Technische Universität München
Department of Mathematics

Bachelor's Thesis

Sequential Monte Carlo for time-dependent Bayesian Inverse Problems

Matthieu Bulté

Supervisor: Ullmann, Elisabeth; Prof. Dr. rer. nat.
Advisor: Latz, Jonas; M.Sc
Submission Date: 15. June 2018

With my signature below, I assert that the work in this thesis has been composed by myself independently and no source materials or aids other than those mentioned in the thesis have been used.

<table>
<tr><td>München, April 17, 2018</td><td></td></tr>
<tr><td>Place, Date</td><td>Signature</td></tr>
</table>

Abstract

Titel auf Englisch wiederholen.

Es folgt die englische Version der Kurzfassung.

Contents

1 Introduction

Study of complex system if often done through mathematical modelling, allowing to simulate, analyse and predict their behaviour. These mathematical models require input parameters, for which only limited or no information is known. Finding these parameters from measurements of the system is called the *inverse problem*. Since measurements are often noisy or sparse, and the mathematical models can be complex and expensive to evaluate, developing sound and efficient mathematical frameworks to treat the inverse problem is a rich and difficult task.

A general description of the inverse problem is following: given a system modelled by a function $\mathcal{G} : X \to Y$, called the *forward response operator*, mapping elements of the *parameter space* X to the *data space* Y, how much information can be extracted from the data $y \in Y$ about the parameter $\hat{x} \in X$ generating it? Or, considering the formula

$$y = \mathcal{G}(\hat{x}),$$

to which extent can the function $\mathcal{G}$ be *inverted*?

Two prominent classes of methods for attempting to answer this question are the *maximum likelihood* methods and the *Bayesian* methods. In maximum likelihood methods, the solution of the inverse problem is given as the maximizer of the likelihood of the observed data. In Bayesian methods, the system is re-modeled probabilistically with random variables, and the solution is given as a marginalization of the model by the observed data using Bayes' formula, as first developed by Laplace [Lap20]. A theoretical and practical comparison of the two methods is given by Kaipio and Somersalo [KS06], including a broad introduction to solving inverse problems found in science and engineering.

The Bayesian approach is a very general modelling and inference framework allowing to approach very different kind of statistical problems, the work by Gelman et al. [GCS+14] give a broad introduction to the field of *Bayesian data analysis*. We focus on using Bayesian inference for solving time-dependent inverse problems found in dynamical systems, based on the framework presented bt Stuart [Stu10]. This will show that under weak assumptions of the model, the solution can be shown to be *well-posed*, using a definition of well-posedness similar to Hadamard's [Had02].

Very often, the solution of an inverse problem given in the Bayesian framework does not admit any analytical solution. Since one is interesting in obtaining summerized statistics about the solution of the inverse problem, such as mean and variance, numerical approximations will involve computing integrals over the parameter space. Since volumes grow exponentially with the number of dimensions, classical methods of numerical integration cannot be used for high-dimensional problems. This phenomenon is called the *curse of dimensionality*.

Fortunately, other numerical approximations were developed that do not suffer fromn the curse of dimensionality. Typically, such approximations work by generating pseudo-random values distributed according to the posterior distribution and use them to approximate the hard integral. Some variation of the law of large number will then provide dimensionality-free error bounds, making these methods suited for high-dimensional problems. A common class of algorithms falling in this category are the *Markov Chain Monte Carlo* (MCMC) methods, presented by Metropolis et al. [MRR⁺53] for a specific class of problems, and later extended to the general case by Hastings in [Has70]. While having dimensionality-free error bounds, MCMC algorithms often need a lot of knowledge and tunning to operate them properly. A simpler method to operate is *importance sampling* (IS) (NOTE: need ref) , where the sampling is done by chosing a auxiliary distribution similar to the target distribution, but from which direct sampling is easier. The discrepancy between the generated samples and a sample generated from the posterior distribution is then corrected by assigning correction weights to the values of the sample. However, chosing an auxiliary close to the target distribution is not always possible, and failing to do so results in a poor estimation of the posterior distribution.

Sequential Monte Carlo (SMC) [DMDJ06] is a method merging ideas of MCMC and IS samplers in an attempt to solve major problems found in these other two methods. This sampler was created to approximate a sequence of distributions, such as those found in data assimilation problems, but can be also be used on an artificial sequence of distributions to interpolate between a simple auxiliary distribution and the true posterior, this is done by Beskos et al. in [BJMS15] for approximating the solution of the Bayesian inverse problem associated to elliptic PDEs. By drawing parallels to particle physics, Del Moral [DM13, DM04] provides convergence results of the algorithm that will be presented in this thesis.

In their work, Allmaras et al. [ABL⁺13] give a case study-based introduction to the whole process of Bayesian techniques for solving inverse problems. This thesis will follow a similar approach, by structuring itself around a simple time-dependent Bayesian inversion problem. The system studied is the simple pendulum, an idealized model for a pendulum in which the mass of the pendulum and its friction are ignored. It can be described by a second-order, non-linear differential equation with a parameter representing the *gravitational acceleration*. We will describe the model of the pendulum, together with the inversion task of estimating the gravitational acceleration from a set of measurements taken in a experiment.

The rest of the thesis is structured as follows. In Section 2 we describe Bayesian inversion in the context of time-dependent problems. Moreover, we show how to formulate the pendulum problem in the Bayesian framework. In Section 3, we present the construction of the SMC algorithm and show the parallels to IS and MCMC. We also present a proof of convergence of the algorithm and discuss possible extensions. We conclude the section by computing and comparing numerical solutions to the pendulum problem. Finally, Section 4 discusses other application areas and current research on SMC algorithms.

2   Bayesian Inverse Problem

2.1   Overview

This section will introduce

A more mathematical formulation of the inverse problem is the following: given some *observation* $y \in Y$, what is the parameter $x \in X$ solving the equation

$$y = \mathcal{G}(x) \tag{1}$$

where $\mathcal{G}$ is the , mapping values from the *parameter space* X to the *data space* Y, with X, Y vector spaces of finite dimension.

thesis is a simple pendulum, an idealized model for a pendulum in which the mass of the pendulum and its friction are ignored, described by the following second-order and non-linear differential equation;

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} = -\frac{g}{l}\sin(\theta), \tag{2}$$

where θ is the angle of the pendulum to its resting point, l is the length of the pendulum and g is the acceleration due to gravity, called *gravitational acceleration*. The inverse problem will be to estimate the gravitational acceleration from an data of an experiment in which the pendulum is let go with no initial velocity from an angle θ_0, and where to measurements are the times $(t_1, \ldots, t_n)$ at which the pendulum reaches a zero angle.

3 Sequential Monte Carlo

4   Case Study

5   Conclusion

A Appendix

List of Figures

List of Tables

References

[ABL+13] Moritz Allmaras, Wolfgang Bangerth, Jean Marie Linhart, Javier Polanco, Fang Wang, Kainan Wang, Jennifer Webster, and Sarah Zedler. Estimating parameters in physical models through bayesian inversion: A complete example. *SIAM Review*, 55(1):149–167, 2013.

[BJMS15] Alexandros Beskos, Ajay Jasra, Ege A Muzaffer, and Andrew M Stuart. Sequential monte carlo methods for bayesian elliptic inverse problems. *Statistics and Computing*, 25(4):727–737, 2015.

[DM04] Pierre Del Moral. *Feynman-kac formulae*. Springer, 2004.

[DM13] Pierre Del Moral. *Mean field simulation for Monte Carlo integration*. Chapman and Hall/CRC, 2013.

[DMDJ06] Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. Sequential monte carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436, 2006.

[GCS+14] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL, 2014.

[Had02] Jacques Hadamard. Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.

[Has70] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[KS06] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*, volume 160. Springer Science & Business Media, 2006.

[Lap20] Pierre Simon Laplace. *Théorie analytique des probabilités*. Courcier, 1820.

[MRR+53] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[Stu10]   A. M. Stuart. Inverse problems: A bayesian perspective. *Acta Numerica*, 19:451–559, 2010.