First look at temporal point processes

Cameron Roach

02/07/2019

Part 1: Temporal point processes

- Advanced concepts: marked TPP, dynamical systems with jumps
- Used a lot in recent machine learning models and control algorithms.

Useful slides here: http://courses.mpi-sws.org/hcml-ws18/. PDFs from that seminar contained in this directory.

Intensity function

Definitions

- $N(t)$ number of events that have occurred up to and including time t.
- $\mathcal{H}(t)$ is the history of when and how many events occurred up to time t.
- $f(t)$ the density function of an event occurring.
- $f^*(t) := f(t|\mathcal{H}(t))$ is probability of event occurring at time t conditional on the history.
- $S^*(t)$ is probability of event not occuring before time t based on $f^*(t)$ distribution.

The intensity function is a rate (## events per unit of time) given by the probability of an event occurring in $[t, t + dt]$ but not before t. It is given by

$$\begin{aligned}\lambda^*(t)dt &= \frac{f^*(t)dt}{S^*(t)} \\ &= \mathbb{E}\left[dN(t)|\mathcal{H}(t)\right]\end{aligned}$$

Intensity functions are useful for parameterizing histories. Easier to do intensity parameterization rather than density parameterization as intensity functions only need to be non-negative and it is easy to combine timelines.

Dynamic processes

Several types of dynamic processes that can be described by intensity functions

- Poisson process (homogeneous and inhomogeneous)
- Terminating (survival) process
- Self-exciting (Hawkes) process

Fitting

All are fit using maximum likelihood.

Sampling

Different processes (Poisson, Hawkes, survival) can be sampled using acceptance-rejection sampling and inverse sampling.

Superposition

Works as you'd expect. Different intensity functions can be superposed to create new intensity functions. These can be used to model how an individual may be affected by neighbours. For example, consider individuals A and B which have a terminating and self-exciting process, respectively. If A is affected by B then the final intensity function for A will be a combination of the terminating and self-exciting intensity functions.

Marks

Marks (the state of a temporal process) occur at various times. Individual shifts between different states?

Simplest case

- $x^*(t_i) \sim p(x)$
- Marks independent of temporal dynamics
- i.i.d.

Stochastic differential equations

Marks given by SDEs with jumps.

- $dx(t) = f(x(t)), t)dt + h(x(t), t)dN(t)$
- Marks dependent on temporal dynamics
- Defined for all values of t.

Part 2: Models & Inference

- Infection cascades

Complex temporal dynamics

- Poisson, self-exciting processes are examples of simple temporal dynamics and have simple intensity functions.
- Recent works make use of **RNNs** to capture more complex dynamics. See [Du et al., 2016; Dai et al., 2016; Mei & Eisner, 2017; Jing & Smola, 2017; Trivedi et al., 2017; Xiao et al., 2017a; 2018].
- Neural Hawkes process [Mei & Eisner, NIPS 2017]

Causal reasoning

- Causal inference using event sequences [Xu et al., 2016; Achab et al., 2017; Kuśmierczyk & Gomez-Rodriguez, 2018].
- "X causes Y in the sense of Granger causality if forecasting future values of Y is more successful while taking X past values into account" [Granger, 1969]

Remember, $k_{u,v}(t)$ is the effect of v's past events on u as of time t. Integrate k over t to obtain a measure of average total # of events of node u whose direct ancestor is an event by node v.

$$g_{u,v} = \int_{\mathbb{R}^+} k_{u,v}(\tau) d\tau \geq 0 \; \forall \; u, v \in \mathcal{U}.$$

Take $G = [g_{u,v}]$ as a measure of direct causal relationship between nodes.

Part 3: Reinforcement Learning and Control

Stochastic Optimal Control

- Red queen algorithm [Zarezade et al., 2017 & 2018]. See http://learning.mpi-sws.org/redqueen/