Coupled Markov chains with applications to Approximate Bayesian Computation for model based clustering

E. Bertoni, M. Caldarini, F. Di Filippo, G. Gabrielli, E. Musiari

Politecnico di Milano

February 12, 2022

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

I. Introduction

Start explaining the initial problem

II. ABC and coupled Markov chain Monte Carlo method: mean and variance unknown

Model: We assume independent prior $\pi(\mu, \sigma^2) = \pi(\mu)\pi(\sigma^2)$ as follows:

$$Y_i \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$$
$$\mu \sim \mathcal{N}(\mu_0, \sigma_0^2)\ \sigma^2 \sim InvGa(a_0, b_0)$$

with $\mu_0 = 8, \sigma_0^2 = 3$ and $a = 1, b = 1$.

For the MH ABC and coupled MCMC algorithm we define as proposal distribution for the parameters' update:

$$\mu^{i+1} \sim \mathcal{N}(\mu^i, 0.1^2)\ log(\sigma^{2(i+1)}) \sim \mathcal{N}(log(\sigma^{2(i)}), 0.1^2)$$

the logarithm trasformation is needed to keep the proposed value of σ^2 within its domain, which is $\mathbb{R}^+$

Posterior: It would be impossible to compute the marginal posterior distribution, thus we obtain the conditional posterior distribution, or full conditional distributions: Given $\mathbf{y} \in \mathbb{R}$

$$\mu|\sigma^2, \mathbf{y} \sim \mathcal{N}(\sigma_n \mu_n, \sigma_n^2)$$
$$\sigma^2)|\mu, \mathbf{y} \sim InvGa(a_n, b_n)$$
$$\text{where } \mu_n = n\bar{y}/\sigma_0^2 + \mu_0/\sigma_0^2$$
$$\text{and } a_n = a + n/2n = 2, b_n = b + n(\mu - \bar{y})^2/2 + \textstyle\sum_{i=1}^n (y - \bar{y})^2/2$$

III. Numerical Experiments

In order to test our method on a more complex problem, we followed the common numerical experiment in the literature: ABC applied on the g-and-k distribution. The goal is to, not only, apply ABC algorithm to this problem, but also the Unbiased coupled Markov chain Monte Carlo method.

i. univariate g-and-k distribution

A classical numerical experiments in the ABC literature is to test the algorithm on the univariate g-and-k distribution. The likelihood is intractable, therefore the distribution is defined in terms of its quantile function:

$$r \in (0,1) \longmapsto a + b(1 + 0.8\left(\frac{1 - exp(-g \cdot z(r))}{1 + exp(-g \cdot z(r))}\right)(1 + z(r)^2)^k \cdot z(r)$$

where $z(r)$ is the r-th quantile of the standard Normal distribution, while a and b are location and scale parameters and g and k are related to skewness and kurtosis.

Sampling from this distribution can be done by generating $z(r)$s as samples from a $\mathcal{N}(0,1)$: $z(r) \sim \mathcal{N}(0,1)$

To complete the model we assume prior probability distribution as follows:

$$a \sim \mathcal{U}([0,10])$$
$$b \sim \mathcal{U}([0,10])$$
$$g \sim \mathcal{U}([0,10])$$
$$k \sim \mathcal{U}([0,10])$$

To test this distribution we sampled a set of observation y_{obs} imposing the following values to the parameters, following the path of Jacob et al.:

a=3, b=1, g=2, k=0.5

A key point in the ABC procedure is the decision of the summary statistics, which in this case, are taken as the 10 equidistant quantiles as well as the minimum.

As before we set

Distance:

$$L^2 - norm\ of\ the\ difference\ of\ S(y)\ and\ s_{obs}$$

Kernel function:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}, \quad K_h(u) = \frac{K(\frac{u}{h})}{h}$$

ii. Implementation: ABC

Initialization:

$\theta^{(0)} = (a^{(0)}, b^{(0)}, g^{(0)}, k^{(0)}) \sim \mathcal{U}([0,10]^4)$

for j in (1,...,n):

- $z^{(0,j)}(r) \sim \mathcal{N}(0,1)$
- $y^{(0,j)} \sim quantilefunction(z^{(0,j)}(r))$

- compute the summaries statistics: $s^{(0)} = S(y^{(0)})$

Untill $Kh(\|s^{(0)} - s_{obs}\|) > 0$

- Generate $\theta^{(0)} \sim \mathcal{U}([0,10]^4)$ from prior density.
- Generate $z^{(0)}(r) \sim \mathcal{N}(0,1)$
- Generate a sample of 1000 observations such that $y \sim quantilefunction(z^{(0)}(r), \theta^{(0)})$
- Compute $s^{(0)} = S(y)$

for i in (1,...,M): where M is the number of iterations

1. given $\theta^{(i-1)} = (a^{(i-1)}, b^{(i-1)}, g^{(i-1)}, k^{(i-1)})$
   $\theta^{(i)} \sim \mathcal{N}(\theta^{(i-1)}, I)$
   for j in (1,...,n):
   $z^{(i,j)}(r) \sim \mathcal{N}(0,1)$
   $y^{(ij)} \sim quantilefunction(z^{(i,j)}(r), \theta^{(i-1)})$
2. Compute the summaries $s^{(i)} = S(y^{(i)})$
3. Accept $\theta^{(i)}$ with probability $\frac{Kh(\|s^{(i)} - s_{obs}\|)\pi(\theta^{(i)})}{Kh(\|s^{(i-1)} - s_{obs}\|)\pi(\theta^{(i-1)})}$

iii. ABC and coupled Markov chain Monte Carlo method: inference on one parameter

Parameters b,g,k are fixed equal to the observation parameters, while parameter a is random The model is the one explained above.

Initialization: $a_1^{(0)} \sim \mathcal{U}([0,10])$
$a_2^{(0)} \sim \mathcal{U}([0,10])$
for j in (1,...,n_{obs}):

- $z^{(0,j)}(r) \sim \mathcal{N}(0,1)$
- $y_1^{(0,j)} \sim quantilefunction(z^{(0,j)}(r))$
- compute $s_1^{(0)} = S(y_1^{(0)})$

for j in (1,...,n_{obs}):

- $z^{(0,j)}(r) \sim \mathcal{N}(0,1)$
- $y_2^{(0,j)} \sim quantilefunction(z^{(0,j)}(r))$
- compute the summaries statistics: $s_2^{(0)} = S(y_2^{(0)})$

Untill $Kh(\|s_1^{(0)} - s_{obs}\|) > 0$:

1. Generate $a_1^{(0)} \sim \mathcal{U}([0,10])$ from prior density.
2. Generate $z_1^{(0)}(r) \sim \mathcal{N}(0,1)$
3. Generate a sample of 1000 observations such that $y_1 \sim quantilefunction(z_1^{(0)}(r), a_1^0)$
4. Compute $s_1^{(0)} = S(y_1)$

Untill $Kh(\|s_2^{(0)} - s_{obs}\|) > 0$:

1. Generate $a_2^{(0)} \sim \mathcal{U}([0,10])$ from prior density.
2. Generate $z_2^{(0)}(r) \sim \mathcal{N}(0,1)$
3. Generate a sample of 1000 observations such that $y_2 \sim quantilefunction(z_2^{(0)}(r), a_2^{(0)})$
4. Compute $s_2^{(0)} = S(y_2)$

for i in (1,...,M):

1. generate $a_1^{(i)}, a_2^{(i)}$ from maximalcoupling$(a_1^{(i-1)}, a_2^{(i-1)})$
   for j in (1,...,n):
   - $z^{(i,j)}(r) \sim \mathcal{N}(0,1)$
   - generate: $y_1^{(ij)} \sim quantilefunction(z^{(i,j)}(r), \theta^{(i-1)})$ $y_2^{(ij)} \sim quantilefunction(z^{(i,j)}(r), \theta^{(i-1)})$
2. Compute the summaries $s_1^{(i)} = S(y_1^{(i)})$ and $s_2^{(i)} = S(y_2^{(i)})$
3. acceptance:
   - Accept $a_1^{(i)}$ with probability $\frac{Kh(\|s_1^{(i)} - s_{obs}\|)\pi(a_1^{(i)})}{Kh(\|s_1^{(i-1)} - s_{obs}\|)\pi(a_1^{(i-1)})}$ otherwise $a_1^{(i)} = a_1^{(i-1)}$
   - Accept $a_2^{(i)}$ with probability $\frac{Kh(\|s_2^{(i)} - s_{obs}\|)\pi(a_2^{(i)})}{Kh(\|s_2^{(i-1)} - s_{obs}\|)\pi(a_2^{(i-1)})}$ otherwise $a_2^{(i)} = a_2^{(i-1)}$

The algorithm used before to obtain a two set of observation from the parameters derived from maximal coupling, i.e. maximal coupling of the ys, is not applicable to this case since it would be difficult to find the best proposal distribution $q(\cdot), p(\cdot)$ given that there's not a density of the random variable y. Instead, we decide on a different approach to obtain two Markov chains of the r.v. y which after meeting would stay together. After sampling the paramters θ_1 and θ_2 from the maximal coupling algorithm, we generate n_{obs} samples $z(r) \sim (N)(0,1)$ and then we compute the two Markov Chains

$y_1 \sim quantilefunction(z(r), \theta_1)$ $y_2 \sim quantilefunction(z(r), \theta_2)$.

Therefore, if θ_1 and θ_2 from maximal coupling are identical, then also the chains y_1 and y_2 will stay together.

iv. ABC and coupled Markov chain Monte Carlo method

Initialization given $\theta = (a, b, g, k)$

$\theta_1^{(0)} \sim \mathcal{U}([0,10]^4)$

$\theta_2^{(0)} \sim \mathcal{U}([0,10]^4)$

for j in (1,...,n_{obs}):

- $z^{(0,j)}(r) \sim \mathcal{N}(0,1)$
- $y_1^{(0,j)} \sim quantilefunction(z^{(0,j)}(r), \theta_1^{(0)})$
- compute $s_1^{(0)} = S(y_1^{(0)})$

for j in (1,...,n_{obs}):

- $z^{(0,j)}(r) \sim \mathcal{N}(0,1)$
- $y_2^{(0,j)} \sim quantilefunction(z^{(0,j)}(r), theta_2^{(0)})$
- compute the summaries statistics: $s_2^{(0)} = S(y_2^{(0)})$

Untill $Kh(\|s_1^{(0)} - s_{obs}\|) > 0$:

1. Generate $\theta_1^{(0)} \sim \mathcal{U}([0,10]^4)$ from prior density.
2. Generate $z_1^{(0)}(r) \sim \mathcal{N}(0,1)$
3. Generate a sample of 1000 observations such that $y_1 \sim quantilefunction(z_1^{(0)}(r), \theta_1^0)$
4. Compute $s_1^{(0)} = S(y_1)$

Untill $Kh(\|s_2^{(0)} - s_{obs}\|) > 0$:

1. Generate $\theta_2^{(0)} \sim \mathcal{U}([0,10]^4)$ from prior density.
2. Generate $z_2^{(0)}(r) \sim \mathcal{N}(0,1)$
3. Generate a sample of 1000 observations such that $y_2 \sim quantilefunction(z_2^{(0)}(r), theta_2^{(0)})$
4. Compute $s_2^{(0)} = S(y_2)$

for i in (1,...,M):

1. generate $\theta_1^{(i)}, \theta_2^{(i)}$ from maximalcoupling$(\theta_1^{(i-1)}, \theta_2^{(i-1)})$
   for j in (1,...,n):
   - $z^{(i,j)}(r) \sim \mathcal{N}(0,1)$
   - generate: $y_1^{(ij)} \sim quantilefunction(z^{(i,j)}(r), \theta_1^{(i-1)})$ $y_2^{(ij)} \sim quantilefunction(z^{(i,j)}(r), \theta_1^{(i-1)})$
2. Compute the summaries $s_1^{(i)} = S(y_1^{(i)})$ and $s_2^{(i)} = S(y_2^{(i)})$
3. acceptance:
   - Accept $\theta_1^{(i)}$ with probability $\frac{Kh(\|s_1^{(i)} - s_{obs}\|)\pi(\theta_1^{(i)})}{Kh(\|s_1^{(i-1)} - s_{obs}\|)\pi(\theta_1^{(i-1)})}$ otherwise $\theta_1^{(i)} = \theta_1^{(i-1)}$
   - Accept $\theta_2^{(i)}$ with probability $\frac{Kh(\|s_2^{(i)} - s_{obs}\|)\pi(\theta_2^{(i)})}{Kh(\|s_2^{(i-1)} - s_{obs}\|)\pi(\theta_2^{(i-1)})}$ otherwise $\theta_2^{(i)} = \theta_2^{(i-1)}$

IV. Implementation

V. Results

$$e = mc^2 \tag{1}$$

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

VI. Discussion

i. Subsection One

A statement requiring citation [Figueredo and Wolf, 2009]. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

ii. Subsection Two

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

References

[Figueredo and Wolf, 2009] Figueredo, A. J. and Wolf, P. S. A. (2009). Assortative pairing and life history strategy - a cross-cultural study. *Human Nature*, 20:317–330.