nphawkes R Package

Modeling spatio-temporal point processes with nphawkes package

Peter Boyd
Oregon State University
useR! 2021

July 2021

Hawkes process

A **self-exciting point process** (Hawkes process [**?**]) is a process in which the occurrence of a point causes the temporary elevation in the occurrence of future points nearby in time or space and time.

- Background (parent) events happen “randomly"
- Triggered (child) events are spawned by background events.

Examples:

- Seismology
- Social networks such as email chains or retweets on Twitter
- Spread of epidemic diseases like COVID-19
- Mass shootings and gang violence

Modeling

Self-exciting processes may be fit nonparametrically using the **Model independent stochastic declustering (MISD)** algorithm [**?**].

- estimates the background rate, $\mu(s)$, or the rate at which events stochastically occur in time.
- estimates the ability of an event to trigger new events in space (h) and time (g), based on the observed magnitude (k).

Goal: estimate the **conditional intensity** function, $\lambda(s, t)$, or the expected rate at which points occur within a very small space-time window given the history of the process, $\mathcal{H}_t$ [**?**, **?**].

$$\lambda(s, t|\mathcal{H}_t) = \mu(s) + \sum_{i:t_i<t} g(t - t_i)h(s - s_i)k(m_i). \tag{1}$$

nphawkes package

The MISD algorithm may be computationally expensive, especially with large data. I have built the `nphawkes` package to easily and quickly fit these nonparametric Hawkes models, along with several related analysis tools.

- Functions within algorithm written in C++ using `Rcpp` package.
- Requires minimal inputs (time) and may flexibly include space and magnitude-type covariate.

Package may be installed as below:

```
install.packages("devtools")
devtools::install_github("boydpe/nphawkes")
library(nphawkes)
```

MISD implementation

We'll apply the `nphawkes::misd()` function to an earthquake catalog containing the Hector Mine earthquake of October 1999 in Southern California and roughly 540 subsequent events.

```
eq = read.csv("hm.csv") # Hector Mine earthquake data
out = misd(dates = eq$t, lat = eq$lat, lon = eq$lon,
           marks = eq$m, time_breaks = NULL,
           space_breaks = NULL, mark_breaks = NULL,
           just_times = TRUE, nonstat_br = TRUE)
```

Results:

- r round(100*out$perc_diag,2)% of data treated as background events.
- Hector Mine earthquake found to trigger r round(colSums(out$p$0)[2] − $diag$(out$p$0)[2]) aftershocks.
- Background events found to occur at a rate of r round(mean(out$br),2) mainshocks per day.
- Algorithm takes roughly 11 seconds to complete.

What's next?

Once a model has been fit, plots can be made to visualize:

- `heat map of nonstationary background rate`
- `histogram estimators of triggering function`
- `conditional intensity over time`

`nphawkes` allows for easy implementation of nonparametric Hawkes models.
Try it out with your own self-exciting data!

References I

#

#