Minimizing distance is often not optimal

Ultimate goal is to approximate posterior distribution

- How to choose query points that are most informative about the posterior distribution
- Active learning strategies can be based on more reasonable goals

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

Acquisition functions

- Lower Confidence Bound Selection Criterion
- Maximum Variance
- Randomized Maximum Variance
- Expected Integrated Variance