

Bayesian Inference for Diffusion Processes with Applications in Epidemiology

Christiane Dargatz¹
Gareth Roberts²

¹ Department of Statistics, Ludwig Maximilians University Munich, Germany

² Department of Statistics, University of Warwick, UK

Diffusion models provide a natural way to describe dynamic systems that change continuously in time. Statistical inference is often based on approximating the likelihood, which faces the difficulty of sparse data situations in many applications such as epidemics, where new infections might be reported only at weekly intervals. This problem can be overcome e.g. by combining the estimation of the model parameters with the imputation of augmented data using an MCMC scheme. Methods which treat the diffusion paths as discretely sampled (e.g. [1]) work sufficiently well for low-dimensional data, but are degenerate in the limit due to a strong dependency between the parameters and the quadratic variation of the diffusion paths [2]. Such techniques are thus not appropriate for very high-dimensional applications, where slow convergence becomes computationally costly.

We present a new approach which breaks down the disrupting dependency structures and hence allows a fast and efficient algorithm for parameter estimation and data imputation also for high-dimensional datasets. The method is illustrated in the framework of modelling the spatio-temporal spread of influenza in Germany.

References

- [1] Golightly A, Wilkinson DJ (2006) Bayesian sequential inference for nonlinear multivariate diffusions. *Statistics and Computing* 16:323-338
- [2] Roberts GO, Stramer O (2001) On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm. *Biometrika* 88:603-621