Posterior concentration for Bayesian Poisson regression trees

Lanpeng Ji [*1] and Xiaofan Peng [†2]

[1]School of Mathematics, University of Leeds, UK
[2]School of Mathematical Sciences, University of Electronic Science and Technology of China, China

Abstract

Regression trees have been one of the widely used non-parametric prediction methods. Trees are powerful, easily interpretable and thus preferred in many applications, including insurance ratemaking [4, 5]. However, traditional greedily grown trees are locally optimal, unstable and susceptible to over-fitting. To address some of these disadvantages, Bayesian framework was implemented to trees by introducing tree priors. Since the seminal works by Chipman, George, and McCulloch [1] and Denison, Mallick, and Smith [2], Bayesian tree-based models have received a lot of attention in both applied and theoretical research. In this talk, we focus on Bayesian CART modelling for insurance frequency data and some theoretical aspect of these studies. We review the main techniques used to prove the posterior concentration (around the true regression function) for Bayesian regression trees [3]. This study helps us understand why/when do Bayesian trees not over-fit and guide the choice of tree priors in practice. We discuss some results on the Bayesian Poisson regression trees in detail.

Keywords: Insurance ratemaking; Bayesian CART; posterior concentration; Poisson distribution.

References

[1] H. A. Chipman, E. I. George, and R. E. McCulloch (1998), "Bayesian CART model search." *Journal of the American Statistical Association*, vol. **93**(443), pp. 935-948.

[2] D. G. Denison, B. K. Mallick, and A. F. Smith (1998), "A Bayesian CART algorithm." *Biometrika*, vol. **85**(2), pp. 363-377.

[3] V. Rockova, S. Van der Pas (2020), "Posterior concentration for Bayesian regression trees and forests." *Annals of Statistics*, vol. **48**(4), pp. 2108-2131.

[*]E-mail address: l.ji@leeds.ac.uk
[†]E-mail address: xfpeng@uestc.edu.cn

[4] Y. Zhang, L. Ji, G. Aivaliotis, C. C. Taylor (2024) "Bayesian CART models for insurance claims frequency." *Insurance: Mathematics and Economics*, vol. **114**, pp. 108-131.

[5] Y. Zhang, L. Ji, G. Aivaliotis, C. C. Taylor (2025) "Bayesian CART models for aggregate claim modeling." *Insurance: Mathematics and Economics*, vol. **125**, pp. 103-136.