Bayesian Inversion for the Drift in Stochastic Differential Equations

Yvo Pokern

(joint work with Tjun Y. Hoh, Ioanna Manolopoulou)

For a stochastic differential equation (SDE) of Ito type of the form

$$\mathrm{d}V_t = \xi(V_t)\mathrm{d}t + \sigma(V_t)\mathrm{d}W_t, \qquad V_0 = v_0,$$

the problem considered here is to estimate the term $\xi(\cdot)$ known as the drift term as well as the term $\sigma(\cdot)$ known as the diffusivity based on observations of the stochastic process $(V_t)_{t\in[0,T]}$ at time points $t_i \in [0,T]$. In the statistical context, it has frequently been assumed that the unknown functions $\xi(\cdot)$ and $\sigma(\cdot)$ can be described by a finite-dimensional parameter $\theta \in \Theta$, usually of small dimension and with each component readily interpretable in the application domain. Interest then shifts to finding good values for this parameter, an approach known as parametric inference. Such a parameterization limits flexibility and if this is to be avoided, a nonparametric approach can be adopted instead, where the whole functions $\xi(\cdot), \sigma(\cdot)$, or at least their values at a set of points are estimated. The Bayesian approach is to assume *a priori*, i.e. before having taken the observations (V_{t_i}) into account, that the function is an element of a suitable function space H and to construct a probability measure π_0, referred to as a *prior*, on this function space which reflects a scientific consensus belief elicited from interaction with application area experts. The SDE gives rise to a probability measure on the observations $P((V_{t_i})|\xi,\sigma)$ which is combined with the prior measure π_0 using Bayes' theorem to yield the posterior measure; see [2] for the Bayesian viewpoint in general and [10] for an exposition in the context of nonparametric estimation and inverse problems.

If observations are available at all points of the interval $[0,T]$, this is known as continuous time observation and a rich theory exists to address this problem, see [6]. This setting has also been considered in [8] as its relative simplicity enables a study of the frequentist behaviour of the Bayesian procedure proposed in dimension one. If observations are available at a finite number of time points with maximal inter-observation time $\Delta t = \max\{t_{i+1} - t_i\}$ and it is acceptable to consider the limit $\Delta t \to 0$, $T \to \infty$, then many results are available, see [9] in the parametric case and e.g. [5] in the nonparametric case.

This work assumes a simpler parametric form for the diffusivity $\sigma(v) \equiv \Sigma \in \mathbb{R}^{2\times 2}$ and presents fully nonparametric estimation of the drift $\xi(\cdot)$. In extension of [7, 8], the state space considered is the two-dimensional torus, or equivalently the unit square with periodic boundary conditions but the work differs from [4] by defining the process on the torus rather than mapping a diffusion on $\mathbb{R}^2$ to the torus by a modulus operation. The prior measure is Gaussian and described by the prior mean function $\xi_0 \equiv 0$ and the prior precision operator

$$\mathcal{A}_0 = \eta_o I + \eta(\partial_x^8 + \partial_y^8),$$

where x and y refer to the two coordinates describing the state space $S = [0,1]^2 \ni (x,y)^T$ and $\eta > 0$, $\eta_o > 0$ are so-called hyperparameters that are chosen to more carefully reflect prior beliefs on the drift.

As in the univariate case, conditional on continuous time observations and diffusivity, the posterior follows a Gaussian measure and the update equations connecting prior to posterior mean and precision are of the same form given in [7], i.e. the posterior mean is given as the solution of

$$\int_S \varphi(v)\mathcal{A}\hat{\xi}(v)\mathrm{d}v = \frac{1}{2}\int_0^T \varphi(V_t)\mathrm{d}V_t \qquad \forall \varphi \in \mathcal{D}(\mathcal{A}), \tag{1}$$

where the posterior precision is given by

$$\mathcal{A} = \mathcal{A}_0 + \gamma_T,$$

where, in turn, γ_T is the empirical measure of the process $\{V_t\}_{t=0}^T$:

$$\int_0^T \varphi(V_t)\mathrm{d}t = \int_S \varphi(v)\mathrm{d}\gamma_T(v) \qquad \forall \varphi \in \mathbb{C}(S).$$

Discretization is carried out via a truncated Fourier representation using preconditioned conjugate gradient methods to solve the PDE (1) and to sample from the posterior measure using the Krylov-based methods reviewed in [1]. The algorithm is complemented by a Langevin-based sampling method for data augmentation and a Gibbs sampler. Finally, an application to animal movement modelling is displayed briefly where position observations of a single Capuchin monkey are obtained at not quite regularly spaced observation times. It is found that acceptable model fit is obtained only upon sub-sampling of the data and the drift appears non-conservative (i.e. it contains a rotational component) which precludes simpler models present in the literature on animal movement ecology where the drift is modelled as the gradient of a potential, e.g. [3].

References

[1] E. Aune, J. Eidsvik, Y. Pokern, *Iterative numerical methods for sampling from high dimensional Gaussian distributions*, Statistics and Computing, (2012).

[2] J. M. Bernardo, A. F. M. Smith *Bayesian Theory*, Wiley (2000).

[3] D. R. Brillinger, H. K. Preisler, A. A. Ager, J. G. Kie *The Use of Potential Functions in Modelling Animal Movement*, In: Data Analysis from Statistical Foundations, Nova Science, Huntington (2007), pp. 369–386

[4] E. Garcia-Portugues, M. Sorensen, K. V. Mardia, T. Hamelryck, *Langevin Diffusion on the Torus: estimation and applications*, arxiv1705.00296

[5] F. Comte, V. Genon-Catalot, Y. Rozenholc, *Penalized nonparametric mean square estimation of the coefficients of diffusion processes*, Bernoulli, **13**, 2, (2007), 514–543.

[6] Y. A. Kutoyants, *Statistical Inference for Ergodic Diffusion Processes*, Springer (2004).

[7] O. Papaspiliopoulos, Y. Pokern, G. O. Roberts, A. M. Stuart, *Nonparametric estimation of diffusions: a differential equations approach*, Biometrika **99**, 2 (2012), 511–531.

[8] Y. Pokern, A. M. Stuart, H. J. van Zanten *Posterior consistency via precision operators for Bayesian nonparametric drift estimation in SDEs*, Stochastic Processes and Their Applications **132**, 2 (2013), 603–628.

[9] B .L. S. Prakasa Rao, *Statistical Inference for Diffusion Type Processes*, Arnold Publishers, London (1999).

[10] A. M. Stuart, *Inverse Problems: A Bayesian perspective*, Acta Numerica **19**, (2010), 451–559.