Dynamic Analysis of Multivariate Time Series Using Conditional Wavelet Graphs

Maria Grith[1]
Matthias Eckardt[2]

[1]Ladislaus von Bortkiewicz Chair of Statistics
[2]Department of Computer Science
Humboldt–Universität zu Berlin

Contributions

- Extend Granger causality and partial correlation graphs for time series to the time-frequency domain using wavelets
- Describe local stationarity in terms of local graphs
- Graph recovery from empirical data (graph structure learning, graph estimation)

Related Literature

Partial correlation graphs for multivariate time series

- generalize classical Gaussian concentration graphical models
- indicate the pairwise conditional linear dependence
- account for the contemporaneous and lagged influences

Granger causal graphs for multivariate time series

- an effect cannot precede its cause in time, (Granger, 1969)
- alternative to intervention-based causality (Pearl, 1995)
- account for lagged influences

Brillinger (1981), Brillinger (1996), Dahlhaus (2000), Eichler (2000), Dahlhaus and Eichler (2003), Eichler (2007), Eckardt (2015) - review study; Barigozzi and Brownless (2014)

Outline

1. Graphical models for time series
2. Granger Causality Graph
3. Partial Correlation Graph
4. Frequency domain representation
5. Wavelet graphs
6. Graph estimation
7. Final remarks

Graphical Models

A **graph** $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of:

- a set of vertices $\mathcal{V} = \{v_1, \ldots, v_k\} < \infty$
- a set of edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, $e_{ij} = (v_i, v_j)$
  - undirected edges $e_{ij} \in \mathcal{E} \Leftrightarrow e_{ji} \in \mathcal{E}$, **undirected graph**
  - directed edges $e_{i \to j} \in \mathcal{E}$, **directed graph**
- optional: *loops, multiple edges (multigraph), mixed graph (directed and undirected edges)*

Usually, $v_i \in \mathcal{V}$ represents a random variable or process.

Graphical Models for Time Series

k-dimensional stationary **multivariate time series** $X_V(t)$

- $X_V(t) = \{X_i(t)\}_{i \in V}$, $t \in \mathbb{Z}$, $V = \{1, \ldots, k\}$
- $X_{V \setminus S}(t) = \{X_i(t)\}_{i \in V \setminus S}$, for any $S \subseteq V$

The **time series graph** of a process X_V

- vertex v_i refers to the X_i component processes of X_V

Linear dependence graphs

- *Conditional orthogonality*: X_i and X_j are conditionally uncorrelated after removing the linear effects of X_S
  $X_i \perp\!\!\!\perp X_j \mid X_{V \setminus S}$

Remark: For Gaussian time series "$\perp\!\!\!\perp$" $\approx$ independence; factorization of the joint distribution in marginals of subgraphs

Granger Causality Graph

- X_i is linearly **nocausal** for X_j relative to the process X_V, denoted by $X_i \nrightarrow X_j \mid X_V$ if

$$X_j(t) \perp\!\!\!\perp \tilde{X}_i(t) \mid \tilde{X}_{V\setminus\{i\}}(t),$$

  for $\tilde{X}_S = \{X_S(z), z < t\}$.
- X_i and X_j are **contemporaneously uncorrelated** relative to the process X_V, denoted by $X_i \nsim X_j \mid X_V$ if

$$X_i(t) \perp\!\!\!\perp X_j(t) \mid \tilde{X}_V(t), X_{V\setminus\{i,j\}}(t).$$

Definition: The Granger causality graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for a stationary process X_V is a mixed graph given by
(i) $e_{i\rightarrow j} \notin \mathcal{E}^{GC} \Leftrightarrow X_i \nrightarrow X_j \mid X_V$,
(ii) $e_{ij} \notin \mathcal{E}^{GC} \Leftrightarrow X_i \nsim X_j \mid X_V$.

Partial Correlation Graph for Time Series

Definition: The partial correlation graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for a stationary process X_V is given by

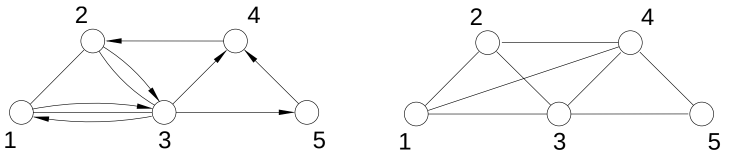
$$e_{ij} \notin \mathcal{E} \Leftrightarrow X_i \perp\!\!\!\perp X_j \mid X_{V\setminus\{i,j\}}$$
$$\Leftrightarrow cov(\varepsilon_{i|V\setminus\{i,j\}}(t), \varepsilon_{j|V\setminus\{i,j\}}(t+u)), \forall u \in \mathbb{Z}$$

$$\varepsilon_{i|V\setminus\{i,j\}} := X_i(t) - \mu_i^{opt} - \sum_{u=-\infty}^{+\infty} d_i^{opt}(u) X_{V\setminus\{i,j\}}(t-u)$$

$$(\mu_i^{opt}, d_i^{opt}) = \arg\min_{\mu_i, d_i} \mathrm{E}(X_i(t) - \mu_i - \sum_{u=-\infty}^{+\infty} d_i(u) X_{V\setminus\{i,j\}}(t-u))^2$$

Example: Five-dimensional VAR(2)-process with parameters

$$A(1)=\begin{pmatrix} \frac{3}{5} & 0 & \frac{1}{5} & 0 & 0 \\ 0 & \frac{3}{5} & 0 & -\frac{1}{5} & 0 \\ \frac{2}{5} & \frac{1}{3} & \frac{3}{5} & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{5} \\ 0 & 0 & \frac{1}{5} & 0 & \frac{2}{5} \end{pmatrix},\ A(2)=\begin{pmatrix} 0 & 0 & -\frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & -\frac{1}{5} \end{pmatrix},\ \Sigma_{\varepsilon}=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & 0 & 0 \\ \frac{1}{2} & 1 & -\frac{1}{3} & 0 & 0 \\ \frac{1}{3} & -\frac{1}{3} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$



Granger causality graph (left) and partial correlation (right) - moralization

Frequency Domain Formulation

Partial cross-spectrum b/w X_i and X_j at frequency $\omega \in [-\pi, \pi]$

$$f_{ij|V\setminus\{i,j\}}(\omega) = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} \left[\sum_{u=-\infty}^{+\infty} \varepsilon_{i|V\setminus\{i,j\}}(t)\varepsilon_{j|V\setminus\{i,j\}}(t+u) \right] e^{-i\omega t}$$

$$= \frac{1}{2\pi} \sum_{u=-\infty}^{+\infty} cov(\varepsilon_{i|V\setminus\{i,j\}}(t), \varepsilon_{j|V\setminus\{i,j\}}(t+u))e^{-i\omega t}$$

- is the Fourier transform of the cross-correlation function
- is a measure of covariance b/w $\varepsilon_{i|V\setminus\{i,j\}}$ and $\varepsilon_{j|V\setminus\{i,j\}}$

$\rightarrow X_i \perp\!\!\!\perp X_j \mid X_{V\setminus\{i,j\}} \Leftrightarrow f_{ij|V\setminus\{i,j\}}(\omega) = 0, \forall\omega$

Partial Spectral Coherence

Observation: The estimation of residuals $\varepsilon_{i|V\setminus\{i,j\}}(t)$ is computationally intensive.

Alternative: If the spectral matrix $f_V(\omega) = \{f_{ij}(\omega)\}_{i,j\in V}$ is regular and $g(\omega) := f(\omega)^{-1}$ then the **partial spectral coherence matrix** is $R(\omega) = -diag(g(\omega))^{-1/2} g(\omega) diag(g(\omega))^{-1/2}$, whose elements can be shown to satisfy

$$R_{ij|V\setminus\{i,j\}}(\omega) = \frac{f_{ij|V\setminus\{i,j\}}(\omega)}{\left[f_{ii|V\setminus\{i,j\}}(\omega) f_{jj|V\setminus\{i,j\}}(\omega)\right]^{\frac{1}{2}}}.$$

$\rightarrow X_i \perp\!\!\!\perp X_j \mid X_{V\setminus\{i,j\}} \Leftrightarrow R_{ij|V\setminus\{i,j\}}(\omega) = 0, \forall\omega \Leftrightarrow g_{ij}(\omega) = 0, \forall\omega$

Vector Autoregressive Processes

$$X(t) = \sum_{j=1}^{p} A_j X(t-j) + \varepsilon(t),\ \ \varepsilon(t) \sim \mathrm{N}(0, \Sigma_\varepsilon)$$

A_j are $k \times k$ matrices. Let $A(z) := I - \sum_{j=1}^{p} A_j z^p$. The *spectral density matrix* of representation $X(t)$ is

$$f(\omega) = \frac{1}{2\pi} A^{-1}(e^{-i\omega}) \Sigma_\varepsilon A^{-1}(e^{i\omega})^\top$$

and

$$g(\omega) = f(\omega)^{-1} = 2\pi A(e^{i\omega})^\top \Gamma_\varepsilon A(e^{-i\omega}),\ \ \Gamma_\varepsilon = \Sigma_\varepsilon^{-1}.$$

Then

$$g_{ij} \Leftrightarrow \sum_{h=0\vee u}^{p \vee p+u} \sum_{j,l=1}^{k} \Gamma_{\varepsilon,jl} A_{ji}(h) A_{lj}(h+u) = 0,\ \ (u = -p, \cdots, p).$$

Localized Partial Correlation Graph

For locally stationary multivariate time series, **wavelet**-based methods

- allow time varying analysis of spectral behavior
- characterize dependence in time-frequency domain
- similar to applying linear filters locally
- local covariance functions, local cross-spectra and local coherence

Remark: If the time series are stationary, their spectral behavior will be constant over time.

Wavelets

- "Mother wavelet" $\psi \in L_2(\mathbb{R})$ s.t.

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \text{ admissibility condition}$$

$$\int_{-\infty}^{\infty} \psi^2(t)dt = \|\psi\|^2 = 1 \text{ 'unit' energy property.}$$

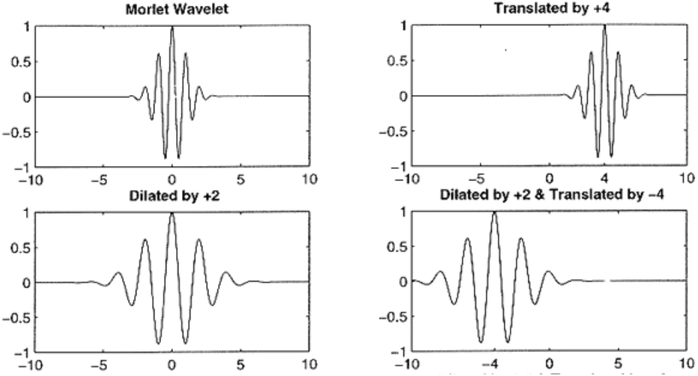
- Families of basis functions $\psi_{\tau,s}(t)$

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right),\ s \in \mathbb{R}^+, \tau \in \mathbb{R} \tag{1}$$

τ location and s scale (pseudo-frequency); $\|\psi_{\tau,s}\| = 1$

Note: We will consider complex wavelets further on.

Example: Morlet Wavelet



Morlet wavelet under translation and dilation

Wavelet Transform

Wavelet coefficients w.r.t. X_i

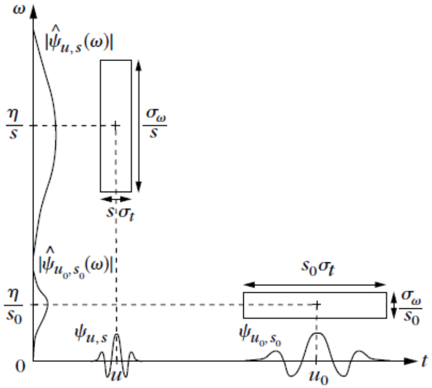
$$W_i(\tau, s) = \langle X_i, \psi_{\tau,s} \rangle = \frac{1}{\sqrt{s}} \sum_{-\infty}^{+\infty} X_i(t) \overline{\psi_{\tau,s}(t)}$$

$\overline{(\cdot)}$ stands for the complex conjugate. Additionally, a frequency domain representation of $W_i(\tau, s)$ follows as

$$W_i(\omega) = \frac{\sqrt{|s|}}{2\pi} \sum_{t=-\infty}^{\infty} X_i(t) \overline{f_{\psi_{s,\tau}}(st)} e^{i\omega t},$$

where $f_{\psi_{s,\tau}}$ is the Fourier transform of the wavelet function $\psi_{\tau,s}(t)$.

'Adaptive' Window



Time-frequency boxes of two wavelet basis

Parseval's Relation: Extension to Wavelets

Recall: The inner product of two time series equals the inner product of their Fourier transform.

- $X_i(t)$ can be recovered from the wavelet transform

$$X_i(t) = \frac{1}{C_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{s^2}W_i(\tau,s)\psi_{\tau,s}(t)d\tau ds$$

- For two processes $X_i(t)$ and $X_j(t)$, the energy in the time domain is preserved in the time-frequency domain

$$\langle X_i X_j\rangle = \frac{1}{C_\psi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\frac{1}{s^2}|W_i(\tau,s)\overline{W_i(\tau,s)}|d\tau ds,$$

  for a finite constant C_ψ satisfying

$$C_\psi = \int_{-\infty}^{\infty}\frac{|\psi(\omega)|^2}{|\omega|}d\omega < \infty.$$

Partial Cross Wavelet

- Cross-wavelet coefficients - can be interpreted as a localized measure of correlation between two time series

$$W_{ij}(\tau, s) = W_i(\tau, s)\overline{W_j}(\tau, s)$$

- Partial cross-wavelet

$$W_{ij|V\setminus\{i,j\}}(\tau, s) = W_{ij}(\tau, s) - W_{iV\setminus\{i,j\}}(\tau, s) W_{V\setminus\{i,j\}V\setminus\{i,j\}}(\tau, s)^{-1} W_{jV\setminus\{i,j\}}(\tau, s)$$

It extends a result for partial cross-spectrum (Brillinger, 1981) and involves inversion of $(k-2)\times(k-2)$ dimensional matrix; alternatively solve via recursion formula.

Partial Wavelet Coherence

- Partial wavelet coherence (PWC)

$$R_{ij|V\setminus\{i,j\}}(\tau,s)=\frac{|W_{ij|V\setminus\{i,j\}}(\tau,s)|}{|W_{ii|V\setminus\{i,j\}}(\tau,s)W_{jj|V\setminus\{i,j\}}(\tau,s)|^{\frac{1}{2}}}$$

$0\le|R_{ij|V\setminus\{i,j\}}(\tau,s)|^2\le 1$, interpreted as a localized correlation in the time-frequency domain

Remark. $X_i \perp\!\!\!\perp X_j \mid X_{V\setminus\{i,j\}} \Leftrightarrow R_{ij|V\setminus\{i,j\}}(\tau,s)=0, \forall s,\tau \Leftrightarrow |W_{ij|V\setminus\{i,j\}}(\tau,s)|=0, \forall s,\tau$

Undirected Wavelet Dependence Graph

For $X_V(t)$ a multivariate stochastic process evolving in discrete time a *undirected wavelet dependence graph* is an undirected multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ in which any $v_i \in \mathcal{V}$ encodes the i-th component $X_i(t)$ of $X_V(t)$ s.t. at fixed scale s

$$\begin{aligned} X_{i,s} \perp\!\!\!\perp X_{j,s} \mid X_{V\setminus\{i,j\},s} &\Leftrightarrow e_{ij,s} \notin \mathcal{E}_s \\ &\Leftrightarrow R_{ij|V\setminus\{i,j\}}(\tau, s) = 0, \forall \tau \end{aligned}$$

where $\mathcal{E}_s$ is a scale-specific subset and it holds that $\mathcal{E} = \cup \mathcal{E}_s$.
Remark: A partial correlation graph can be obtained from the

multigraph by replacing any multiedge by a single edge.

Factorization of Wavelet Spectral Matrix

Wavelet spectral matrix $WS(\tau,\omega) = \{WS_{ij}(\tau,\omega)\}_{i,j\in V}$, where entries are frequency specific equivalents of $W_{i,j}(\tau,s)$. For fixed τ (we omit indexing τ for exposition purposes)

$$WS(\tau,\omega) = \Psi_\tau \overline{\Psi_\tau}^\top,$$

where Ψ_τ, the minimum-phase spectral density matrix, produces a causal filter B_τ with a causal inverse s.t.

$$\Psi_\tau(e^{i2\pi\omega}) = \sum_{k=0}^{\infty} B_{\tau,k}(e^{ik2\pi\omega}),$$

error covariance matrix $\Sigma_{\tau,\varepsilon} = B_{\tau,0}B_{\tau,0}^\top$, minimum-phase transfer function $H_\tau = \Psi_\tau B_{\tau,0}^{-1}$. In time domain, $\Psi_\tau(z) = \sum_{k=0}^{\infty} B_{\tau,k} z^k$, with $\Psi_\tau(0) = B_{\tau,0}$ upper triangular matrix with positive diagonal.

Granger Causality Spectra

Geweke (1982), Geweke (1984)

- Pairwise Granger causality (PGC)

$$GC_{i\to j}(\tau,\omega) = \log \frac{WS_{jj}(\tau,\omega)}{WS_{jj}(\tau,\omega) - \left(\Sigma_{\tau,ii} - \Sigma^2_{\tau,ij}/\Sigma_{\tau,jj}\right)|H_{\tau,ij}(\omega)|^2},$$

- Conditional Granger causality (CGC)

$$GC_{i\to j|V\setminus\{i,j\}}(\tau,\omega) = \log \frac{\Sigma_{\tau,jj}(X_i,X_j)}{Q_{jj}(\tau,\omega)\Sigma_{\tau,jj}(X_i,X_j,X_{V\setminus\{i,j\}})\overline{Q_{jj}}^{\top}(\tau,\omega)},$$

where $\Sigma_{\tau,jj}(X_i,X_j)$ and $\Sigma_{\tau,jj}(X_i,X_j,X_{V\setminus\{i,j\}})$ are the variance of the error when regressing X_j on past values of X_i and $X_{V\setminus j}$, Q_{jj} are functions of $\Sigma_{\tau,\varepsilon}$ and H_τ, (see Ding et al., 2006).

Directed Wavelet Dependence Graph

For $X_V(t)$ a multivariate stochastic process evolving in discrete time a *directed wavelet dependence graph* is a directed multiedge graph $\mathcal{G}^{GC} = (\mathcal{V}, \mathcal{E}^{GC})$ in which any $v_i \in \mathcal{V}$ encodes the i-th component $X_i(t)$ of $X_V(t)$ s.t. at fixed scale s

$$\begin{aligned} X_{i,s} \nrightarrow_I X_{j,s} \mid X_{V\setminus\{i,j\},s} &\Leftrightarrow e_{i\to v_j} \notin \mathcal{E}_s^{GC} \\ &\Leftrightarrow GC_{i\to j|V\setminus\{i,j\},s}(\tau) = 0, \forall\tau \end{aligned}$$

where $GC_{ij|V\setminus\{ij\},s}(\tau)$ scale specific version of the CGC, $\mathcal{E}_s^{GC}$ is a scale-specific subset and it holds that $\mathcal{E}^{GC} = \cup\mathcal{E}_s^{GC}$.

Remark: A Granger causality graph can be obtained by replacing same-directional subset of an multiedge by at most one directed edge; together with an undirected simple graph obtained from $\Sigma_{\tau,\varepsilon}$.

Model Selection and Parameter Estimation

- Identify null entries of the precision matrix, Dempster (1972)
- Sparsity: shrinkage, computational savings
- Main approaches
  - Hypothesis testing (Edwards, 2000)
  - Simultaneous confidence interval (Drton and Perlman, 2004)
  - Neighborhood search (Meinshausen and Bühlmann, 2006)
  - Graphical Lasso: Friedman, Hastie and Tibshirani (2008)
  - Bayesian approaches (Wong et al., 2003; Dobra et al., 2004)
  - Greedy methods (Pradeep et al, 2012)
  - Measure method approaches, e.g. Frobenius norm (Rothman et al., 2008; Lam and Fan, 2008)

Conclusions

Wavelet methods

- useful to analyze time-varying nonstationary time series
- recover linear filters and error covariance matrices from spectral representations
- easy to derive the graph structure if new components are added to the MTS

Challenges

- Graph estimation
- Directed graphs for contemporaneous/instantaneous correlations

Barigozzi, M. and Brownlees, C.
NETS: Network Estimation for Time Series
Working paper, 2014

Brillinger, D.R.
Time Series: Data Analysis and Theory
New York: Holt, Rinehart and Winston. 1981

Brillinger, D.R.
Remarks Concerning Graphical Models for Time Series and Point Processes
Revista de Econometrica **16**: 1–23, 1996

Dahlhaus, R.
Graphical Interaction Models for Multivariate Time Series
Metrika **51**: 157–172, 2000

Dahlhaus, R. and Eichler, M.
Causality and Graphical Models in Time Series Analysis
In: P.J. Green, N.L. Hjort, and S. Richardson (Eds.): Highly Structured Stochastic Systems. Oxford University Press, Oxford. 115–137, 2003

Dempster, A.P.
Covariance selection
Biometrics **51**: 157–175, 1972

Ding, M., Chen, Y., Bressler, S.
Granger Causality: Basic Theory and Application to Neuroscience
In: Schelter, B., Winterhalder, M., Timmer, J. (Eds.), Handbook of Time Series Analysis: Recent Theoretical Developments and Applications. Wiley-VCH, Berlin. 437–459, 2006

Eichler, M.
Granger-Causality Graphs for Multivariate Time Series
Technical report, University of Heidelberg, Germany, 2000

Eichler, M.
Granger Causality and Path Diagrams for Multivariate Time Series
Journal of Econometrics **137**: 334–353, 2007

Eckardt, M.
Reviewing Graphical Modelling of Multivariate Temporal Processes
In Adalbert Wilhelm and Hans A. Kestler. (eds.), Analysis of Large and Complex Data, Springer, In press

Geweke, J.
Measurement of Linear-Dependence and Feedback between Multiple Time-Series
J. Am. Stat. Assoc. **77**: 304–313, 1982

Geweke, J.
Inference and Causality in Economic Time Series
In Z. Griliches and M.D. Intriligator (eds.), Handbook of Econometrics, Vol. 2, North-Holland, Amsterdam. 1101–1144, 1984

Granger, C.W.J.
Investigating Causal Relations by Econometric Models and Cross-Spectral Methods
Econometrica. **37**: 424–438, 1969

Pearl, J.
Causality: Models, Reasoning, and Inference
Cambridge University Press, Cambridge, UK. 2000