Active learning

Example : BOLFI

Journal of Machine Learning Research 17 (2016) 1-47 Submitted 1/15; Revised 8/15; Published 8/16

Bayesian Optimization for Likelihood-Free Inference of Simulator-Based Statistical Models

Michael U. Gutmann MICHAEL.GUTMANN@HELSINKI.FI
Helsinki Institute for Information Technology HIIT
Department of Mathematics and Statistics, University of Helsinki
Department of Information and Computer Science, Aalto University

Jukka Corander JUKKA.CORANDER@HELSINKI.FI
Helsinki Institute for Information Technology HIIT
Department of Mathematics and Statistics, University of Helsinki

Editor: Nando de Freitas

Abstract

Our paper deals with inferring simulator-based statistical models given some observed data. A simulator-based model is a parametrized mechanism which specifies how data are generated. It is thus also referred to as generative model. We assume that only a finite number of parameters are of interest and allow the generative process to be very general; it may be a noisy nonlinear dynamical system with an unrestricted number of hidden variables. This weak assumption is useful for devising realistic models but it renders statistical inference very difficult. The main challenge is the intractability of the likelihood function. Several likelihood-free inference methods have been proposed which share the basic idea of identifying the parameters by finding values for which the discrepancy between simulated and observed data is small. A major obstacle to using these methods is their computational cost. The cost is largely due to the need to repeatedly simulate data sets and the lack of knowledge about how the parameters affect the discrepancy. We propose a strategy which combines probabilistic modeling of the discrepancy with optimization to facilitate likelihood-free inference. The strategy is implemented using Bayesian optimization and is shown to accelerate the inference through a reduction in the number of required simulations by several orders of magnitude.

Keywords: intractable likelihood, latent variables, Bayesian inference, approximate Bayesian computation, computational efficiency

1. Introduction

We consider the statistical inference of a finite number of parameters of interest $\boldsymbol{\theta} \in \mathbb{R}^d$ of a simulator-based statistical model for observed data $\mathbf{y}_o$ which consist of n possibly dependent data points. A simulator-based statistical model is a parametrized stochastic data generating mechanism. Formally, it is a family of probability density functions (pdfs) $\{p_{\mathbf{y}|\boldsymbol{\theta}}\}_{\boldsymbol{\theta}}$ of unknown analytical form which allow for exact sampling of data $\mathbf{y}_{\boldsymbol{\theta}} \sim p_{\mathbf{y}|\boldsymbol{\theta}}$. In practical terms, it is a computer program which takes a value of $\boldsymbol{\theta}$ and a state of the random number generator as input and returns data $\mathbf{y}_{\boldsymbol{\theta}}$ as output. Simulator-based models are also called implicit models because the pdf of $\mathbf{y}_{\boldsymbol{\theta}}$ is not specified explicitly (Diggle and Gratton, 1984), or generative models because they specify how data are generated.

Bayesian Analysis (2019) 14, Number 2, pp. 595–622

Efficient Acquisition Rules for Model-Based Approximate Bayesian Computation

Marko Järvenpää*, Michael U. Gutmann†, Arijus Pleska‡, Aki Vehtari§, and Pekka Marttinen¶

Abstract. Approximate Bayesian computation (ABC) is a method for Bayesian inference when the likelihood is unavailable but simulating from the model is possible. However, many ABC algorithms require a large number of simulations, which can be costly. To reduce the computational cost, Bayesian optimisation (BO) and surrogate models such as Gaussian processes have been proposed. Bayesian optimisation enables one to intelligently decide where to evaluate the model next but common BO strategies are not designed for the goal of estimating the posterior distribution. Our paper addresses this gap in the literature. We propose to compute the uncertainty in the ABC posterior density, which is due to a lack of simulations to estimate this quantity accurately, and define a loss function that measures this uncertainty. We then propose to select the next evaluation location to minimise the expected loss. Experiments show that the proposed method often produces the most accurate approximations as compared to common BO strategies.

Keywords: approximate Bayesian computation, intractable likelihood, Gaussian processes, Bayesian optimisation, sequential experiment design.

1 Introduction

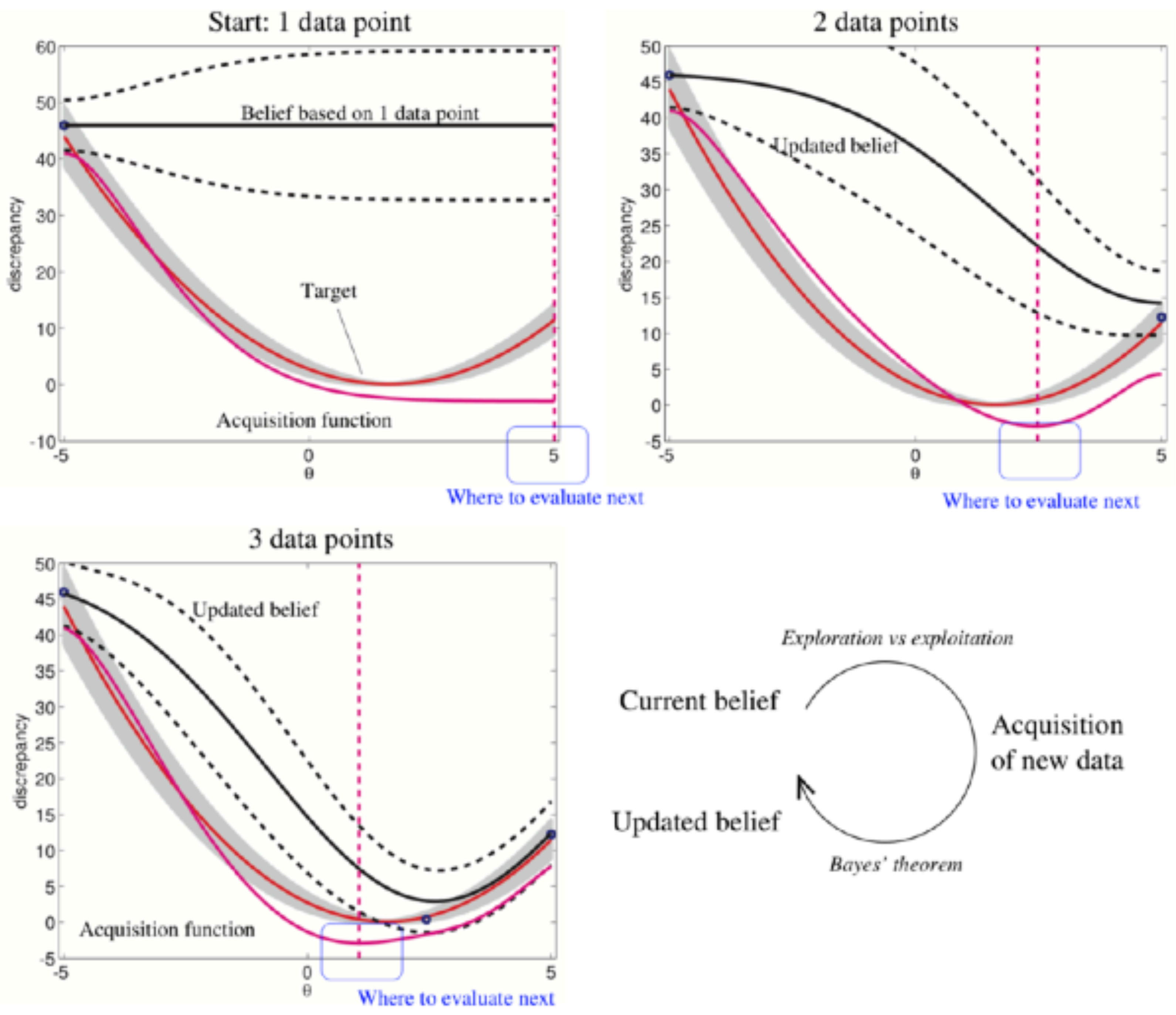
We consider the problem of Bayesian inference of some unknown parameter $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$ of a simulation model. Such models are typically not amenable to any analytical treatment but they can be simulated with any parameter $\boldsymbol{\theta} \in \Theta$ to produce data $\mathbf{x}_{\boldsymbol{\theta}} \in \mathcal{X}$. Simulation models are also called simulator-based or implicit models (Diggle and Gratton, 1984). Our prior knowledge about the unknown parameter $\boldsymbol{\theta}$ is represented by the prior probability density $\pi(\boldsymbol{\theta})$ and the goal of the analysis is to update our knowledge about the parameters $\boldsymbol{\theta}$ after we have observed data $\mathbf{x}_{obs} \in \mathcal{X}$.

If evaluating the likelihood function $\pi(\mathbf{x} \,|\, \boldsymbol{\theta})$ is feasible, the posterior distribution can be computed directly using Bayes' theorem

$$\pi(\boldsymbol{\theta} \,|\, \mathbf{x}_{obs}) = \frac{\pi(\boldsymbol{\theta})\pi(\mathbf{x}_{obs} \,|\, \boldsymbol{\theta})}{\int_\Theta \pi(\boldsymbol{\theta}')\pi(\mathbf{x}_{obs} \,|\, \boldsymbol{\theta}')\,\mathrm{d}\boldsymbol{\theta}'} \propto \pi(\boldsymbol{\theta})\pi(\mathbf{x}_{obs} \,|\, \boldsymbol{\theta}). \quad (1)$$

*Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, marko.j.jarvenpaa@aalto.fi
†School of Informatics, University of Edinburgh, michael.gutmann@ed.ac.uk
‡Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, pleska.arijus@gmail.com
§Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, aki.vehtari@aalto.fi
¶Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, pekka.marttinen@aalto.fi

 https://doi.org/10.1214/18-BA1121



BOLFI

Example : MA(2)

- Model parameters vs discrepancy with GP
- Find parameter values that minimize discrepancy function
  - Black-box optimization
  - Acquisition strategies balance exploration and exploitation