Hawkes Processes Tutorial

Adapted from Gomez-Rodriguez [3]

Tiago Santos,
Machine Learning Meetup

Graz, 26. April 2018

Motivation and Applications

See page 6 of Lecture 1 in [3]

See page 33 of Lecture 2 in [3]

Generalized Problem Formulation

Suppose:

1. Discrete **event stream** of timestamps
   - Irrespective of application scenario [4, 2, 1]
2. Non-trivial temporal dynamics and dependencies:
   - Dependence of own event history
   - Dependence of other event histories

Generalized Problem Formulation

Suppose:

1. Discrete **event stream** of timestamps
   - Irrespective of application scenario [4, 2, 1]
2. Non-trivial temporal dynamics and dependencies:
   - Dependence of own event history
   - Dependence of other event histories

When facing such a problem, consider Hawkes processes!

Univariate Point Processes and Hawkes Processes

See page 12 of Lecture 1 in [3]

Intensity Function

Since it is cumbersome to model event counts over time, we model event **intensity** over time:

$$\lambda^*(t)dt = \mathbb{E}[dN(t)|\mathcal{H}(t)]$$

Intensity Function

Since it is cumbersome to model event counts over time, we model event **intensity** over time:

$$\lambda^*(t)dt = \mathbb{E}[dN(t)|\mathcal{H}(t)]$$

$\lambda^*(t)$ is expected value of (infinitesimal) change in event count over time, given event history

$\Leftrightarrow \lambda^*(t)$ is **number of events per time interval**, and this changes over time!

See page 21 of Lecture 1 in [3]

See page 24 of Lecture 1 in [3]

See page 28 of Lecture 1 in [3]

See page 29 of Lecture 1 in [3]

Hawkes (Self-Exciting) Process

Typical choices for kernel function $\kappa_\omega(t)$ include power law and exponential kernel:

$$\kappa_\omega(t) = e^{-\beta t}$$

Hence we get:

$$\lambda^*(t) = \mu + \sum_{t_i<t} \alpha e^{-\beta(t-t_i)}$$

Hawkes (Self-Exciting) Process

Typical choices for kernel function $\kappa_\omega(t)$ include power law and exponential kernel:

$$\kappa_\omega(t) = e^{-\beta t}$$

Hence we get:

$$\lambda^*(t) = \mu + \sum_{t_i < t} \alpha e^{-\beta(t - t_i)}$$

What can we do with these models?

- Fit models to real data by maximizing log-likelihood
- Sample from fitted process via inverse sampling

See page 22 of Lecture 1 in [3]

See page 30 of Lecture 1 in [3]

Multivariate Hawkes Processes

See page 35 of Lecture 1 in [3]

Multivariate Hawkes Process

M-variate Hawkes Process definition:

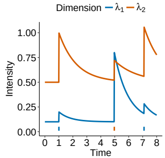
$$\lambda^{*m}(t) = \mu_m + \sum_{n=1}^{M} \sum_{t_i^n < t} \alpha_{mn} e^{-\beta_{mn}(t - t_i^n)}$$

Multivariate Hawkes Process

M-variate Hawkes Process definition:

$$\lambda^{*m}(t) = \mu_m + \sum_{n=1}^{M} \sum_{t_i^n < t} \alpha_{mn} e^{-\beta_{mn}(t - t_i^n)}$$

2-variate Hawkes Process sample for $T = 8$:

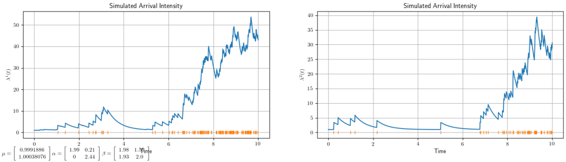


Parameter values: $\mu = \begin{pmatrix} 0.1 \\ 0.5 \end{pmatrix}$, $\alpha = \begin{pmatrix} 0.1 & 0.7 \\ 0.5 & 0.2 \end{pmatrix}$, $\beta = \begin{pmatrix} 1.2 & 1.0 \\ 0.8 & 0.6 \end{pmatrix}$

A few words of caution!

Pitfalls & Counter-Measures

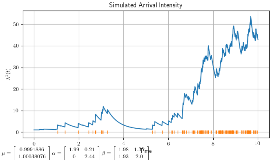
- Assure stationary of Multivariate Hawkes, otherwise:

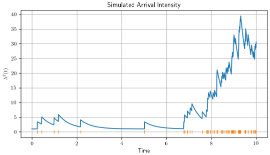


- Stationarity test: Spectral radius $\rho < 1$
- Fitting β: Hyperparameter optimization
- Fit quality: Measure with Q-Q plot

Pitfalls & Counter-Measures

- Assure stationary of Multivariate Hawkes, otherwise:





- Stationarity test: Spectral radius $\rho < 1$
- Fitting β: Hyperparameter optimization
- Fit quality: Measure with Q-Q plot
- Alternative approaches:
  - Dynamical systems (e.g. branching processes)
  - Information-theory (e.g. transfer entropy)

Resources

- Python package: Tick

  https://github.com/X-DataInitiative/tick
- C++ package: PtPack

  https://github.com/dunan/MultiVariatePointProcess
- Hawkes network inference: Pyhawkes

  https://github.com/slinderman/pyhawkes
- Models from papers:
  - Distilling Information Reliability and Source Trustworthiness from Digital Traces

    http://btabibian.com/projects/reliability/
  - Modeling the Dynamics of Online Learning Activity https://github.com/Networks-Learning/hdhp.py

References

[1] E. Bacry, I. Mastromatteo, and J.-F. Muzy.
Hawkes processes in finance.
Market Microstructure and Liquidity, 1(01):1550005, 2015.

[2] D. J. Daley and D. Vere-Jones.
An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods.
Springer Science & Business Media, 2003.

[3] M. Gomez-Rodriguez.
Machine learning for dynamic social network analysis seminar.
`http://learning.mpi-sws.org/uc3m-seminar/`, 2017.
Accessed: 2018-02-10.

[4] Y. Ogata.
Statistical models for earthquake occurrences and residual analysis for point processes.
Journal of the American Statistical association, 83(401):9–27, 1988.