Faculty of Science

Recurrent event time modeling: predictive and causal models

Niels Richard Hansen, Patricia Reynaud-Bouret,
Vincent Rivoirard and Alexander Sokol
Department of Mathematical Sciences

December 14, 2016

Current interests in Hawkes processes

Hawkes processes ... have become very popular in empirical high frequency finance this last decade.

Hawkes Processes in Finance
E. Bacry et al., Market Microstructure and Liquidity, 2015

Current interests in Hawkes processes

Hawkes processes ... have become very popular in empirical high frequency finance this last decade.

Hawkes Processes in Finance
E. Bacry et al., Market Microstructure and Liquidity, 2015

... applications include the user viewing records in an IPTV system, and the patient records in hospitals, ...

Learning Granger Causality for Hawkes Processes
H. Xu et al., ICML, 2016

Current interests in Hawkes processes

Hawkes processes ... have become very popular in empirical high frequency finance this last decade.

Hawkes Processes in Finance
E. Bacry et al., Market Microstructure and Liquidity, 2015

... applications include the user viewing records in an IPTV system, and the patient records in hospitals, ...

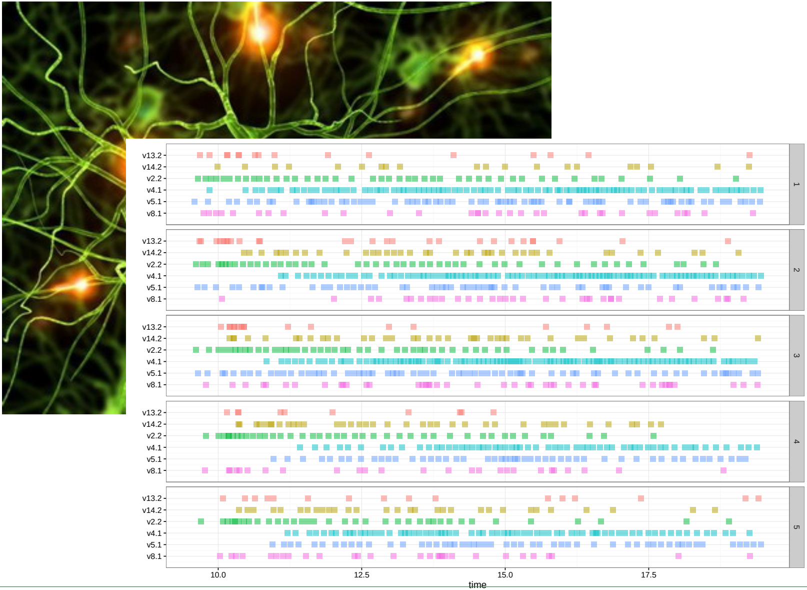
Learning Granger Causality for Hawkes Processes
H. Xu et al., ICML, 2016

In a variety of settings, we observe a series of events associated with a group of actors whose interactions trigger future events. The interactions between these actors can be modeled using a network. For example: **Social networks; Biological neural networks; Financial networks; Epidemiological networks; Seismological networks.**

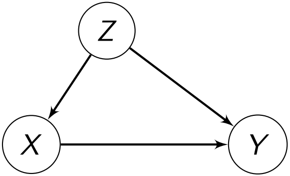
Tracking Dynamic Point Processes on Networks
E. C. Hall and R. M. Willet, IEEE Transactions on Information Theory, 2016

Neuron spike data example

Neuron spike data example

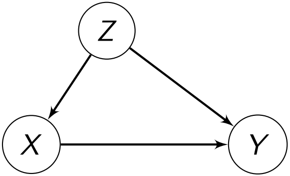


Directed acyclic graphs and structural equations



$$
\begin{aligned}
Z &= \varepsilon_Z \\
X &= F(Z, \varepsilon_X) \\
Y &= G(X, Z, \varepsilon_Y)
\end{aligned}
$$

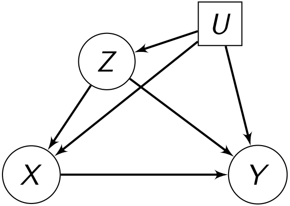
Directed acyclic graphs and structural equations



$$
\begin{aligned}
Z &= \varepsilon_Z \\
X &= F(Z, \varepsilon_X) \\
Y &= G(X, Z, \varepsilon_Y)
\end{aligned}
$$

$$\varepsilon_Z \perp\!\!\!\perp \varepsilon_X \perp\!\!\!\perp \varepsilon_Y.$$

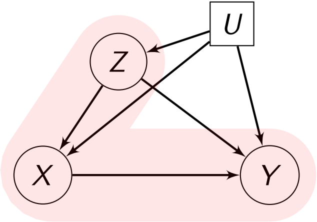
Directed acyclic graphs and structural equations



$$\begin{aligned} U &= \varepsilon_U \\ Z &= H(U, \varepsilon_Z) \\ X &= F(Z, U, \varepsilon_X) \\ Y &= G(X, Z, U, \varepsilon_Y) \end{aligned}$$

$$\varepsilon_U \perp\!\!\!\perp \varepsilon_Z \perp\!\!\!\perp \varepsilon_X \perp\!\!\!\perp \varepsilon_Y.$$

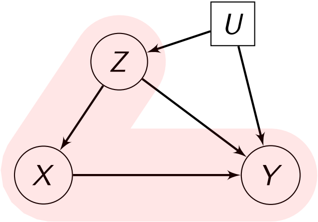
Directed acyclic graphs and structural equations



$$
\begin{aligned}
U &= \varepsilon_U \\
Z &= H(U, \varepsilon_Z) \\
X &= F(Z, U, \varepsilon_X) \\
Y &= G(X, Z, U, \varepsilon_Y)
\end{aligned}
$$

$$\varepsilon_U \perp\!\!\!\perp \varepsilon_Z \perp\!\!\!\perp \varepsilon_X \perp\!\!\!\perp \varepsilon_Y.$$

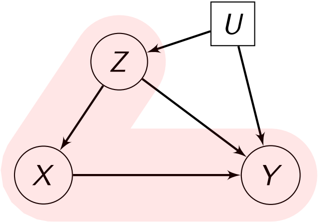
Directed acyclic graphs and structural equations



$$
\begin{aligned}
U &= \varepsilon_U \\
Z &= H(U, \varepsilon_Z) \\
X &= F(Z, \varepsilon_X) \\
Y &= G(X, Z, {\color{red}U}, \varepsilon_Y)
\end{aligned}
$$

$$\varepsilon_U \perp\!\!\!\perp \varepsilon_Z \perp\!\!\!\perp \varepsilon_X \perp\!\!\!\perp \varepsilon_Y.$$

Directed acyclic graphs and structural equations

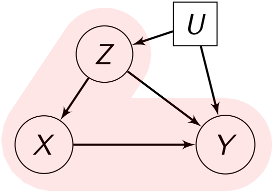


$$
\begin{aligned}
U &= \varepsilon_U \\
Z &= H(U, \varepsilon_Z) \\
X &= F(Z, \varepsilon_X) \\
Y &= G(X, Z, U, \varepsilon_Y)
\end{aligned}
$$

$$\varepsilon_U \perp\!\!\!\perp \varepsilon_Z \perp\!\!\!\perp \varepsilon_X \perp\!\!\!\perp \varepsilon_Y.$$

Directed acyclic graphs (DAGs) and structural equation models (SEMs) specify probability models with a potential causal interpretation.

DAGs, causality and Markov properties

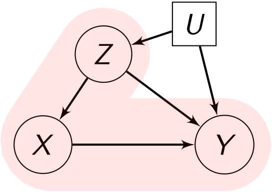


$$Y = G(X, Z, U, \varepsilon_Y)$$

Pearl's do-conditioning defines causal effects

$$Y \mid do(X = x) \stackrel{\mathcal{D}}{=} G(x, Z, U, \varepsilon_Y).$$

DAGs, causality and Markov properties



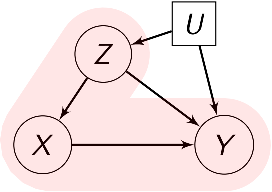
$$Y = G(X, Z, U, \varepsilon_Y)$$

Pearl's do-conditioning defines causal effects

$$Y \mid do(X = x) \stackrel{\mathcal{D}}{=} G(x, Z, U, \varepsilon_Y).$$

Pearl's back-door criterion is a sufficient condition for identifiability of causal effects in the presence of unobserved confounders.

DAGs, causality and Markov properties



$$Y = G(X, Z, U, \varepsilon_Y)$$

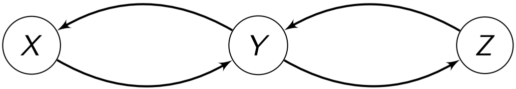
Pearl's do-conditioning defines causal effects

$$Y \mid do(X = x) \stackrel{\mathcal{D}}{=} G(x, Z, U, \varepsilon_Y).$$

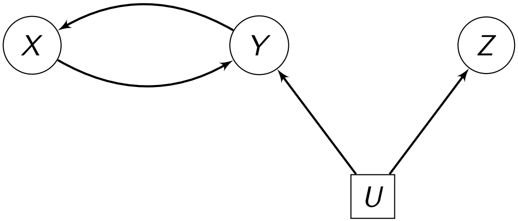
Pearl's back-door criterion is a sufficient condition for identifiability of causal effects in the presence of unobserved confounders.

A causal model implies Markov properties (cond. independences), which can be exploited by causal search algorithms.

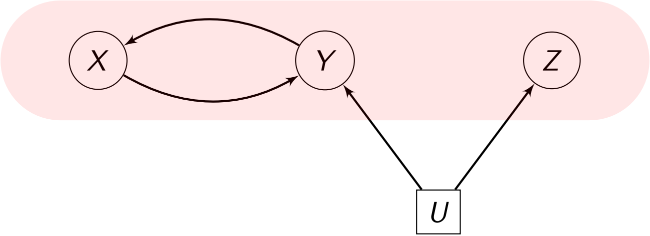
Dynamics, feedback and latent variables



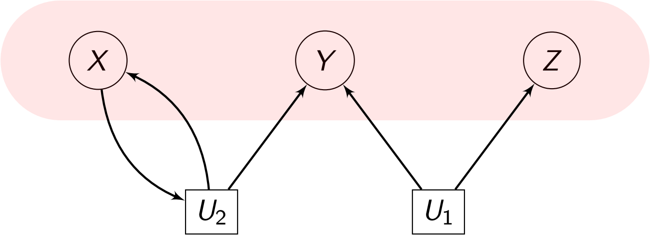
Dynamics, feedback and latent variables



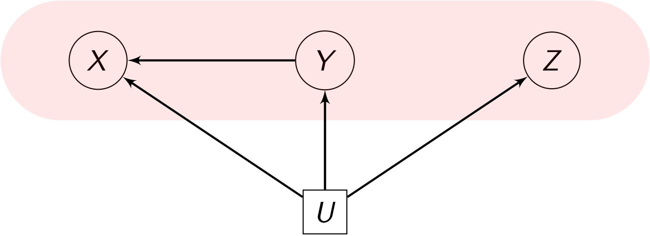
Dynamics, feedback and latent variables



Dynamics, feedback and latent variables



Dynamics, feedback and latent variables



Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

- How can we handle feedback that doesn't fit into the framework of DAGs?

Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

- How can we handle feedback that doesn't fit into the framework of DAGs?
    - But it does if time is included explicitly.

Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

- How can we handle feedback that doesn't fit into the framework of DAGs?
  - But it does if time is included explicitly.
- What are the consequences of latent variables for statistical inference?

Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

- How can we handle feedback that doesn't fit into the framework of DAGs?
  - But it does if time is included explicitly.
- What are the consequences of latent variables for statistical inference? Two solutions:
  - model latent variables explicitly,
  - or characterize (parametrize) marginalized distributions.

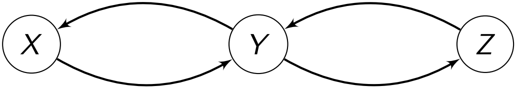
Questions

We are interested in models that are not just predictive but also have a causal interpretation. That is, they are predictive about intervention effects.

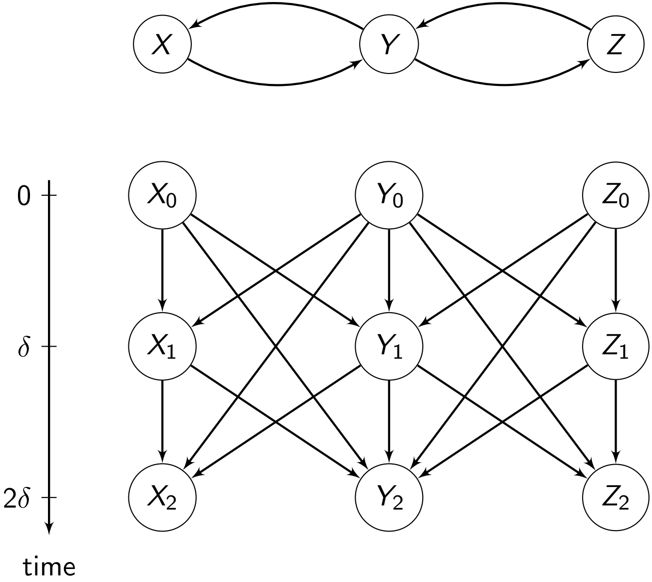
- How can we handle feedback that doesn't fit into the framework of DAGs?
  - But it does if time is included explicitly.
- What are the consequences of latent variables for statistical inference? Two solutions:
  - model latent variables explicitly,
  - or characterize (parametrize) marginalized distributions.

We consider dynamic Markov properties and graphs for event times.

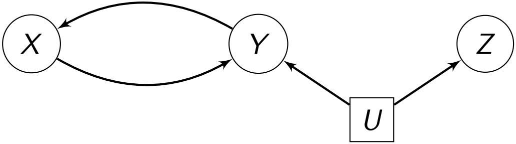
Dynamics, feedback and latent variables



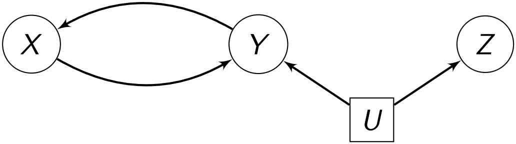
Dynamics, feedback and latent variables

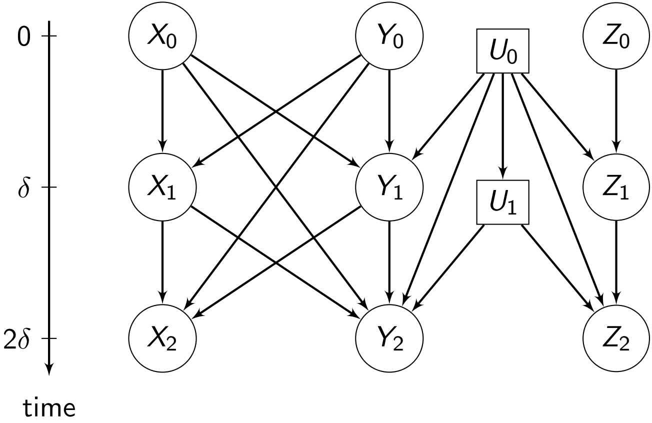


Dynamics, feedback and latent variables



Dynamics, feedback and latent variables





Important observation

For a stationary dynamic process

dynamic Markov properties do not translate into Markov properties for the invariant distribution – the latter may have no conditional independencies.

Important observation

For a stationary dynamic process

dynamic Markov properties do not translate into Markov properties for the invariant distribution – the latter may have no conditional independencies.

It is thus particularly difficult to draw causal inference about dynamical processes from cross-sectional samples.

Events and intensity models

The counting process for the kth event track will be modeled in terms of an intensity:

$$P(N^k_{t+\delta} - N^k_t = 1 \mid \mathcal{F}_t) \simeq \lambda^k_t \delta.$$

Events and intensity models

The counting process for the kth event track will be modeled in terms of an intensity:

$$P(N^k_{t+\delta} - N^k_t = 1 \mid \mathcal{F}_t) \simeq \lambda^k_t \delta.$$

The process λ^k_t is predictable w.r.t. the filtration $\mathcal{F}_t$.

Events and intensity models

The counting process for the kth event track will be modeled in terms of an intensity:

$$P(N^k_{t+\delta} - N^k_t = 1 \mid \mathcal{F}_t) \simeq \lambda^k_t \delta.$$

The process λ^k_t is predictable w.r.t. the filtration $\mathcal{F}_t$.

The filtration $\mathcal{F}_t$ holds information on all event tracks up to time t.

Events and intensity models

The counting process for the kth event track will be modeled in terms of an intensity:

$$P(N^k_{t+\delta} - N^k_t = 1 \mid \mathcal{F}_t) \simeq \lambda^k_t \delta.$$

The process λ^k_t is predictable w.r.t. the filtration $\mathcal{F}_t$.

The filtration $\mathcal{F}_t$ holds information on all event tracks up to time t.

In the neuron science literature,

$$N^k_{t+\delta} - N^k_t \mid \mathcal{F}_t \sim \mathrm{Poi}(\lambda^k_t \delta)$$

is applied to binned data for small enough δ (nonlinear, non-Gaussian time series).

Local Independence

Let $V = \{1, \ldots, d\}$ denote the set of track indices.

We say that k is conditionally locally independent of j and write

$$j \not\to k \mid V \backslash \{j, k\} \tag{1}$$

if λ_t^k does not depend on track j.

Local Independence

Let $V = \{1, \dots, d\}$ denote the set of track indices.

We say that k is conditionally locally independent of j and write

$$j \not\to k \mid V \backslash \{j, k\} \tag{1}$$

if λ_t^k does not depend on track j.

Definition (Local Independence Graph)

A graph $G = (V, E)$ is called a local independence graph if

$$(j, k) \notin E \Longrightarrow j \not\to k \mid V \backslash \{j, k\}.$$

A local independence graph directly encodes pairwise dynamic Markov properties.

The global Markov property

Let $A, B, C \subseteq V$ be disjoint.

We say that B is locally independent of A given C and write

$$A \not\rightarrow B \mid C \tag{2}$$

if $E(\lambda_t^k \mid \sigma(N_{0:t}^{A \cup B \cup C}))$ does not depend on tracks in A for $k \in B$.

The global Markov property

Let $A, B, C \subseteq V$ be disjoint.

We say that B is locally independent of A given C and write

$$A \not\rightarrow B \mid C \tag{2}$$

if $E(\lambda_t^k \mid \sigma(N_{0:t}^{A \cup B \cup C}))$ does not depend on tracks in A for $k \in B$.

Theorem (Didelez[1])

If C δ-separates A from B in a local independence graph then (2) holds.

[1]Graphical models for marked point processes based on local independence. JRSS-B 70(1), 2008.

The global Markov property

- Local independence graphs can have cycles and can represent feedback.

The global Markov property

- Local independence graphs can have cycles and can represent feedback.
- They can be given a causal interpretation e.g. via interventions.

The global Markov property

- Local independence graphs can have cycles and can represent feedback.
- They can be given a causal interpretation e.g. via interventions.
- δ-separation generalizes d-separation known for DAGs.

The global Markov property

- Local independence graphs can have cycles and can represent feedback.
- They can be given a causal interpretation e.g. via interventions.
- δ-separation generalizes *d*-separation known for DAGs.
- Some technical conditions are required.

The global Markov property

- Local independence graphs can have cycles and can represent feedback.
- They can be given a causal interpretation e.g. via interventions.
- δ-separation generalizes d-separation known for DAGs.
- Some technical conditions are required.

Definition

A graph G is faithful if

$$C \ \delta\text{-separates } A \text{ from } B \iff A \not\rightarrow B \mid C.$$

The global Markov property

- Local independence graphs can have cycles and can represent feedback.
- They can be given a causal interpretation e.g. via interventions.
- δ-separation generalizes *d*-separation known for DAGs.
- Some technical conditions are required.

Definition

A graph G is faithful if

$$C \ \delta\text{-separates } A \text{ from } B \iff A \not\to B \mid C.$$

Faithfulness + the global Markov property allows for inference about G from marginal observations.

Linear filters and the Hawkes process

The intensity can be specified in terms of linear filters of preceeding events

$$\lambda_t^k = \beta_0^k + \sum_{j=1}^{p} \sum_{i:\tau_i^j < t} g^{jk}(t - \tau_i^j)$$

Linear filters and the Hawkes process

The intensity can be specified in terms of linear filters of preceeding events

$$\lambda_t^k = \beta_0^k + \sum_{j=1}^{p} \sum_{i:\tau_i^j < t} g^{jk}(t - \tau_i^j)$$

This gives the (linear) Hawkes process.

The nonlinear Hawkes process

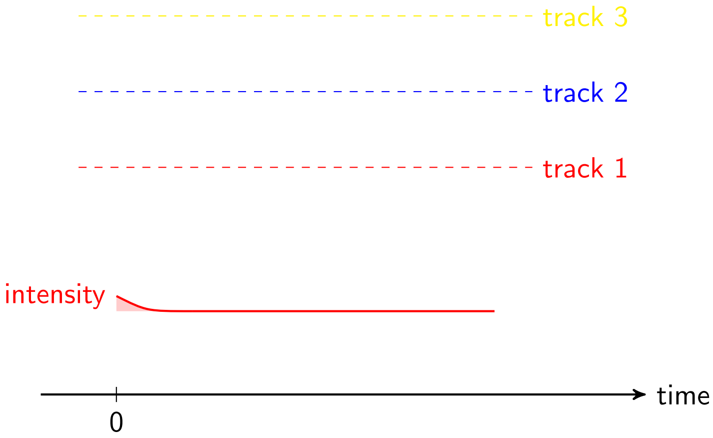
The intensity

$$\lambda_t^k(g) = \varphi\Big(\beta_0^k + \sum_{j=1}^{p} \sum_{i:\tau_i^j < t} g^{ik}(t - \tau_i^j)\Big),$$

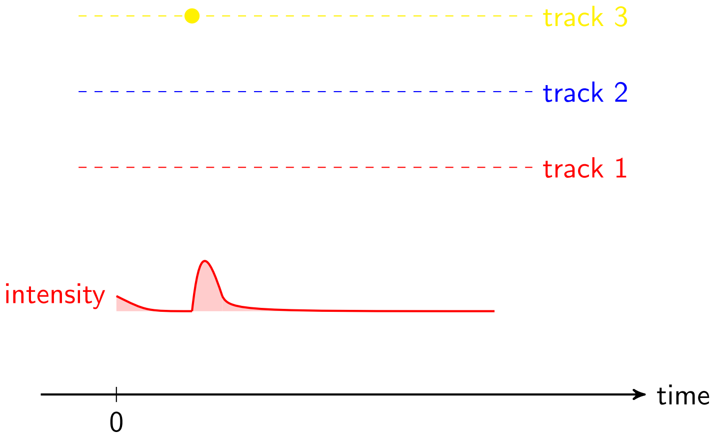
for $\varphi : \mathbb{R} \to [0, \infty)$ gives the nonlinear Hawkes process[1].

[1] Brémaud, P. and Massoulié, L. *Stability of nonlinear Hawkes processes*. Ann. Probab. 24(3), 1996..

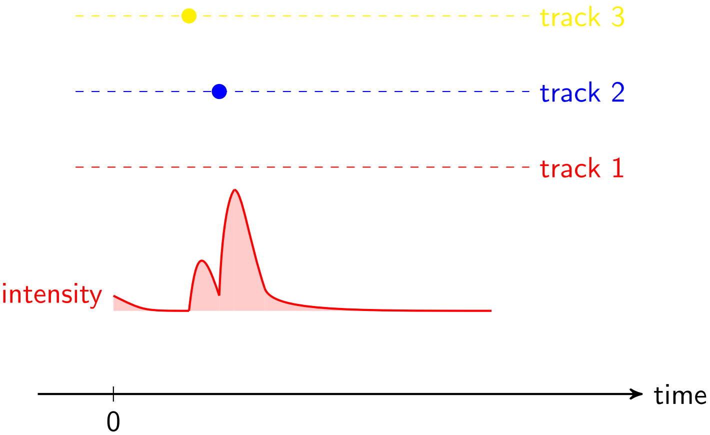
Linear filters



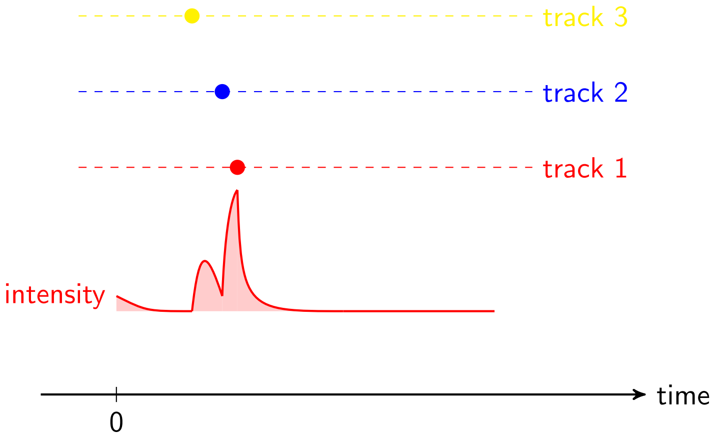
Linear filters



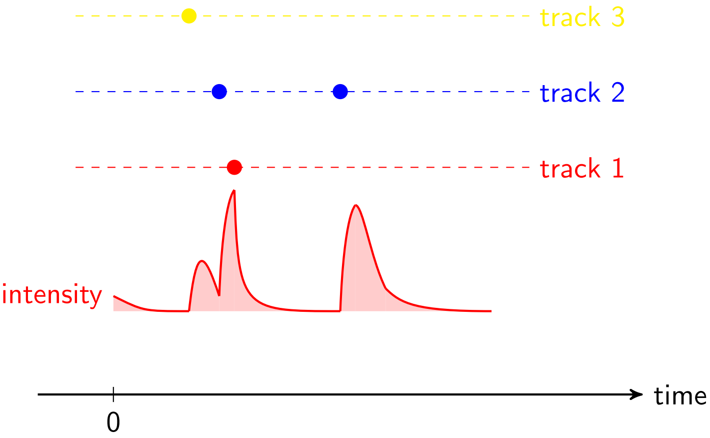
Linear filters



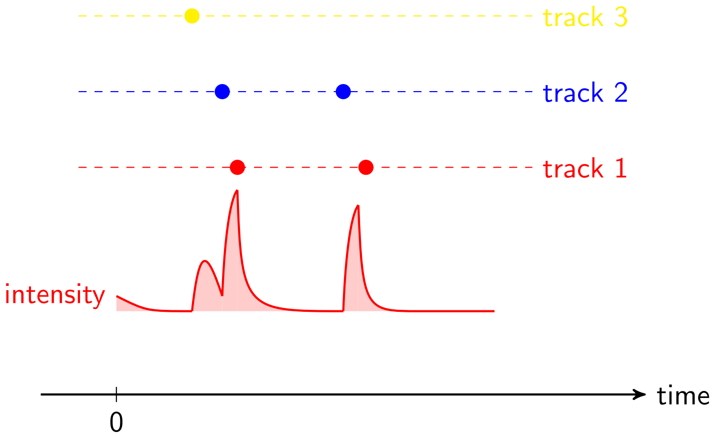
Linear filters



Linear filters



Linear filters



Local independence graphs for Hawkes processes

Simple observation: The graph defined by

$$(j, k) \in E \iff g^{jk} \neq 0$$

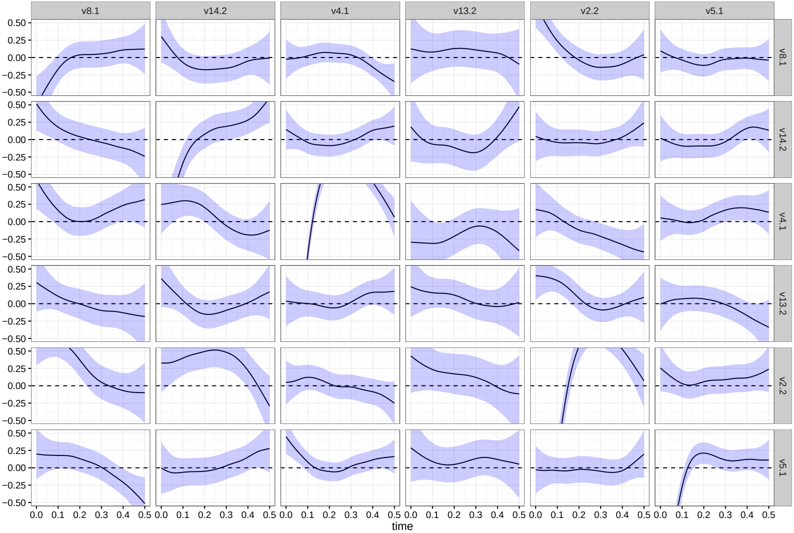
is a local independence graph.

R code – using package ppstat

A penalized *B*-spline basis expansion is available via the `ppSmooth` function.

```
pps <- ppSmooth(c(v8.1, v14.2, v4.1, v13.2, v2.2, v5.1) ~
                s(v8.1) + s(v14.2) +  s(v4.1) +
                s(v13.2) + s(v2.2) + s(v5.1),
                data = spikeData,
                family = Hawkes("log"),
                support = 0.5,
                lambda = 10
)
```

Results – basis approximation with *B*-splines



Testing

We can test

$$H_0 : g^{jk} = 0$$

using that test statistic

$$(\hat{\beta}_{\lambda}^{jk})^T \left(\hat{\Sigma}_{\lambda}^{jk}\right)^{-1} \hat{\beta}_{\lambda}^{jk}$$

with an approximating χ^2-distribution.

Testing

We can test

$$H_0 : g^{jk} = 0$$

using that test statistic

$$(\hat{\beta}_\lambda^{jk})^T \left(\hat{\Sigma}_\lambda^{jk}\right)^{-1} \hat{\beta}_\lambda^{jk}$$

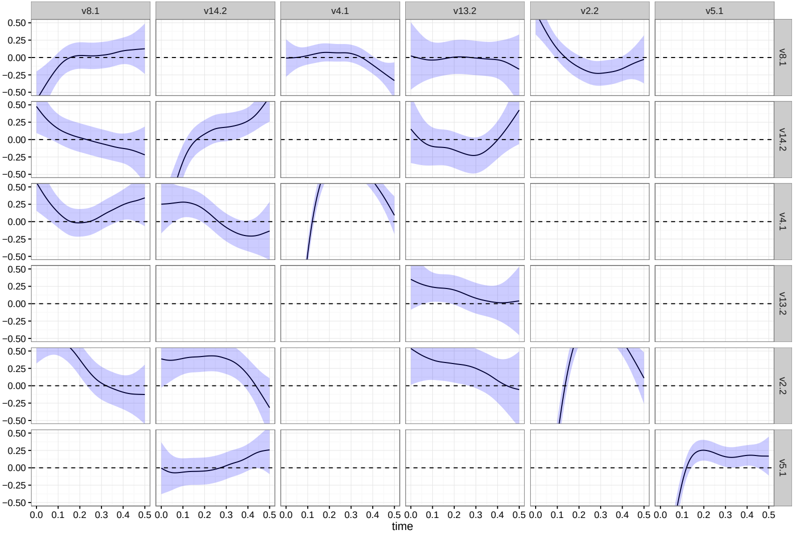
with an approximating χ^2-distribution.

Here, λ is the smoothing parameter and $\hat{\Sigma}_\lambda^{jk}$ is the sandwich estimator for the covariance matrix for the parameters belonging to the jkth filter function.

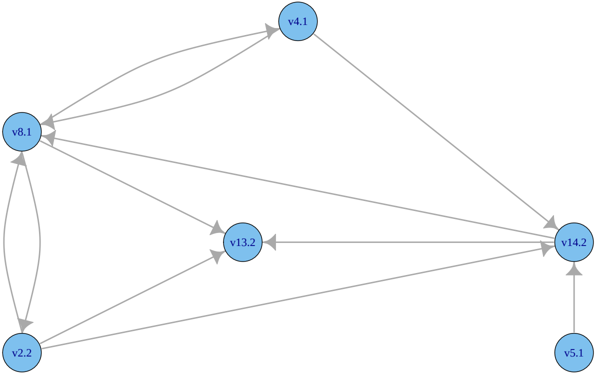
Test results

```
condTest(pps)
              v8.1        v14.2          v4.1         v13.2          v2.2          v5.1
v8.1  0.51 (38, 39)  0.2 (47, 40)    0 (91, 53) 0.03 (46, 30) 0.03 (65, 45) 0.22 (59, 52)
v14.2    0 (88, 40)   0 (110, 42) 0.51 (53, 54)    0 (84, 28) 0.79 (38, 46) 0.36 (55, 52)
v4.1     0 (92, 53)    0 (90, 53)   0 (550, 63) 0.06 (51, 37) 0.19 (63, 54) 0.11 (72, 58)
v13.2 0.25 (35, 30) 0.07 (45, 32) 0.87 (29, 39) 0.36 (29, 27) 0.44 (40, 39) 0.46 (36, 36)
v2.2     0 (77, 46)    0 (84, 48) 0.09 (69, 54) 0.01 (61, 38)   0 (240, 52) 0.16 (63, 53)
v5.1   0.4 (54, 52)   0 (100, 52) 0.15 (70, 59) 0.99 (18, 35) 0.07 (68, 52)   0 (180, 55)
```

Refit after selection



Local independence graph



Does the graph have a causal interpretation?

- Yes, if the observable event tracks have a distribution with a causally valid graph.

Does the graph have a causal interpretation?

- Yes, if the observable event tracks have a distribution with a causally valid graph.
- No, not in general if the causally valid graph contains unobserved processes.

Does the graph have a causal interpretation?

- Yes, if the observable event tracks have a distribution with a causally valid graph.
- No, not in general if the causally valid graph contains unobserved processes.

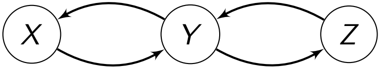
Earlier work

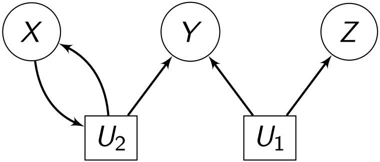
NRH, P. Reynaud-Bouret and V. Rivoirard. Lasso and probabilistic inequalities for multivariate point processes, Bernoulli 21(1), 2015.

NRH. Nonparametric likelihood based estimation of linear filters for point processes. Statistics and Computing 25(3), 2015.

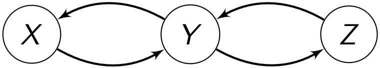
has focused on predictive models.

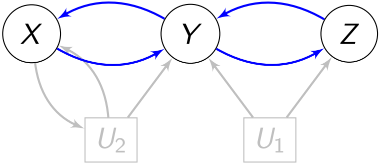
Markov properties for a faithful graph





Markov properties for a faithful graph

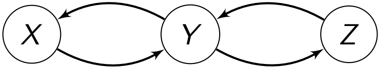


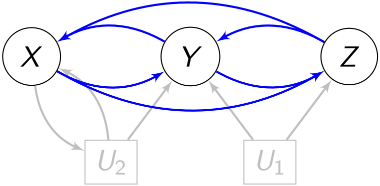


$$X \to Y \mid Z$$
$$Y \to X \mid Z$$
$$Z \to Y \mid X$$
$$Y \to Z \mid X$$

$$X \to Y \mid Z$$
$$Y \to X \mid Z$$
$$Z \to Y \mid X$$
$$Y \to Z \mid X$$

Markov properties for a faithful graph

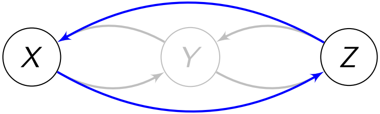


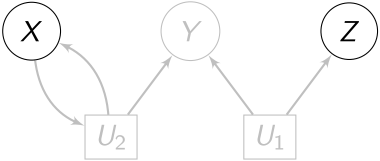


$$X \rightarrow Y \mid Z$$
$$Y \rightarrow X \mid Z$$
$$Z \rightarrow Y \mid X$$
$$Y \rightarrow Z \mid X$$
$$X \not\rightarrow Z \mid Y$$
$$Z \not\rightarrow Y \mid Y$$

$$X \rightarrow Y \mid Z$$
$$Y \rightarrow X \mid Z$$
$$Z \rightarrow Y \mid X$$
$$Y \rightarrow Z \mid X$$
$$X \rightarrow Z \mid Y$$
$$Z \rightarrow X \mid Y$$

Markov properties for a faithful graph





$$X \to Y \mid Z$$
$$Y \to X \mid Z$$
$$Z \to Y \mid X$$
$$Y \to Z \mid X$$
$$X \not\to Z \mid Y$$
$$Z \not\to Y \mid Y$$
$$X \to Z \mid \emptyset$$
$$Z \to X \mid \emptyset$$

$$X \to Y \mid Z$$
$$Y \to X \mid Z$$
$$Z \to Y \mid X$$
$$Y \to Z \mid X$$
$$X \to Z \mid Y$$
$$Z \to X \mid Y$$
$$X \not\to Z \mid \emptyset$$
$$Z \not\to X \mid \emptyset$$

The Empirical Causal Analysis (ECA) algorithm[1]

Initiate with the complete directed graph on V. Let $\mathrm{pa}(j)$ denote the parents of vertex j.

- $k = 0$: Test $j \not\to k \mid \emptyset$. Remove the edge (j, k) if the test is not rejected.

[1] An empirical version of the *Causal Analysis algorithm* in C. Meek, Toward Learning Graphical and Causal Process Models, UAI, 2014,

The Empirical Causal Analysis (ECA) algorithm[1]

Initiate with the complete directed graph on V. Let $\mathrm{pa}(j)$ denote the parents of vertex j.

- $k = 0$: Test $j \not\rightarrow k \mid \emptyset$. Remove the edge (j, k) if the test is not rejected.
- $k = 1$: Test $j \not\rightarrow k \mid m$ for $(j, k), (m, k) \in E$ $(j, m \in \mathrm{pa}(j))$. Remove the edge (j, k) if the test is not rejected.

[1] An empirical version of the *Causal Analysis algorithm* in C. Meek, Toward Learning Graphical and Causal Process Models, UAI, 2014,

The Empirical Causal Analysis (ECA) algorithm[1]

Initiate with the complete directed graph on V. Let $\mathrm{pa}(j)$ denote the parents of vertex j.

- $k = 0$: Test $j \not\to k \mid \emptyset$. Remove the edge (j, k) if the test is not rejected.
- $k = 1$: Test $j \not\to k \mid m$ for $(j, k), (m, k) \in E$ $(j, m \in \mathrm{pa}(j))$. Remove the edge (j, k) if the test is not rejected.
- ...

[1] An empirical version of the *Causal Analysis algorithm* in C. Meek, Toward Learning Graphical and Causal Process Models, UAI, 2014,

The Empirical Causal Analysis (ECA) algorithm[1]

Initiate with the complete directed graph on V. Let $\mathrm{pa}(j)$ denote the parents of vertex j.

- $k = 0$: Test $j \not\to k \mid \emptyset$. Remove the edge (j, k) if the test is not rejected.
- $k = 1$: Test $j \not\to k \mid m$ for $(j, k), (m, k) \in E$ $(j, m \in \mathrm{pa}(j))$. Remove the edge (j, k) if the test is not rejected.
- ...
- k: Test $j \not\to k \mid C$ for $\{j\} \cup C \subseteq \mathrm{pa}(j)$ with $|C| = k$. Remove the edge (j, k) if the test is not rejected.

[1]An empirical version of the *Causal Analysis algorithm* in C. Meek, Toward Learning Graphical and Causal Process Models, UAI, 2014,

The Empirical Causal Analysis (ECA) algorithm[1]

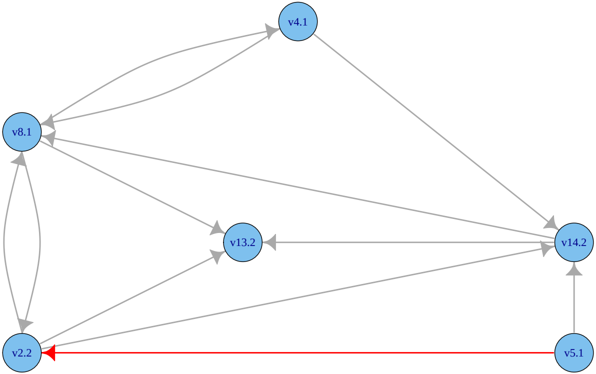
Initiate with the complete directed graph on V. Let $\mathrm{pa}(j)$ denote the parents of vertex j.

- $k = 0$: Test $j \not\to k \mid \emptyset$. Remove the edge (j, k) if the test is not rejected.
- $k = 1$: Test $j \not\to k \mid m$ for $(j, k), (m, k) \in E$ $(j, m \in \mathrm{pa}(j))$. Remove the edge (j, k) if the test is not rejected.
- ...
- k: Test $j \not\to k \mid C$ for $\{j\} \cup C \subseteq \mathrm{pa}(j)$ with $|C| = k$. Remove the edge (j, k) if the test is not rejected.

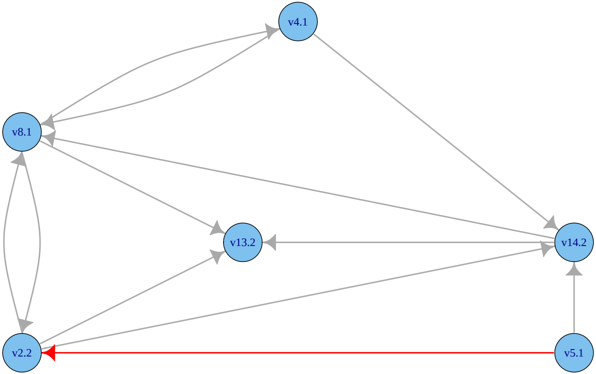
The algorithm stops when $\mathrm{pa}(j) < k + 1$ for all j.

[1] An empirical version of the *Causal Analysis algorithm* in C. Meek, Toward Learning Graphical and Causal Process Models, UAI, 2014,

Local independence graph

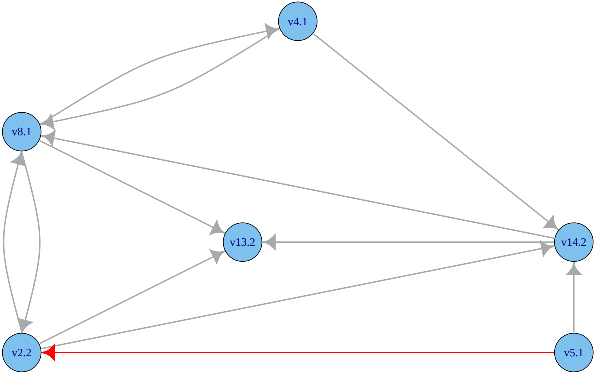


Local independence graph



The ECA algorithm is a PC-type algorithm.

Local independence graph



The ECA algorithm is a PC-type algorithm.
No MAGs, PAGs or (R)FCI-type algorithm yet ☹ ☹ ☹.

Problems

Proposed tests rely on joint as well as marginal Hawkes process models.

The class of Hawkes processes is not closed under marginalization.

Problems

Proposed tests rely on joint as well as marginal Hawkes process models.

The class of Hawkes processes is not closed under marginalization.

- Marginal independence can be tested using Ripley's K-function under stationarity.

Problems

Proposed tests rely on joint as well as marginal Hawkes process models.

The class of Hawkes processes is not closed under marginalization.

- Marginal independence can be tested using Ripley's K-function under stationarity.
- Nonparametric tests for conditional local independence of event tracks are not abundant

Problems

Proposed tests rely on joint as well as marginal Hawkes process models.

The class of Hawkes processes is not closed under marginalization.

- Marginal independence can be tested using Ripley's K-function under stationarity.
- Nonparametric tests for conditional local independence of event tracks are not abundant (that's true for conditional independence tests as well without e.g. a Gaussian assumption).

Problems

Proposed tests rely on joint as well as marginal Hawkes process models.

The class of Hawkes processes is not closed under marginalization.

- Marginal independence can be tested using Ripley's K-function under stationarity.
- Nonparametric tests for conditional local independence of event tracks are not abundant (that's true for conditional independence tests as well without e.g. a Gaussian assumption).
- Causal effect estimates (intervention effects) have not been considered.

Extensions

Conditional local independence was not defined in terms of factorization.

It was defined in terms of functional dependence of compensators for counting processes.

Extensions

Conditional local independence was not defined in terms of factorization.

It was defined in terms of functional dependence of compensators for counting processes.

A formal generalization to semimartingales is possible in terms of the characteristic triplet. We have resolved some issues for SDEs[1], such as interventions being well-defined.

[1]A. Sokol and NRH. *Causal interpretation of stochastic differential equations*. Electron. J. Probab. 19, 2014.

Extensions

Conditional local independence was not defined in terms of factorization.

It was defined in terms of functional dependence of compensators for counting processes.

A formal generalization to semimartingales is possible in terms of the characteristic triplet. We have resolved some issues for SDEs[1], such as interventions being well-defined.

Open problem: Does generalizations beyond the counting / event processes enjoy the global Markov property?

[1]A. Sokol and NRH. *Causal interpretation of stochastic differential equations*. Electron. J. Probab. 19, 2014.

Acknowledgement

Thanks to my collaborators:

Patricia Reynaud-Bouret
Vincent Rivoirard
Alexander Sokol

Acknowledgement

Thanks to my collaborators:

Patricia Reynaud-Bouret
Vincent Rivoirard
Alexander Sokol

Research sponsered by VILLUM FONDEN.

Causal inference with incomplete data

NRH, Steffen Lauritzen and Jonas Peters

`http://www.math.ku.dk/~richard/causality/`

Acknowledgement

Thanks to my collaborators:

Patricia Reynaud-Bouret
Vincent Rivoirard
Alexander Sokol

Research sponsered by VILLUM FONDEN.

Causal inference with incomplete data

NRH, Steffen Lauritzen and Jonas Peters

http://www.math.ku.dk/~richard/causality/

Thanks