Approximate Bayesian Computation Overview

Davi Sales Barreira

FGV - Escola de Matemática Aplicada, Rio de Janeiro, Brasil
davisbarreira@gmail.com

Abstract. Approximate Bayesian Computation (ABC) methods are known as likelihood-free techniques, thus are a useful approach in problems that the likelihood is intractable, e.g., likelihood not available in closed form, or likelihood too expensive to calculate. In this article, we present an overview of the method by replicating the paper Approximate Bayesian computational methods by Marin et al. (2012).

Keywords: Approximate Bayesian Computation · likelihood-free · Monte Carlo.

1 Introduction

1.1 Original ABC

The Approximate Bayesian Computation method was originally described by Rubin (1984) as a thought experiment to explain how to sample from a posterior distribution with a frequency interpretation. The method became proeminent due to the fact that it circumvents the need to calculate the likelihood function in order to obtain the posterior distribution. This can be a very useful feature in scenarios where the likelihood is intractable or too expensive to calculate. One example is in the case where one has latent variables, thus, the likelihood is expressed as:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{y}) = \int \ell^*(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{u}) d\boldsymbol{u} \tag{1}$$

with $\boldsymbol{y}$ being the observed variable, $\boldsymbol{u}$ the latent variable and $\boldsymbol{\theta}$ is the parameter of interest.

Tavaré et al. (1997) introduced the ABC algorithm as a rejection techinique to obtain the posterior distribution without the explicit calculation of the likelihood. This original algorithm is given below.

Algorithm 1: Original ABC method

---
for *i=1 to N* **do**
    repeat
        Sample $\boldsymbol{\theta}' \sim \pi(\cdot)$
        Generate $\boldsymbol{z} \sim p(\cdot \mid \boldsymbol{\theta}')$
    until $\boldsymbol{y} = \boldsymbol{z}$;
end

The proof that the algorithm indeed results in an iid sample from the posterior is shown below. Let $\boldsymbol{y}$ be the observed, $\boldsymbol{\theta}$ the parameter of interest and $\boldsymbol{z}$ the generated samples.

$$f(\boldsymbol{\theta}_i) \propto \sum_{\boldsymbol{z} \in \mathbb{D}} \pi(\boldsymbol{\theta}_i)p(\boldsymbol{z} \mid \boldsymbol{\theta}_i)\mathbb{I}_{\boldsymbol{y}}(\boldsymbol{z}) = \pi(\boldsymbol{\theta}_i)p(\boldsymbol{y} \mid \boldsymbol{\theta}_i) \propto \pi(\boldsymbol{\theta}_i \mid \boldsymbol{y}) \tag{2}$$

The original ABC formulation only works for the case where $\boldsymbol{y}$ is discrete taking finite values, and therefore, an exact match is possible to be obtained in a finite number of simulations. Pritchard et al. (1999) then extended the method to a more general form considering an approximation instead of an exact match. This extended algorithm is shown below, where

– η is a function defining a statistic (e.g. the mean),
– ρ is a distance function,
– ϵ is an acceptance tolerance.

Algorithm 2: ABC method for discrete and continuous distributions

---

for *i=1 to N* **do**
    repeat
        Sample $\boldsymbol{\theta}' \sim \pi(\cdot)$
        Generate $\boldsymbol{z} \sim p(\cdot \mid \boldsymbol{\theta}')$
    until $\rho[\eta(\boldsymbol{y}), \eta(\boldsymbol{z})] \leq \epsilon$;
end

For this ABC algorithm, instead of the actual posterior, we get

$$\pi_\epsilon(\boldsymbol{\theta}, \boldsymbol{z} \mid \boldsymbol{y}) = \frac{\pi(\boldsymbol{\theta})p(\boldsymbol{z} \mid \boldsymbol{\theta})\mathbb{I}_{A_{\epsilon,\boldsymbol{y}}}(\boldsymbol{z})}{\int_{A_{\epsilon,\boldsymbol{y}} \times \boldsymbol{\theta}} \pi(\boldsymbol{\theta})p(\boldsymbol{z} \mid \boldsymbol{\theta})dzd\boldsymbol{\theta}} \tag{3}$$

where, $A_{\epsilon,\boldsymbol{y}} = \{\boldsymbol{z} \in \mathcal{D} \mid \rho[\eta(\boldsymbol{z}), \eta(\boldsymbol{y}) \leq \epsilon]\}$. Hence, for a tolerance (ϵ) "small enough", we expect a good approximation of the real posterior.

$$\pi_\epsilon(\boldsymbol{\theta} \mid \boldsymbol{y}) = \int \pi_\epsilon(\boldsymbol{\theta}, \boldsymbol{z} \mid \boldsymbol{y})d\boldsymbol{z} \approx \pi(\boldsymbol{\theta} \mid \boldsymbol{y}) \tag{4}$$

1.2 Moving Average

We will use the Moving Average model, also denoted as MA(q), for assessing the performance of the ABC methods. The MA(q) process is a stochastic process defined by:

$$y_k = u_k + \sum_{i=1}^{q} \theta_i u_{k-i} \tag{5}$$

where $(u_k)_{k \in \mathbb{Z}} \overset{iid}{\sim} N(0, 1)$. The true posterior distribution of MA(2) and MA(1) models can be numerically computed, since the likelihood function is indeed

avaliable. Therefore, the approximations obtained through ABC can be compared with the true posterior.

For a $q = 2$, imposing the standard identifiability condition we obtain the following conditions:

$$-2 < \theta_1 < 2, \qquad \theta_1 + \theta_2 > -1, \qquad \theta_1 - \theta_2 < 1. \tag{6}$$

Hence, we use an uniform distribution over this triangular region as prior for $\boldsymbol{\theta}$. The likelihood of $\boldsymbol{y} \mid \boldsymbol{\theta}$ is more complex because of the need to integrate $\boldsymbol{u}$.

Bibliography

Marin, J.-M., Pudlo, P., Robert, C. P., and Ryder, R. J. (2012). Approximate bayesian computational methods. *Statistics and Computing*, 22(6):1167–1180.

Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., and Feldman, M. W. (1999). Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, 16(12):1791–1798.

Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Statist.*, 12(4):1151–1172.

Tavaré, S., Balding, D. J., Griffiths, R. C., and Donnelly, P. (1997). Inferring coalescence times from dna sequence data. *Genetics*, 145(2):505–518.