Stat Comput (2016) 26:679–692
DOI 10.1007/s11222-015-9551-z

Approximate Bayesian computation with composite score functions

Erlis Ruli · Nicola Sartori · Laura Ventura

Received: 23 May 2014 / Accepted: 19 February 2015 / Published online: 1 March 2015

Abstract Both approximate Bayesian computation (ABC) and composite likelihood methods are useful for Bayesian and frequentist inference, respectively, when the likelihood function is intractable. We propose to use composite likelihood score functions as summary statistics in ABC in order to obtain accurate approximations to the posterior distribution. This is motivated by the use of the score function of the full likelihood, and extended to general unbiased estimating functions in complex models. Moreover, we show that if the composite score is suitably standardised, the resulting ABC procedure is invariant to reparameterisations and automatically adjusts the curvature of the composite likelihood, and of the corresponding posterior distribution. The method is illustrated through examples with simulated data, and an application to modelling of spatial extreme rainfall data is discussed.

Keywords Complex model · Composite marginal likelihood · Likelihood-free inference · Pairwise likelihood · Tangent exponential model · Unbiased estimating function

Electronic supplementary material The online version of this article (doi:10.1007/s11222-015-9551-z) contains supplementary material, which is available to authorized users.

E. Ruli (✉) · N. Sartori · L. Ventura
Department of Statistical Sciences, University of Padua, Padua, Italy
e-mail: ruli@stat.unipd.it

N. Sartori
e-mail: sartori@stat.unipd.it

L. Ventura
e-mail: ventura@stat.unipd.it

1 Introduction

The summary of the data on a given model offered by the likelihood function is the key ingredient of all likelihood-based inferential methods. However, likelihood inference, both frequentist and Bayesian, is difficult or even impossible to perform when the likelihood is analytically or computationally intractable. This usually occurs in the presence of complex models, such as models with complicated dependence structures or in models with many latent variables.

In these situations, for frequentist or Bayesian inference, surrogates of the ordinary likelihood are useful. A notable contribution is given by composite likelihoods (Lindsay 1988), which are based on the composition of suitable lower dimensional densities, such as bivariate marginal (Cox and Reid 2004), conditional or full conditional densities (Varin et al. 2011). The use of composite likelihoods has been widely advocated in different complex applications of frequentist inference (see Varin et al. 2011, for a general review, and Larribe and Fearnhead 2011, for a review in genetics).

From a Bayesian perspective, when the computation of the likelihood is impracticable, but it is easy to simulate from the model, an approximation of the posterior distribution can be obtained by approximate Bayesian computation (ABC). The idea of ABC is to simulate from the model for different parameter values, and to keep those values that produce simulated datasets that approximately match the observed data (see Beaumont 2010; Marin et al. 2012). The most popular ABC approach is to consider an approximate matching of some summary statistics, evaluated at the observed and simulated data, by means of suitable distances. When the statistics are sufficient for the parameters of the model, this method leads to the exact posterior distribution as the distance tends to zero. However, in realistic applications sufficient statistics