

Contributed Discussion on Article by Pratola

Abdolreza Mohammadi ^{*} and Maurits Kaptein ^{*}

Abstract. The author should be commended for his outstanding contribution to the literature on Bayesian regression tree models. The author introduces three innovative sampling approaches which allow for efficient traversal of the model space. In this response, we add a fourth alternative.

Keywords: Markov chain Monte Carlo, Birth-death process, continuous time Markov process, Bayesian regression tree.

1 Background

The algorithm introduced in section 2.4 consists of a combination of birth/death and tree rotation proposals. Essentially, the *birth/death* mechanism explores trees with a different dimension, while the *rotation* mechanism explores alternative trees with the same dimensions. The specific birth/death mechanism proposed is known as reversible jump MCMC (RJ-MCMC) (Green, 1995) and is based on an ergodic discrete-time Markov chain. This algorithm is efficient only if the acceptance rate is high. As the author points out, this is not always the case.

This issue can be overcome by adopting birth-death MCMC (BD-MCMC) which is based on a *continuous*-time Markov process, as an alternative to RJ-MCMC. In this sampling scheme the algorithm explores the model space by jumping to a larger dimension (birth) or lower dimension (death) where each of these events is modeled as independent Poisson processes. The birth and death events thus occur in continuous time and their rates determine the stationary distribution of the process; see Figure 1. In BD-MCMC the moves between models are always accepted making the algorithm extremely efficient. Cappé et al. (2003) have shown, on appropriate rescaling of time, that the RJ-MCMC converges to a continuous time birth-death chain. One advantage of BD-MCMC is its ability to transit to low probability regions that can form a kind of “springboard” for the algorithm to flexibly move from one mode to another. The BD-MCMC algorithm has already been used effectively in the context of graphical models (Mohammadi and Wit, 2015; Mohammadi et al., 2017; Mohammadi and Wit, 2016) and mixture distributions (Stephens, 2000; Mohammadi et al., 2013).

2 Extension to BD-MCMC sampler

To implement the BD-MCMC mechanism we need to prove that the stationary distribution of the birth-death process converges to our target posterior distribution

$$Pr(T|data) \propto \pi(T)L(T).$$

^{*} Dept. of Methodology and Statistics, Tilburg University, The Netherlands a.mohammadi@uvt.nl
m.c.kaptein@uvt.nl

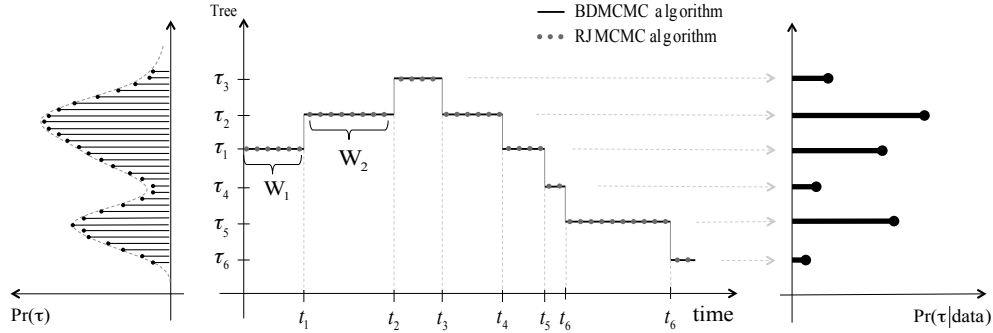


Figure 1: Visualization of the BD-MCMC algorithm. The left panel shows the true posterior distribution and the middle panel compares to progression of BD-MCMC and RJ-MCMC. $\{W_1, W_2, \dots\}$ denote waiting times and $\{t_1, t_2, \dots\}$ denote jumping times for the BD-MCMC algorithm, while the dots visualize the discrete RJ-MCMC samples. The right panel presents the estimated posterior probabilities of the trees which are the proportional to the waiting times (see, Cappé et al., 2003, Section 2.5).

Mohammadi and Wit (2015) show that this will be the case if the balance condition holds (Mohammadi and Wit, 2015, Appendix 1). It is easy to see that the proof of this condition for graphical models provided by Mohammadi and Wit (2015) is also valid for regression trees by noting that a tree is a special case of a graph. Let Θ_T be the tree-model space and $\Theta_{G_{max}}$ graph space in which G_{max} is a graph with the same number of nodes as a tree T with the maximum number of nodes. Clearly, $\Theta_T \subset \Theta_{G_{max}}$. This argument supports the implementation of the same birth/death mechanism as proposed by Mohammadi and Wit (2015) for trees: adding/removing a graph edge in the case of a birth/death event corresponds to adding/deleting a node of the tree (see Figure 2).

Following Mohammadi and Wit (2015), the birth and death rates are

$$B_n(T) = \frac{\pi(T^{+n})L(T^{+n})}{\pi(T)L(T)},$$

$$D_n(T) = \frac{\pi(T^{-n})L(T^{-n})}{\pi(T)L(T)},$$

in which T^{+n}/T^{-n} is a tree T with one more/less node n . As birth and death follow a Poisson processes, the time between two events has an exponential distribution and the probability of birth and death events are proportional to their rates.

We thus propose to replace the accept-reject mechanism of RJ-MCMC by a continuous time birth-death mechanism. In this new birth/death scheme the births and deaths occur at a higher rate when the components explain more of the data; a desirable feature not present in the RJ-MCMC approach. We believe that *combining* the BD-MCMC with the rotation mechanism will increase the efficiency of the traversal of

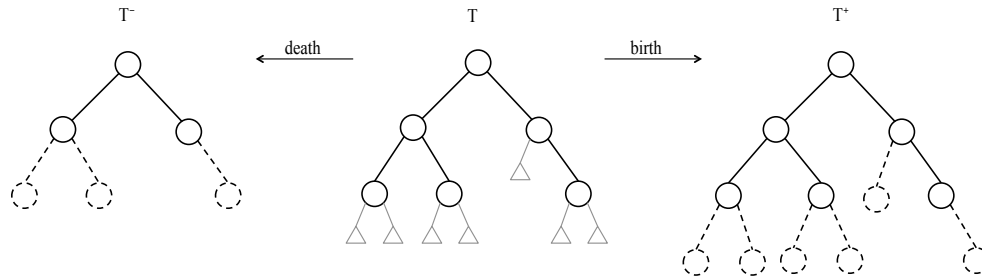


Figure 2: The birth-death mechanism for adding or deleting nodes of the tree.

model space (for a performance comparison of RJ-MCMC and BD-MCMC in graphical models see Section 4.1 of [Mohammadi and Wit, 2015](#)).

References

- Cappé, O., Robert, C., and Rydén, T. (2003). “Reversible Jump, Birth-and-Death and More General Continuous Time Markov Chain Monte Carlo Samplers.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(3): 679–700. [1](#), [2](#)
- Green, P. J. (1995). “Reversible jump Markov chain Monte Carlo computation and Bayesian model determination.” *Biometrika*, 82(4): 711–732. [1](#)
- Mohammadi, A., Abegaz Yazew, F., van den Heuvel, E., and Wit, E. C. (2017). “Bayesian modeling of Dupuytren disease using Gaussian copula graphical models.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(3): 1–17. To be published. [1](#)
- Mohammadi, A., Salehi-Rad, M., and Wit, E. (2013). “Using mixture of Gamma distributions for Bayesian analysis in an M/G/1 queue with optional second service.” *Computational Statistics*, 28(2): 683–700. [1](#)
- Mohammadi, A. and Wit, E. C. (2015). “Bayesian Structure Learning in Sparse Gaussian Graphical Models.” *Bayesian Analysis*, 10(1): 109–138. [1](#), [2](#), [3](#)
- (2016). “BDgraph: An R Package for Bayesian Structure Learning in Graphical Models.” *Journal of Statistical Software*. To be published. [1](#)
- Stephens, M. (2000). “Bayesian analysis of mixture models with an unknown number of components—an alternative to reversible jump methods.” *Annals of statistics*, 40–74. [1](#)