Approximate Bayesian Computation

Davi Barreira

FGV - Escola de Matemtica Aplicada

Table of contents

Objective & Motivation

The objective of this presentation is to give an overview of the Approximate Bayesian Computation (ABC) algorithm through the replication of the paper **Approximate Bayesian computational methods** by Marin et al. (2012).

The paper talks about different variants of ABC by estimating the posterior of Moving Average models.

Objective & Motivation

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.

- Coalecent models in population genetics (Tavaré et al., 1997);
- Species dynamics (Jabot and Lohier, 2016);
- Real-world model of HIV transmission (McKinley et al., 2018).

Objective & Motivation

In some settings where we have latent variables, the likelihood is expressed as:

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

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

Original ABC Algorithm

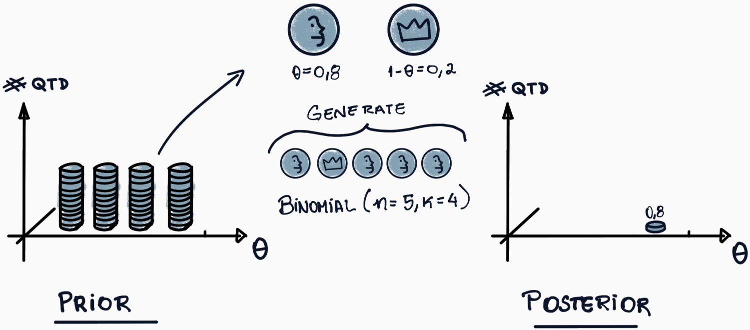
Rubin (1984) described the ABC algorithm as a thought experiment to explain how to sample from a posterior distribution. Tavaré et al. (1997) is usually considered the paper responsible for the proposing ABC for infering the posterior distribution.

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

Original ABC Algorithm

Below we have an schematic drawing with an example of the ABC method for Beta/Binomial model.



Original ABC Algorithm

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.

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

Original ABC Algorithm

Pritchard et al. (1999) extended the original algorithm to the case of continuos sample spaces.

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

- η: function defining a statistic (e.g. the mean),
- ρ: a distance function,
- ϵ: acceptance tolerance.

Original ABC Algorithm

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

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

Where, $A_{\epsilon,\boldsymbol{y}} = \{\boldsymbol{z} \in \mathbb{D} \mid \rho[\eta(\boldsymbol{z}), \eta(\boldsymbol{y}) \leq \epsilon].\}$

Hence, for a tolerance (ϵ) "small enough", we expect a good approximation.

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

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:

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

Where $(u_k)_{k \in \mathbb{Z}} \overset{iid}{\sim} N(0,1)$. For a $q = 2$, imposing the standard identifiability condition we obtain the following conditions:

$$-2 < \theta_1 < 2, \qquad \theta_1 + \theta_2 > -1, \qquad \theta_1 - \theta_2 < 1.$$

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 $\boldsymbol{u}$.

Moving Average

We generate a synthetic sample of length 100 using $(\theta_1, \theta_2) = (0.6, 0.2)$. For $q = 2$ we can also numerically calculate the real posterior and the marginal distributions.

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \pi(\boldsymbol{\theta}) p(\boldsymbol{y} \mid \boldsymbol{\theta}), \qquad \boldsymbol{y} \mid \boldsymbol{\theta} \sim MVN(0, \Sigma)$$

$$\Sigma = \begin{bmatrix} 1+\theta_1^2+\theta_2^2 & \theta_1+\theta_2\theta_1 & \theta_2 & 0 & 0 & 0 & \dots & 0 \\ \theta_1+\theta_2\theta_1 & 1+\theta_1^2+\theta_2^2 & \theta_1+\theta_2\theta_1 & \theta_2 & 0 & 0 & \dots & 0 \\ \theta_2 & \theta_1+\theta_2\theta_1 & 1+\theta_1^2+\theta_2^2 & \theta_1+\theta_2\theta_1 & \theta_2 & 0 & \dots & 0 \\ 0 & \theta_2 & \theta_1+\theta_2\theta_1 & 1+\theta_1^2+\theta_2^2 & \theta_1+\theta_2\theta_1 & \theta_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \theta_2 & \theta_1+\theta_1\theta_2 & 1+\theta_1^2+\theta_2^2 \end{bmatrix}$$

References i

Jabot, F. and Lohier, T. (2016). Non-random correlation of species dynamics in tropical tree communities. *Oikos*, 125(12):1733–1742.

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

McKinley, T. J., Vernon, I., Andrianakis, I., McCreesh, N., Oakley, J. E., Nsubuga, R. N., Goldstein, M., and White, R. G. (2018). Approximate bayesian computation and simulation-based inference for complex stochastic epidemic models. *Statist. Sci.*, 33(1):4–18.

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.