Computational Statistics and Data Analysis 97 (2016) 98–113

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Sequentially Constrained Monte Carlo

Shirin Golchi *, David A. Campbell

University of British Columbia, Canada
Simon Fraser University, Canada

ARTICLE INFO

Article history:
Received 24 February 2015
Received in revised form 1 September 2015
Accepted 24 November 2015
Available online 28 November 2015

Keywords:
Particle filtering
Constraints
Differential equation models
Parameter estimation
Approximate Bayesian Computation
Sequential Monte Carlo

ABSTRACT

Constraints can be interpreted in a broad sense as any kind of explicit restriction over the parameters. While some constraints are defined directly on the parameter space, when they are instead defined by known behavior on the model, transformation of constraints into features on the parameter space may not be possible. Incorporation of constraints into the model often leads to truncations in the parameter space and multimodality which in turn cause difficulties in posterior sampling. A variant of the Sequential Monte Carlo algorithm is proposed by defining a sequence of densities through the imposition of the constraint. Particles generated from an unconstrained or mildly constrained distribution are filtered and moved through sampling and resampling steps to obtain a sample from the fully constrained target distribution. General and model specific forms of constraints enforcing strategies are defined. The Sequentially Constrained Monte Carlo algorithm is demonstrated on constraints defined by monotonicity of a function, densities constrained to low dimensional manifolds, adherence to a mechanistic differential equation model, and Approximate Bayesian Computation.

Crown Copyright © 2015 Published by Elsevier B.V. All rights reserved.

1. Introduction

Constraints are tools to incorporate external information and ensure parameters remain interpretable. In non-linear or high dimensional models, parameter constraints are complex to define, even when constraints on the model expectation surface may remain simple. Constraints can be defined in a broad sense as any explicit restriction over the parameter or model space. A few examples are: inequality constraints over model parameters to enforce physical laws such as conservation of mass and energy; ensuring monotonicity or convexity of functions in regression or non-parametric smoothing; conforming to theoretical behavior governed by a model.

Imposition of constraints can induce multimodality and/or zero probability regions resulting in challenges in sampling random variable $\boldsymbol{\theta}$ from a target distribution and in the Bayesian context may also lead to disagreements between the prior and likelihood. In this paper we extend the utility of Sequential Monte Carlo (SMC) samplers (Del Moral et al., 2006) by defining a sequence of distributions by their enforcement of a constraint through the proposed Sequentially Constrained Monte Carlo (SCMC) algorithm. We connect a "simple" distribution, $\pi_0(\boldsymbol{\theta})$ to the target distribution, $\pi_T(\boldsymbol{\theta})$, via a path defined by the strictness of constraint enforcement, thereby generalizing the usual transitions of SMC between $\pi_0(\boldsymbol{\theta})$ and $\pi_T(\boldsymbol{\theta})$. Furthermore, we show general applicability of SCMC by creatively defining constraints.

To showcase SCMC, we begin with the toy problem of polynomial regression on noisy observations with constraints over the first and second derivatives. Sequentially imposing the constraints by defining a "soft" positivity constraint over

* Correspondence to: University of British Columbia - Okanagan, 1177 Research Rd, ASC 310, Kelowna, BC, Canada.
E-mail address: golchi.shirin@gmail.com (S. Golchi).

http://dx.doi.org/10.1016/j.csda.2015.11.013
0167-9473/Crown Copyright © 2015 Published by Elsevier B.V. All rights reserved.

Algorithm 1 Sequential Monte Carlo Sampler

Input: Forward and backward kernels, K_t and L_t.

Generate an initial sample $\boldsymbol{\theta}_0^{1:N} \sim \pi_0$;

$W_0^j \leftarrow \frac{1}{N}, j = 1, \dots, N$;

for $t := 1, \dots, T$ **do**

- **if** $\text{ESS} = \left(\sum_{j=1}^N \left(W_{t-1}^j\right)^2\right)^{-1} < \frac{N}{2}$ **then**
- resample $\boldsymbol{\theta}_{t-1}^{1:N}$ with weights $W_{t-1}^{1:N}$
- $W_{t-1}^{1:N} \leftarrow \frac{1}{N}$
- **end if**
- Sample $\boldsymbol{\theta}_t^{1:N} \sim K_t$;
- $W_t^j \leftarrow W_{t-1}^j w_t^j$ where $w_t^j = \frac{\eta_t(\boldsymbol{\theta}_t^j) L_{t-1}(\boldsymbol{\theta}_t^j, \boldsymbol{\theta}_{t-1}^j)}{\eta_{t-1}(\boldsymbol{\theta}_{t-1}^j) K_t(\boldsymbol{\theta}_{t-1}^j, \boldsymbol{\theta}_t^j)}, j = 1, \dots, N$;
- Normalize $W_t^{1:N}$.

end for

Return: Particles $\boldsymbol{\theta}_{1:T}^{1:N}$.

the derivative polynomials produces more accurate predictions while satisfying the monotonicity and convexity/concavity constraints. The second example involves sampling a bivariate density constrained to take non-zero probabilities only when both variables lie on lower dimensional manifolds. In our third application we estimate a mixture of discrete and continuous parameters of an ordinary differential equation model where we generalize the usual definition of constraint to include model adherence. In this case, full constraint enforcement produces multiple disjoint modes. The fourth example focuses on parameter estimation for a chaotic stochastic differential equation model where we define the constraint through a form of sequentially expanding model adherence criteria in an ABC algorithm. The first two applications define a general strategy for enforcing constraints, whereas the last two constraints showcase problem specific strategies.

The rest of the paper is organized as follows; in Section 2, we provide a background on SMC samplers and the commonly used versions of it. We explain the choice of the sequence of densities that outlines the SCMC in Section 3. In Section 4, a general strategy is defined through a probit based soft constraint model. This SCMC strategy is applied to the derivative constrained polynomial regression model and a 1 dimensional model embedded on a nonlinear manifold in the 2 dimensional sampling space while avoiding the need for a Jacobian of the transformation. Some application specific constraint strategies are introduced in Section 5. Parameter estimation for ODE models in a sequential framework is explained in Section 5.1. In Section 5.2, the Sequentially Constrained Approximate Bayesian Computation (SCMC ABC) algorithm is introduced and Section 6 follows with concluding discussion.

2. Sequential Monte Carlo

SMC samplers are a family of algorithms that can be used in many challenging scenarios where random walk Markov chain Monte Carlo (MCMC) methods fail in efficiently sampling $\boldsymbol{\theta}$ from its target distribution. SMC algorithms take advantage of a sequence of bridging distributions that bridge between $\pi_0(\boldsymbol{\theta})$, a distribution that is straightforward to sample, and $\pi_T(\boldsymbol{\theta})$, a difficult to sample target distribution. In Bayesian inference, these distributions are typically the prior $\pi_0(\boldsymbol{\theta})$ and posterior $\pi_T(\boldsymbol{\theta})$.

SMC discretizes a sequence of densities

$$\left\{\pi_t(\boldsymbol{\theta}) = \frac{\eta_t(\boldsymbol{\theta})}{Z_t}\right\}_{t=0}^T$$

between $\pi_0(\boldsymbol{\theta})$ and $\pi_T(\boldsymbol{\theta})$ with possibly unknown normalizing constant Z_t and kernel η_t which can be evaluated for a given $\boldsymbol{\theta}$. The initial sample of particles, $\boldsymbol{\theta}_0^{1:N} \sim \pi_0(\boldsymbol{\theta})$ are filtered through iterative jittering and importance resampling steps to eventually obtain a sample from $\pi_T(\boldsymbol{\theta})$ as outlined in Algorithm 1 (Del Moral et al., 2006).

Algorithm 1, is very general in the sense that many possible choices could be made for the inputs of the algorithm. The choice of the inputs, especially the forward kernels used for jittering the sample, $K_t(\cdot)$, and the backward kernels, L_t, that ensure the weights are defined according to the posterior at time t, can change the order of the steps in Algorithm 1. A variety of options for the forward and backward kernels and the resulting expressions for the incremental weights, w_i, are provided by Del Moral et al. (2006). In the following, we explain the specific choices that are made for all our examples.

At algorithmic stage t, the forward kernel, K_t is chosen to be a MCMC kernel of invariant distribution π_t. The associated backward kernel recommended by Del Moral et al. (2006) for this choice of K_t is

$$L_{t-1}(\theta_t, \theta_{t-1}) = \frac{\pi_t(\boldsymbol{\theta}_{t-1}) K_t(\boldsymbol{\theta}_{t-1}, \boldsymbol{\theta}_t)}{\pi_t(\boldsymbol{\theta}_t)}.$$

Download English Version:

https://daneshyari.com/en/article/6869337

Download Persian Version:

https://daneshyari.com/article/6869337

Daneshyari.com