Monte Carlo EM Algorithm for Handling Missing Data in Hawkes Process

Jingtian Yu
Department of Statistics, Oregon State University

Sharmodeep Bhattacharyya
Department of Statistics, Oregon State University

Sarah Emerson
Department of Statistics, Oregon State University

Rob Trangucci
Department of Statistics, Oregon State University

Shirshendu Chatterjee
Department of Mathematics, The City University of New York

Abstract

The Hawkes process is a self-exciting point process used in a diverse range of fields, including finance, neuroscience, and social networks. The occurrence of missing data is quite common in point process data. The event intensity of the Hawkes process depends on past occurrences, making estimation with incomplete observations a challenge as unobserved events affect the intensity function. Much past work on estimation of Hawkes processes with missing data focuses on scenarios in which the missing events can be confined to a known interval, typically non-overlapping with the observation window or a subset of the observation window. The more general scenario in which missing events can occur at any point within the observation window remains unexplored. We address this more general missingness scenario by developing a likelihood-based estimation approach that incorporates imputation steps tailored to accommodate the missing mechanism, ensuring better handling of estimation bias in the incomplete data scenarios. We provide an extensive simulation study demonstrating the superior performance of our proposed estimation method. We also provide examples in real-life event time datasets.

Keywords: Hawkes process, incomplete data, missing mechanism, Monte Carlo EM algorithm