HawkesProcesses.jl

Dean Markwick

May 22, 2020

Abstract

HawkesProcesses.jl is a Julia package for using Hawkes processes to model event time data. This package provides functions to plot the intensity, calculate that likelihood and fit the parameters using a latent variable Bayesian algorithm from [Ros16] and extending it to the generic case.

1 Introduction

A Hawkes process is a point process where the intensity function is self-exciting. The intensity at time t given the history of the process, H_t can be written as

$$\lambda(t \mid H_t) = \mu(t) + \kappa \sum_{t_i < t} g(t - t_i), \tag{1}$$

where μ is the background rate, κ is a constant parameter and $g(t)$ is some positive function.

2 Likelihood

For a collection of event times $t_i, \dots, t_n$ in the time interval $[0, T]$ the log likelihood can be written as

$$\mathcal{L} = \sum_{i=1}^{n} \log \lambda(t_i \mid H_{t_i}) - \int_0^T \lambda(t \mid H_t) \mathrm{d}t,$$

which is also the general log likelihood for a point process from [DVJ03]. By inserting the Hawkes process intensity function into this equation we can simplify it down.

The integral of the intensity can be expanded as

$$\begin{aligned} \int_0^T \lambda(t \mid H_t) \mathrm{d}t &= \int_0^T \mu(t) \mathrm{d}t + \kappa \int_0^T \sum_{t_j < t} g(t - t_j) \\ &= M(T) + \kappa \sum_{i=1}^{n} \left(G(T - t_i) - G(0) \right), \end{aligned}$$

where $M(T) = \int_0^T \mu(t)\mathrm{d}t$ and $G(t) = \int g(t)\mathrm{d}t$. As we are forcing g to be a probability density defined on 0 to T, G is just the distribution function and thus $G(0) = 0$. Thus the final likelihood can be written as

$$\mathcal{L} = \sum_{i=1}^{n} \log \lambda(t_i \mid H_{t_i}) - M(T) - \kappa \sum_{i=1}^{n} G(T - t_i). \tag{2}$$

3 Latent Variable B

Using the superposition principle of Poisson processes the Hawkes process can be viewed as many different inhomogeneous Poisson processes layered up each other. After each event t_i a new Poisson process with intensity is spawned with intensity $\kappa g(t - t_i)$ between $[t_i, T]$. Under this formulation we can estimate whether an event was spawned from the background rate ($\mu(t)$) or whether it was spawned from another event during the excitation. Therefore we introduce a variable $\mathbf{B} = B_1, \ldots, B_n$ for each event $t_i, \ldots, t_n$ that indicates who the parent is. As this can never be observed in practise it is a latent variable and instead must be estimated. The estimation probabilities can be written as

$$\begin{aligned} \Pr(B_i = 0 \mid t_i, H_t, \theta) &= \frac{\mu(t)}{\lambda(t_i \mid H_t)}, \\ \Pr(B_i = j \mid t_i, H_t, \theta) &= \frac{\kappa g(t_i - t_j)}{\lambda(t_i \mid H_t)}, \quad j = \{1, 2, \ldots, i-1\}. \end{aligned} \tag{3}$$

where $B_i = 0$ indicates the event t_i came from the background rate and $B_i = j$ indicates that t_j is the parent to t_i. These latent variables allow us to split the likelihood and build a more efficient MCMC sampler.

4 Latent Variable Sampling

From the generic likelihood in (2) we introduce the branching structure from (3) to separate the likelihood

$$p(t_i, \ldots t_n \mid \theta, \mathbf{B}) = e^{-\mu T} \mu^{|S_0|} \prod_{j=1}^{n} \left(e^{-\kappa G(T - t_j)} \kappa^{|S_j|} \prod_{t_i \in S_j} g(t_i - t_j) \right), \tag{4}$$

therefore given a sample of the latent variable $\mathbf{B}$ we can split out the likelihood into its three components and sample new parameters from each component accordingly.

4.1 Constant Background

For all events in S_0 the likelihood is simple to write

$$p(\mid B) = e^{-\mu T} \mu^{|S_0|}$$

and thus the posterior distribution is

$$p(\mu \mid B) = \text{Gamma}\left(\alpha_\mu + \mid S_0 \mid, \beta_\mu + 1\right). \tag{5}$$

4.2 Constant Kappa

From the latent likelihood in (4) we can pull out the middle term that concerns κ only

$$\begin{aligned}
p(\mid \theta, B, h) &= \prod_{j=1}^{n} e^{-\kappa g(T-t_j)} \kappa^{|S_j|} \\
&= \kappa^{\sum j=1^n |S_j|} e^{-\kappa \sum_{j=1}^{n} G(T-j_j)}, \\
\tilde{G} &\equiv \sum_{j=1}^{n} G(T-j_j), \\
p(t_1, \ldots, t_n \mid \theta, \mathbf{B}, h) &= \kappa^{\sum_{j=1}^{n} |S_j|} e^{-\kappa \tilde{G}}, \\
&= \text{Gamma}(\sum_{j=1}^{n} \mid S_j \mid +1, \tilde{G}),
\end{aligned}$$

we define our prior on κ as a gamma distribution also which means we can exploit the conjugacy to give the posterior distribution of κ as

$$p(\kappa \mid t_1, \ldots t_n, \mathbf{B}, g) = \text{Gamma}\left(\sum_{j=1}^{n} \mid S_j \mid + \alpha_\kappa, \tilde{G} + \beta_\kappa\right), \tag{6}$$

where $\alpha_\kappa, \beta_\kappa$ are the prior parameters for κ.

In practise $\tilde{G} \approx n$ as $G(T - t_j) = 0$ for the events that are far away from the boundary.

4.3 Exponential Kernel

We have to assume a form of g to proceed with inference. An exponential kernel is most commonly used

$$g(t \mid \beta) = \beta \exp(-\beta t),$$

we also define $\tau_i = t_i - t_j$ as the time of a child relative to its parent for all events with a parent that isn't the background. From both of these we can define the likelihood of the observations as

$$\begin{aligned}
p(\mid \theta) &= \prod_{j=1}^{n} \prod_{t_i \in S_j} g(t_i - t_j) \\
&= \beta^{N_{\text{child}}} \exp\left(-\beta \sum_{i=1}^{N_{\text{child}}} \tau_i\right),
\end{aligned}$$

which again is conjugate to the gamma distribution and therefore

$$p(\beta \mid \tau_1, \ldots, \tau_{N_{\text{child}}}, \mathbf{B}) = \text{Gamma}\left(N_{\text{child}} + \alpha_\beta, \sum_i^n \tau + \beta_\beta\right), \qquad (7)$$

where $\alpha_\beta, \beta_\beta$ are the prior parameters for the kernel.

4.4 The Algorithm

Using these posterior distributions the parameters of the Hawkes process can be updated using the following algorithm.

Algorithm 1: Sampling a Hawkes process using the latent variables

Result: *its* samples of μ, κ, β and $\mathbf{B}$
while $i \leq its$ **do**
  Sample new latent variables $\mathbf{B}$ using (3). ;
  Sample new μ using (5) ;
  Sample new κ using (6) ;
  Sample new β using (7) ;

References

[DVJ03] D.J Daley and D Vere-Jones. *An Introduction to the Theory of Point Processes*. en. Probability and its Applications. New York: Springer-Verlag, 2003. ISBN: 978-0-387-95541-4. URL: `http://link.springer.com/10.1007/b97277` (visited on 07/11/2016).

[Ros16] Gordon J Ross. *Bayesian Estimation of the ETAS Model for Earthquake Occurrences*. 2016.