wiley-VCH, Berlin, pp. 335-372.
- Eichler, M. (2007), Granger-causality and path diagrams for multivariate time series, *Journal of Econometrics* 137, 334-353.
- Eichler, M. (2007), Causal inference from time series: what can be learned from Granger causality? To appear in: G. Glymour, W. Wang, D. Westerståhl (eds), Proceedings of the 13th International Congress of Logic, Methodology and Philosophy of Science.
- Eichler, M., and Didelez, V. (2007), Causal reasoning in graphical time series models, *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*.