

Bayesian inference on individual-based models by controlling the random inputs

Michael Spence¹, Paul Blackwell²

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 38

¹ School of Mathematics and Statistics, University of Sheffield
M.A.Spence@sheffield.ac.uk
p.blackwell@sheffield.ac.uk

Abstract

Complex models are becoming increasingly popular in ecological modelling. However, quantifying uncertainty, estimating parameters and so on for a model of this sort are complicated by the fact that their probabilistic behaviour is often implicit in its rules or programs, rather than made explicit as in a more conventional statistical or stochastic model.

In a complex stochastic model, the output is dependent on both the parameters and the random inputs i.e. the random numbers used to resolve decisions or generate stochastic quantities within the model. By treating these random inputs as nuisance parameters, often we can turn the model into a deterministic model where small movements in the parameter space result in small changes in the model output. When this is the case it will allow us to use Approximate Bayesian Computation methods with MCMC in order to perform parameter estimation. Controlling the random inputs allows us to move better in the parameter space and improves the mixing of the Markov Chain.

We will use these methods to estimate parameters in an individual-based model which is used to model the population dynamics of a group-living bird, the woodhoopoe.

Keywords: Approximate Bayesian Computation; Intractable likelihood; Complex Models; Individual based models.

1 Introduction

In ecology the need for answering the question “what makes something happen?” as opposed to “what actually happens?” is becoming increasingly popular. These questions lead to building complex models where the different aspects of the system are modelled separately and give rise to the collective behaviour of the system. A natural approach in ecology is to model each individual separately in the system. These are called individual-based models (IBMs) [2].

IBMs generally model behaviour through a series of rules or algorithms, rather than describing it in a formal mathematical way. They are developed with algorithms that are not well tuned from the beginning but require parameters that are either not precise enough in the literature, or simply not concretely measurable [3]. As the probabilistic behaviour of the model is implicit in the rules of the model, the likelihood is generally intractable.

Using Pattern Orientated Methods [4], parameter values can be found indirectly by changing a number of parameters at once and seeing if the model output matches some observed data [6][7]. As IBMs have many parameters that can have complex and interacting effects on the output, this approach may be unproductive [1] so other methods of parameter estimation are required.

Current methods for performing parameter estimation for a model of this kind involve tuning the parameters and trying to match the model output to observed patterns. A range of potential parameter values are proposed and then tested by dividing the potential parameter space up and seeing if the the output matches the observed data.

A more probabilistic version of this is ABC. In ABC parameter values are proposed from a distribution, usually the prior, and the model is run using these parameters. If the model output is similar to the realized data then that parameter is accepted. This is continued until n parameter values have been accepted. Although this is more probabilistic it can be computationally very inefficient if the prior is not very informative.

We propose a method of examining the model by controlling the random inputs in the model which will allow us to turn the stochastic model into a deterministic one. We use this method, coupled with Approximate Bayesian Computation (ABC) methods, to perform parameter estimates on these types of models.

2 Controlling random inputs

Given a stochastic model M with input parameters θ and output

$$X \sim M(\cdot|\theta), \tag{1}$$

X is drawn from a random distribution. However if we condition on the parts that cause the stochasticity, the random inputs \mathbf{u} , then

$$X \sim M(\cdot|\theta, \mathbf{u}) \tag{2}$$

will be deterministic.

By controlling the random inputs we aim to ensure that small changes in the parameters result in small changes in the model output. However, depending on how the inputs are used, a change in a parameter may cause a submodel to require a different number of random inputs to what it required before that could result in a large change in the output.

For example consider a model where only the annual numbers of births and the weights at birth are generated. Let parameter θ control the number of births and u_i , ($i = 1, 2, \dots$) be the sequence of random inputs. Suppose when $\theta = \alpha$, using u_1 there is only one birth. The weight of the birth is determined by u_2 . The next time the number of births will be determined using u_3 . Now suppose that when $\theta = \alpha + \epsilon$, where ϵ is small, using u_1 there may now be two births whose weights involve u_2 and u_3 . The next time the number of births will be determined using u_4 . This causes all of the random inputs to be out of sequence which could result in a large change in the output. This change in output is caused by the change in sequence of random inputs not the parameter θ .

One way to get round this is to control all of the inputs individually so that each submodel will have its own sequence of inputs and then the inputs for each submodel will remain the same regardless to what happened in earlier submodels.

3 Woodhoopoe model

Woodhoopoes are birds that can be found in sub-Saharan Africa [4]. They live in groups just like wolves, with one dominant pair which are the only ones that breed. Neuert et al. [5] used an Individual-based model in order to model the population and group dynamics of the woodhoopoes which was simplified by Railsback and Grimm [4] for use as examples in their recent book.

In the model individual woodhoopoes live in groups with one dominant male and one dominant female. Each woodhoopoe's aim is to become a dominant

animal in one of the groups. Each month, each woodhoopoe dies with a probability of θ_1 . When a dominant dies the eldest subordinate, if there is any, will become the dominant animal. Younger subordinates will leave their group in order to try and become a dominant in another group with probability θ_2 but if they leave the safety of the group they become more vulnerable to predators and have a predation mortality probability of θ_3 .

4 Summary of the talk

In the talk we will introduce this method of controlling the random inputs, show a few examples of how to do it on toy models and then, along with ABC methods, use it to perform parameter estimation on Railsback and Grimm's [4] woodhoopoe model.

References

- [1] V. Grimm and S. Railsback. **Individual-based Modelling and Ecology**. *Princeton Series in Theoretical and Computational Biology*; 2005
- [2] H. Kaiser. **The dynamics of populations as a result of the properties of individual animals**. *Fortschritte der Zoologie*; 1979; 25; pp. 109-136.
- [3] C. Piou, V. Berger and V. Grimm. **Proposing an information criterion for individual-based models developed in a pattern-oriented modelling framework**. *Ecological Modelling*; 2009; 220; pp. 1957-1967.
- [4] S. Railsback and V. Grimm. **Agent-based and individual-based modeling a practical introduction**. *Princeton University Press*; 2012.
- [5] C. Neuert, M. du Plessis, V. Grimm and C. Wissel **Welche ökologischen Faktoren bestimmen die Gruppengröße bei *Phoeniculus purpureus* (Gemeiner Baumhopf) in Südafrika? Ein individuenbasierte Modell**. *Verhandlungen der Gesellschaft für Ökologie*; 1995; 24; pp. 145-149.
- [6] T. Wiegand, F. Jeltsch, I. Hanski and V. Grimm. **Using pattern-oriented modeling for revealing hidden information: A key for recording ecological theory and application**. *Oikos*; 2003; 100; pp. 209-222.
- [7] T. Wiegand, E. Revilla and F. Knauer **Dealing with uncertainty in spatial explicit population models**. *Biodiversity and Conservation*; 2004; 13; pp. 53-78.