Report on the manuscript

Fast Automatic Bayesian Cubature Using Lattice Sampling

by

R. Jagadeeswaran and F. J. Hickernell,

submitted to 'Statistics and Computing'.

The authors study a stopping criterion for the numerical integration of realizations of Gaussian random fields with parametric families of covariance kernels. They present an abstract set of assumptions concerning the kernel and the nodes of the underlying cubature formula that allow a fast estimation of the unknown parameters and, hereby, a fast computation of credible intervals. Furthermore, it is shown that the assumptions are satisfied for shift-invariant kernels and nodes from certain integration lattices.

There is no theory available that covers the whole process (estimation and cubature with credible interval); instead, the authors present numerical experiments to demonstrate the performance of three variants of their new algorithm.

I recommend to ask for a revision of the paper, which takes into account my comments and remarks below.

Comments and Remarks

p. 3, Sec. 2.2 The application of Lemma 1 to automatic Bayesian cubature is not clear to me, since integrand data is accumulated sequentially, which does not correspond to applying a single linear mapping to all integrands (realizations) f.

Sec. 2.2.1, 2.2.2 I suggest to include some references for the results that are presented in Sections 2.2.1 and 2.2.2.

p. 5, Sec. 2.4 Why are Sobol points chosen for the Matern kernel? Is there some theoretical link?

In the example the integral (22) is considered over a compact set. What is the benefit of Genz's transformation, compared to a simple affine linear transformation, in this case?

p. 6 Why Hermitian? Are you considering complex-valued kernels?

p. 7, top left Definition of v_1^*?

...where ... (lower case)

Sec. 3.2–3.4 Does the operation count also include the computational cost for the minimization problem to estimate θ? I suspect that this is not the case, since there is no assumption concerning the dependence of C_θ on θ. The same question arises in the setting of Section 4.

Sec. 4.1 and 4.2 As the authors state, larger r implies a greater degree of smoothness of the kernel. This should lead to a greater amount of smoothness of the integrands, and thus to a fast convergence of suitable cubature formulas. It might be that this is not exploited by the integration lattices (plus periodization), as studied in the present paper. Please comment.

p. 9, eqn. (37) $i, j = 1$

p. 9, bottom, right Assumption (25c) ...

p. 12, center left ...differs depends ...?

Sec. 5.1 Is there some kind of a universal periodization that 'works' for every value of r?

Sec. 5.2 You take $r = 2$ or $r = 1$ for the periodization in these examples. Does this mean that $\theta = (2, \gamma)$ is considered for the shift-invariant kernel, or do you still consider $\theta = (r, \gamma)$ with $r \in \mathbb{N}$ as the unknown parameter?

I suggest not to use r also in the drift term in the option pricing example.