SEQUENTIAL INFERENCE METHODS FOR NON-HOMGENEOUS POISSON PROCESSES WITH STATE-SPACE PRIOR

Chenhao Li, Simon J. Godsill

Signal Processing and Communications Laboratory, Department of Engineering, University of Cambridge

UNIVERSITY OF CAMBRIDGE

Abstract

The non-homogeneous Poisson process (NHPP) is a point process with time-varying intensity across its domain, the use of which arises in numerous domains in signal processing, machine learning and many other fields. However, its applications are largely limited by the intractable likelihood and the high computational cost of existing inference schemes. We present an online inference framework that utilises generative Poisson data and sequential Markov Chain Monte Carlo (SMCMC) algorithm, which achieves improved performance in both synthetic and real datasets.

Non-homogeneous Poisson process

A non-homogeneous Poisson process over the domain $S = \mathbb{R}^D$ of arbitrary dimension D possesses the following attributes:

- A varying intensity function $\boldsymbol{\lambda(s)}$
- Counting measure $N(\mathcal{T})$ evaluated over $\mathcal{T} \subset S$ follows a Poisson distribution with parameter $\lambda_{\mathcal{T}} = \int_{\mathcal{T}} \lambda(s)\,\mathrm{d}s$
- The counts of events in any disjoint subsets $\mathcal{T}_i \subset S$ are independent random variables

which result in the following likelihood given a set of K events denoted as $\{s_k\}_{k=1}^K$ in the region $\mathcal{T}$:

$$p(\{s_k\}_{k=1}^K \mid \lambda(s), \mathcal{T}) = \exp\Big\{-\int_{\mathcal{T}} \mathrm{d}s\, \lambda(s)\Big\} \prod_{k=1}^{K} \lambda(s_k)$$

GOAL: obtain a Bayesian estimate of the intensity function $\lambda(s)$. But wait, the likelihood is requiring the knowledge of $\lambda(s)$ across $\mathcal{T}$ while $\lambda(s)$ is exactly what we want to know –INTRACTABILITY!

Fortunately, we can still generate this process:

1. Generate homogeneous Poisson points $\{s_n\}_{n=1}^N$ with an upper-bound intensity λ^*.
2. Evaluate the local intensities $\{\lambda(s_n)\}_{n=1}^N$ at these homogeneous points.
3. Perform thinning operation to obtain the unbiased simulation of the NHPP.

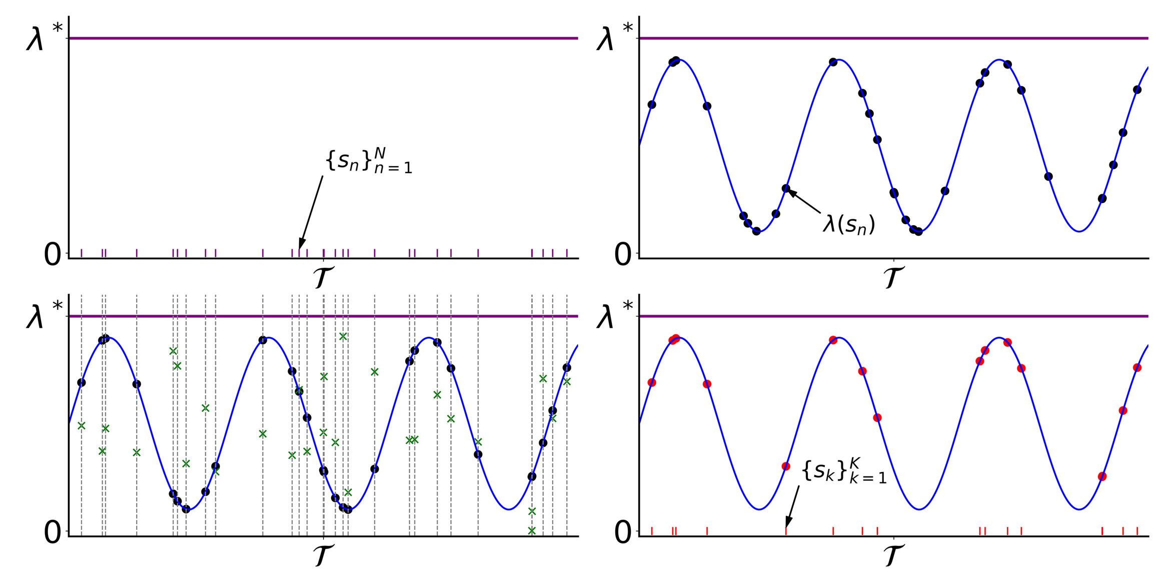

Figure: NHPP Generation Scheme

		S-LD	KDE	SGCP
$\lambda_1(s)$	mse	**0.0257**	0.129	0.0704
	$\mathcal{L}(p)$	**1.825**	–	-9.440
	Time (s)	15.86	**0.01**	60.23
$\lambda_2(s)$	mse	**0.6531**	0.8599	1.5257
	$\mathcal{L}(p)$	**-248.1**	–	-326.6
	Time (s)	60.05	**0.05**	1326.28

Table: Numerical results for models. Bold is the best.

Model

Inspired by the generative procedure, the Sigmoid Gaussian Cox Process (SGCP) was introduced in [1]. The model places a Gaussian process transformed by a maximum-intensity-scaled sigmoid function over the intensity function to construct the generative joint probability using the concept of thinning. Define $i_n \in \{0,1\}$ as the indicator associated with each (homogeneous) Poisson points, taking value 0 for the events retained in thinning (input data points) and value 1 for 'latent' events rejected in thinning. We have

$$p(\{s_n\}_{n=1}^N, \mathbf{g}_{1:N}, \{i_n\}_{n=1}^N \mid \lambda^*, \mathcal{T}, \theta) = \underbrace{(\lambda^*)^N \exp\{-\lambda^*|\mathcal{T}|\}}_{(1)} \times \underbrace{\prod_{n=1}^{N} \sigma\{(-1)^{i_n} g(s_n)\}}_{(2)} \times \underbrace{p(\mathbf{g}_{1:N} \mid \{s_n\}_{n=1}^N, \theta)}_{(3)}$$

where $\{s_n\}_{n=1}^N$ is the time-ordered list of homogeneous Poisson points and $\mathbf{g}_{1:N}$ is the vector of corresponding $g(s_n)$ stochastic process values. In our model, we replace the heavy joint $\mathcal{GP}$ prior with a LTI state-space prior of the form $\mathrm{d}\mathbf{g}(t) = \mathbf{A}\mathbf{g}(t)dt + \mathbf{h}\,\mathrm{d}W(t)$, where its Markovian property avoids the $\mathcal{O}(N^3)$ complexity and aids the sequential inference formulations. Solved with stochastic integration, for two timestamps $Q > P$ we obtain:

$$\mathbf{g}(Q) = e^{\mathbf{A}(Q-P)}\Big[\mathbf{g}(P) + \int_0^{Q-P} e^{-\mathbf{A}\tau}\mathbf{h}\,\mathrm{d}W_\tau\Big] \qquad p\big(\mathbf{g}(Q) \mid \mathbf{g}(P)\big) = \mathcal{N}\Big(\mathbf{g}(Q) \mid \mu(Q,P), C(Q,P)\Big)$$

More specifically, a powerful Langevin dynamics model is adopted in which $\mathbf{g}_t = [g_{1,t}\ g_{2,t}]^T$ contains a stochastic trend (velocity) term $g_{2,t}$; $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & \theta \end{bmatrix}$; and $\mathbf{h} = [0\ \ \sigma]^T$.

Inference

We perform the posterior inference with the following salient features:

- Sequential Markov Chain Monte Carlo algorithm to make online inference and handle high-dimensional sample space
- Mixture of Metropolis-within-Gibbs proposal and joint proposal
- Perform batch inference with clustered points to recover the temporal correlation
- Gibbs update of the maximum intensity λ^* in each batch

Results

Numerical results obtained from synthetic datasets are shown in Table 1 with:

- $\lambda_1(s) = 2\exp\{-s/15\} + \exp\{-((s-25)/10)^2\}$ on the interval [0, 50] with 55 events.
- $\lambda_2(s)$ from a Langevin governed process with parameters $\theta = -0.5$, $\sigma = 0.5$ on interval [0, 100] with 156 events.

Furthermore, the proposed model is tested over a small set of Limit Order Book (LOB) data taken from the EUR-USD FOREX market on the 2nd of September 2015.

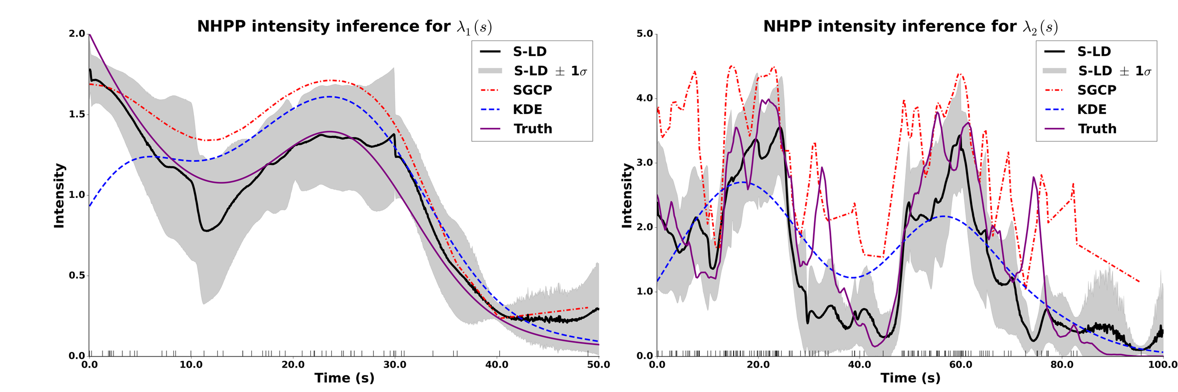

Figure: $\lambda_1(s)$ Figure: $\lambda_2(s)$

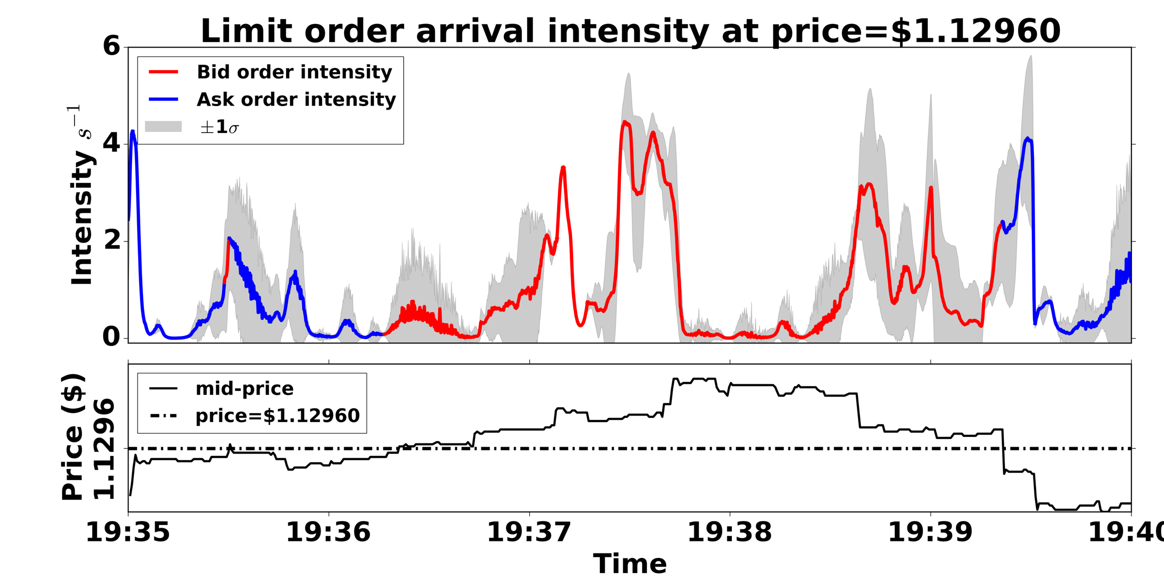

Figure: LOB Result

Conclusion

We have introduced a novel method for inferring the intensity function in a NHPP, which shows improved accuracy and efficiency compared to competing approaches. Moreover it is able to perform online inference in computationally heavy tasks which are challenging for the SGCP approach. The sequential batch scheme further restores the local location information that is crucial in the inference of NHPP and is usually ignored in standard sequential inference schemes.

Future Work

- Iteratively update the Langevin hyperparameters to adapt to the datasets
- Make use of the first derivative $g_{2,t}$ for practical applications

References

[1] R. Adams, I. Murray, and D. MacKay. Tractable nonparametric Bayesian inference in Poisson processes with Gaussian process intensities. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 9–16. ACM, 2009.