Graphical Causal Models for Time Series Econometrics: Some Recent Developments and Applications

Alessio Moneta

Max Planck Institute of Economics, Jena

10 December 2009

MAX-PLANCK-GESELLSCHAFT

Mini Symposium: Causality and Time Series Analysis
NIPS 2009, Vancouver

Scope

- Application of methods of causal search to the problem of finding the appropriate causal order for the Structural Vector Autoregressive models (SVAR).

Overview

- VAR and SVAR model
- Causal search methods: graphical models
- Application to the linear/Gaussian setting
- Extensions:
  - Nonparametric setting
  - Non-Gaussian case: application of a method based on ICA

Overview

- ▷ VAR and SVAR model
- ▷ Causal search methods: graphical models
- ▷ Application to the linear/Gaussian setting
- ▷ Extensions:
  - Nonparametric setting
  - Non-Gaussian case: application of a method based on ICA

VAR vs. SVAR model

Basic VAR model (reduced-form):

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \quad (1)$$

- Y_t: $(y_{t1}, \ldots, y_{tk})'$;
- A_j $(j = 1, \ldots, p)$ are $k \times k$ coefficient matrices;
- u_t is the vector white noise process;
- $E(u_t u_t') = \Sigma_u$.

VAR vs. SVAR model

Basic VAR model (reduced-form):

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \tag{1}$$

- Y_t: $(y_{t1}, \ldots, y_{tk})'$;
- A_j $(j = 1, \ldots, p)$ are $k \times k$ coefficient matrices;
- u_t is the vector white noise process;
- $E(u_t u_t') = \Sigma_u$.

VAR vs. SVAR model

Basic VAR model (reduced-form):

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \tag{1}$$

- Y_t: $(y_{t1}, \ldots, y_{tk})'$;
- A_j $(j = 1, \ldots, p)$ are $k \times k$ coefficient matrices;
- u_t is the vector white noise process;
- $E(u_t u_t') = \Sigma_u$.

VAR vs. SVAR model

Basic VAR model (reduced-form):

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \tag{1}$$

- Y_t: $(y_{t1}, \ldots, y_{tk})'$;
- A_j $(j = 1, \ldots, p)$ are $k \times k$ coefficient matrices;
- u_t is the vector white noise process;
- $E(u_t u_t') = \Sigma_u$.

VAR vs. SVAR model

Wold representation (in case of stationarity):

$$Y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j}, \tag{2}$$

where $\Phi_j = \sum_{i=1}^{j} \Phi_{j-i} A_i$

But for any nonsingular $k \times k$ matrix P we get:

$$Y_t = \sum_{j=0}^{\infty} \Phi_j P P^{-1} u_{t-j} = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \tag{3}$$

where $\varepsilon_{t-j} = P^{-1} u_{t-j}$ and $\Psi_j = \Phi_j P$ $(j = 0, 1, 2, ...)$.

VAR vs. SVAR model

Wold representation (in case of stationarity):

$$Y_t = \sum_{j=0}^{\infty} \Phi_j u_{t-j}, \tag{2}$$

where $\Phi_j = \sum_{i=1}^{j} \Phi_{j-i} A_i$

But for any nonsingular $k \times k$ matrix P we get:

$$Y_t = \sum_{j=0}^{\infty} \Phi_j P P^{-1} u_{t-j} = \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j}, \tag{3}$$

where $\varepsilon_{t-j} = P^{-1} u_{t-j}$ and $\Psi_j = \Phi_j P$ $(j = 0, 1, 2, \ldots)$.

VAR vs. SVAR model

If we premultiply equation (1) by P^{-1} we get

$$P^{-1}Y_t = P^{-1}A_1 Y_{t-1} + \ldots + P^{-1}A_p Y_{t-p} + P^{-1}u_t. \tag{4}$$

SVAR model (structural form):

$$\Gamma_0 Y_t = \Gamma_1 Y_{t-1} + \ldots + \Gamma_p Y_{t-p} + \varepsilon_t, \tag{5}$$

where $\Gamma_0 = P^{-1}, \Gamma_j = P^{-1}A_j$ $(j = 1, \ldots, p)$.

$\triangleright$ choice of P (Γ_0) based on information about the contemporaneous causal structure.

VAR vs. SVAR model

If we premultiply equation (1) by P^{-1} we get

$$P^{-1}Y_t = P^{-1}A_1 Y_{t-1} + \ldots + P^{-1}A_p Y_{t-p} + P^{-1}u_t. \tag{4}$$

SVAR model (structural form):

$$\Gamma_0 Y_t = \Gamma_1 Y_{t-1} + \ldots + \Gamma_p Y_{t-p} + \varepsilon_t, \tag{5}$$

where $\Gamma_0 = P^{-1}$, $\Gamma_j = P^{-1}A_j$ $(j = 1, \ldots, p)$.

▷ choice of P (Γ_0) based on information about the contemporaneous causal structure.

VAR vs. SVAR model

If we premultiply equation (1) by P^{-1} we get

$$P^{-1}Y_t = P^{-1}A_1Y_{t-1} + \ldots + P^{-1}A_pY_{t-p} + P^{-1}u_t. \tag{4}$$

SVAR model (structural form):

$$\Gamma_0 Y_t = \Gamma_1 Y_{t-1} + \ldots + \Gamma_p Y_{t-p} + \varepsilon_t, \tag{5}$$

where $\Gamma_0 = P^{-1}$, $\Gamma_j = P^{-1}A_j$ $(j = 1, \ldots, p)$.

$\triangleright$ choice of P (Γ_0) based on information about the contemporaneous causal structure.

VAR vs. SVAR model

- choice of P (Γ_0) in the literature:
  - *Choleski* decomposition such that: P is lower diagonal and $\Omega = E(\varepsilon_t \varepsilon_t') = I_k$.
  - a priori (theoretical, institutional) zero-restrictions;
- Our proposal: inferring P (Γ_0) starting from the estimated residuals $\hat{u}_t$
  - conditional independence relations $\longrightarrow$ causal relationships (graphical models)
  - independent component analysis (in case of non-Gaussianity)

VAR vs. SVAR model

▷ choice of P (Γ_0) in the literature:

- *Choleski* decomposition such that: P is lower diagonal and $\Omega = E(\varepsilon_t \varepsilon_t') = I_k$.
- a priori (theoretical, institutional) zero-restrictions;

▷ Our proposal: inferring P (Γ_0) starting from the estimated residuals $\hat{u}_t$

- conditional independence relations $\longrightarrow$ causal relationships (graphical models)
- independent component analysis (in case of non-Gaussianity)

Graphical models

Graphs have two functions:

- Representation of causal structures
- Representation of causal independence relations

Graphical models

Representation of causal structures:

▷ edge: causal influence

- Undirected edges: $X — Y$ (ambiguous causal influence between X and Y)
- Directed edges: $X \longrightarrow Y$, $X \longleftarrow Y$
- Bi-directed edges: $X \longleftrightarrow Y$

▷ DAGs: Directed acyclic graphs (only directed edges).

Graphical models

Representation of conditional independence relations:

▷ edge: statistical dependence

- $X \perp\!\!\!\perp Y \mid \emptyset$: no edge between X and Y
- $X \not\!\perp\!\!\!\perp Y$: $X - Y$; $X \longrightarrow Y$; $X \longleftarrow Y$
- $X \perp\!\!\!\perp Z \mid Y$: $X \longrightarrow Y \longrightarrow Z$; $X \longleftarrow Y \longleftarrow Z$; $X \longleftarrow Y \longrightarrow Z$
- $X \not\!\perp\!\!\!\perp Z \mid Y$: $X \longrightarrow Y \longleftarrow Z$

Graphical models

Rules of Inference:

▷ edge: from conditional independence $\mapsto$ causal relationships

- DAGs among $V_1, \ldots, V_k$
  - Causal Markov Condition: *conditioned on its parents every node is independent of its nondescendants*; or: *conditioned on its direct causes every variable is independent of its non-effects*
  - Faithfulness Condition: *every conditional independence relation among* $V_1, \ldots, V_k$ *is entailed by the Causal Markov Condition*

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges

- search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges

- $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
- orient edges in order to avoid cycles

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges

- search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges

- $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
- orient edges in order to avoid cycles

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges
  - search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges
  - $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
  - orient edges in order to avoid cycles

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges

- search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges

- $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
- orient edges in order to avoid cycles

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges

- search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges

- $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
- orient edges in order to avoid cycles

Search algorithm:

PC, SGS, modified PC algorithm (Spirtes, Glymour and Scheines 2000):

- Input: conditional independence tests
- Output: set of Markov equivalent DAGs

▷ Start: complete undirected graph among $V_1, \ldots, V_k$

▷ First step: elimination of edges whenever $\perp\!\!\!\perp$

▷ Second step: statistical orientation of edges

  - search for *unshielded colliders*: $X \longrightarrow Y \longleftarrow Z$

▷ Third step: logical orientation of edges

  - $X \longrightarrow Y — Z \quad \mapsto \quad X \longrightarrow Y \longrightarrow Z$
  - orient edges in order to avoid cycles

GMs applied to SVAR

- Cfr. Swanson and Granger (1997), Bessler and Lee (2002), Demiralp and Hoover (2003) (among others)
- Estimate reduced form VAR

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \tag{6}$$

- King-Stock-Plosser-Watson (1991) updated data set
  US data 1947:2 - 1994:1 (quarterly data)

$$Y = \begin{pmatrix} C \\ I \\ M \\ Y \\ R \\ \Delta P \end{pmatrix} \begin{matrix} \text{per capita consumption} \\ \text{per capita investment} \\ \text{money M2 / price} \\ \text{per capita private income} \\ \text{nominal interest rate} \\ \text{price inflation} \end{matrix}$$

- Taking into account non-stationarity / cointegration
- Get the matrix of residuals $\hat{U}_t$

GMs applied to SVAR

- Cfr. Swanson and Granger (1997), Bessler and Lee (2002), Demiralp and Hoover (2003) (among others)
- Estimate reduced form VAR

$$Y_t = A_1 Y_{t-1} + \ldots + A_p Y_{t-p} + u_t. \tag{6}$$

- King-Stock-Plosser-Watson (1991) updated data set
  US data 1947:2 - 1994:1 (quarterly data)

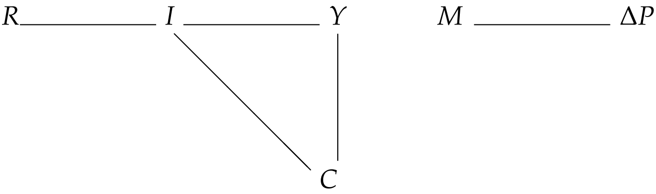
$$Y = \begin{pmatrix} C \\ I \\ M \\ Y \\ R \\ \Delta P \end{pmatrix} \begin{matrix} \text{per capita consumption} \\ \text{per capita investment} \\ \text{money M2 / price} \\ \text{per capita private income} \\ \text{nominal interest rate} \\ \text{price inflation} \end{matrix}$$

- Taking into account non-stationarity / cointegration
- Get the matrix of residuals $\hat{U}_t$

GMs applied to SVAR

- Causal graph among $u_{1t}, \ldots, u_{kt}$ $\equiv$ causal graph among $y_{1t}, \ldots, y_{kt}$
- C.I. relations among $u_{1t}, \ldots, u_{kt}$ tested via Wald tests on zero-partial correlations
- Gaussianity assumption

Results (Moneta 2008):

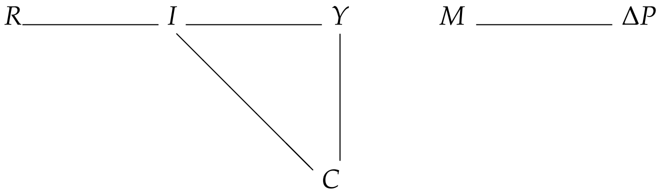


Configurations $R \longrightarrow I \longleftarrow Y$ and $R \longrightarrow I \longleftarrow C$ are excluded.

Possibilities:

- sensitivity analysis
- bootstrap analysis (cfr. Demiralp, Hoover and Perez 2008)

Results (Moneta 2008):



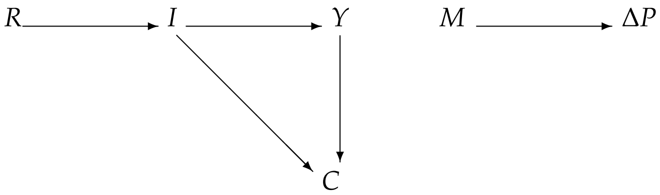
Configurations $R \longrightarrow I \longleftarrow Y$ and $R \longrightarrow I \longleftarrow C$ are excluded.

Possibilities:

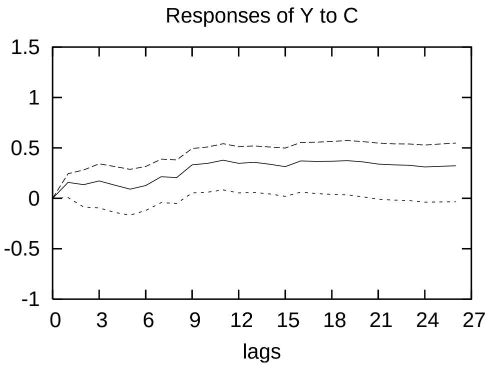
- ▷ sensitivity analysis
- ▷ bootstrap analysis (cfr. Demiralp, Hoover and Perez 2008)

Results (Moneta 2008):

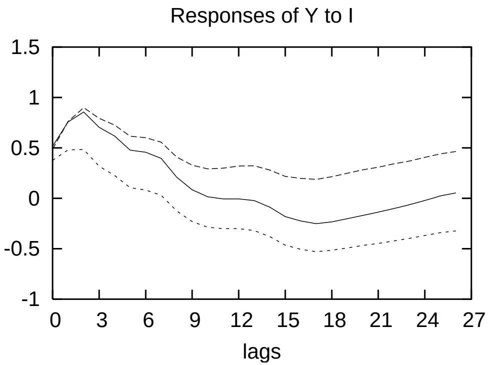
One of the 16 DAGs:



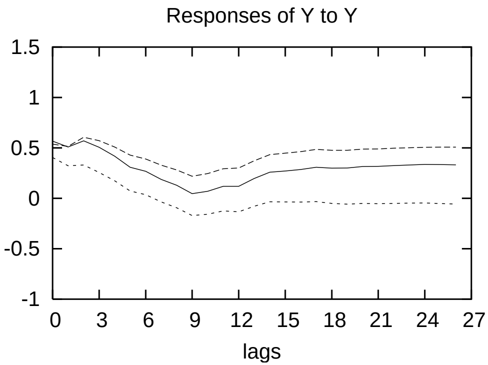
Impulse Response Analysis



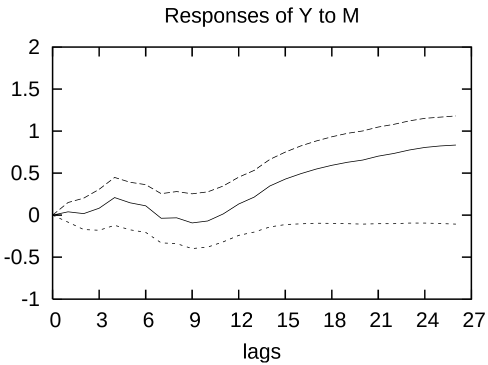
Impulse Response Analysis



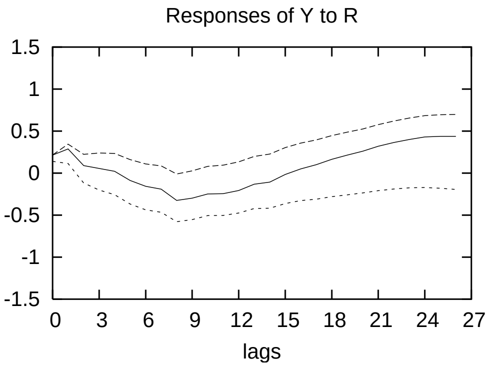
Impulse Response Analysis



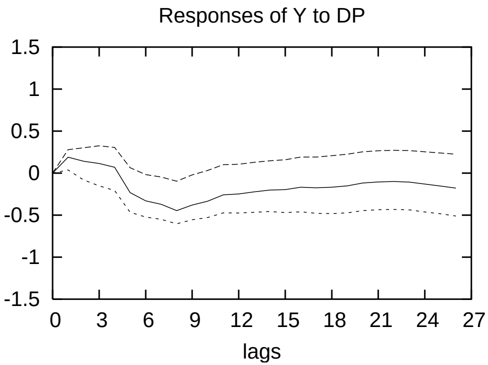
Impulse Response Analysis



Impulse Response Analysis



Impulse Response Analysis



Extensions

- Non-parametric approach
  - Start from non-parametric tests of conditional independence
- Semi-parametric approach
  - Assumption of linearity and non-Gaussianity

Extensions

- Non-parametric approach
  - Start from non-parametric tests of conditional independence
- Semi-parametric approach
  - Assumption of linearity and non-Gaussianity

Nonparametric approach

Chlaß and Moneta (2009):

- $X \perp\!\!\!\perp Y \mid Z$ iff $f(X|Y,Z) = f(X|Z)$
- Test $\hat{f}(X,Y,Z)\hat{f}(Z) = \hat{f}(X,Z)\hat{f}(Y,Z)$.
- Distance measures between kernel density functions:
  - Euclidean distance (Szekely and Rizzo 2004; Baringhaus and Franz 2004)
  - Weighted Hellinger distance (Su and White 2008)
- Curse of dimensionality
- Bootstrap procedure

Nonparametric approach

Chlaß and Moneta (2009):

- $X \perp\!\!\!\perp Y \mid Z$ iff $f(X|Y,Z) = f(X|Z)$
- Test $\hat{f}(X,Y,Z)\hat{f}(Z) = \hat{f}(X,Z)\hat{f}(Y,Z)$.
- Distance measures between kernel density functions:
  - Euclidean distance (Szekely and Rizzo 2004; Baringhaus and Franz 2004)
  - Weighted Hellinger distance (Su and White 2008)
- Curse of dimensionality
- Bootstrap procedure

Nonparametric approach

Chlaß and Moneta (2009):

- $X \perp\!\!\!\perp Y \mid Z \quad \text{iff} \quad f(X|Y,Z) = f(X|Z)$
- Test $\hat{f}(X,Y,Z)\hat{f}(Z) = \hat{f}(X,Z)\hat{f}(Y,Z)$.
- Distance measures between kernel density functions:
  - Euclidean distance (Szekely and Rizzo 2004; Baringhaus and Franz 2004)
  - Weighted Hellinger distance (Su and White 2008)
- Curse of dimensionality
- Bootstrap procedure

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Causal search based on ICA

Moneta, Entner, Hoyer and Coad (2009)

▷ VAR-LiNGAM algorithm (Hyvärinen, Shimizu and Hoyer 2008)

- Estimate the reduced-form VAR
- Check that the residuals are non-Gaussian
- Use an ICA algorithm to decompose the residuals matrix
- Order the variables (residuals) so as to obtain a matrix P that is close to lower triangular
- Once the instantaneous effects are identified, identify lagged effects
- No need of assumptions such as Faithfulness.

Empirical Application

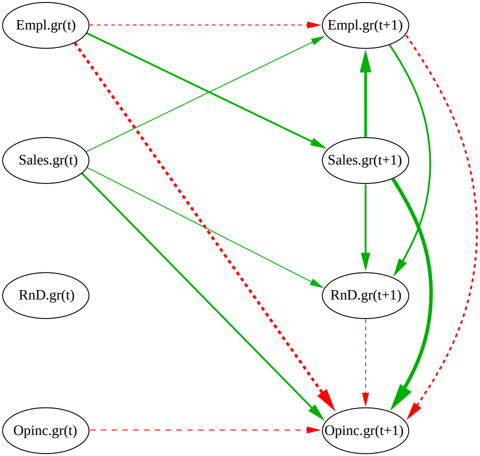
Panel VAR on Firm Growth and R&D Expenditures

Coad-Rao (2007) data set: US firm 1973:2004 (manufacturing sector).

- Employees (growth rate)
- Total sales (growth rate)
- R&D expenditure (growth rate)
- Profits (growth rate)

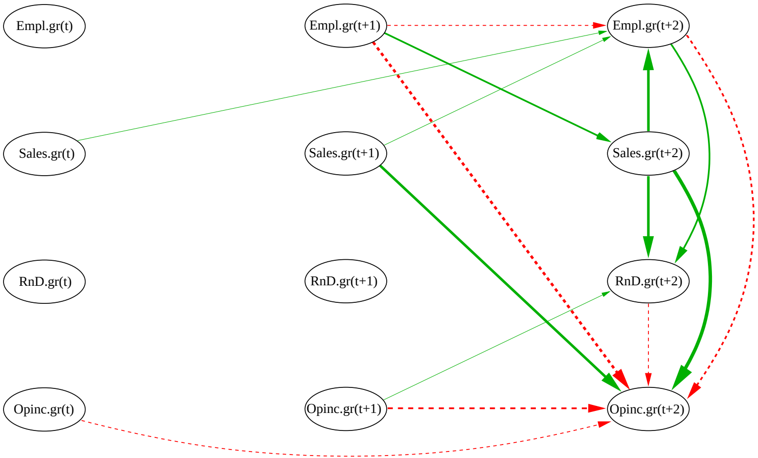
LAD (least absolute deviation) estimation

Structural coefficients one-lag model:



Solid green arrows: positive causal influence; dashed red arrows: negative influence. Only major effects are displayed.

Structural coefficients two-lags model:



Solid green arrows: positive causal influence; dashed red arrows: negative influence. Only major effects are displayed.

Conclusions

- ▷ Identification of causal structure in time series / non-experimental settings
- ▷ SVAR model: possibility of applying causal search to i.i.d. residuals
- ▷ Graphical causal search applied to linear / Gaussian data
- ▷ Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- ▷ Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Conclusions

- ▷ Identification of causal structure in time series / non-experimental settings
- ▷ SVAR model: possibility of applying causal search to i.i.d. residuals
- ▷ Graphical causal search applied to linear / Gaussian data
- ▷ Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- ▷ Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Conclusions

- ▷ Identification of causal structure in time series / non-experimental settings
- ▷ SVAR model: possibility of applying causal search to i.i.d. residuals
- ▷ Graphical causal search applied to linear / Gaussian data
- ▷ Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- ▷ Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Conclusions

- ▷ Identification of causal structure in time series / non-experimental settings
- ▷ SVAR model: possibility of applying causal search to i.i.d. residuals
- ▷ Graphical causal search applied to linear / Gaussian data
- ▷ Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- ▷ Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Conclusions

- Identification of causal structure in time series / non-experimental settings
- SVAR model: possibility of applying causal search to i.i.d. residuals
- Graphical causal search applied to linear / Gaussian data
- Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Conclusions

- ▷ Identification of causal structure in time series / non-experimental settings
- ▷ SVAR model: possibility of applying causal search to i.i.d. residuals
- ▷ Graphical causal search applied to linear / Gaussian data
- ▷ Assumptions (rules of inference): Causal Markov and Faithfulness Condition
- ▷ Extensions:
  - Nonparametric conditional independence tests. No distributional assumptions but same rules of inference.
  - Causal search based on ICA. Semiparametric approach. Weaker assumptions.

Thank you!

References

Baringhaus, L. and C. Franz, 2004, On a new multivariate two-sample test, *Journal of Multivariate Analysis*, 88(1), 190-206.

Bessler, D.A., and S. Lee, 2002, Money and prices: US data 1869-1914 (a study with directed graphs), *Empirical Economics*, 27, 427-446.

Chlaß, N., and A. Moneta, 2009, Can Graphical Causal Inference Be Extended to Nonlinear Settings?, An Assessment of Conditional Independence Tests, in in M. Dorato, M. Redei, M. Suárez (eds.) EPSA Epistemology and Methodology of Science, Springer Verlag.

Demiralp, S., and K.D. Hoover, 2003, Searching for the Causal Structure of a Vector Autoregression, *Oxford Bulletin of Economics and Statistics*, 65, 745-767.

Demiralp, S., K.D. Hoover and S.J. Perez, 2008, A Bootstrap Method for Identifying and Evaluating a Structural Vector Autoregression, *Oxford Bulletin of Economics and Statistics*.

Hyvärinen, A. and S. Shimizu and P. O. Hoyer, 2008, Causal modelling combining instantaneous and lagged effects: an identifiable model based on non-Gaussianity, Proceedings of the 25th International Conference on Machine Learning, 424-431.

References

Moneta, A. 2008, Graphical Causal Models and VARs: an Empirical Assessment of the Real Business Cycles Hypothesis, *Empirical Economics*.

Moneta, A., D. Entner, P. Hoyer, and A. Coad, 2009, Causal Inference by Independent Component Analysis with Applications to Micro- and Macroeconomic Data, mimeo.

Szekely, G. J. and M.L. Rizzo, 2004, Testing for Equal Distributions in High Dimension, *InterStat*, November (5).

Spirtes, P., C. Glymour, and R. Scheines, 2000, *Causation, Prediction, and Search*, The MIT Press.

Su, L. and H. White, 2008, A Nonparametric Hellinger Metric Test for Conditional Independence, *Econometric Theory*, 24, 829-864.

Swanson, N.R., and C.W.J. Granger, 1997, Impulse Response Function Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions, *Journal of the American Statistical Association*, 92(437), 357-367.

.