Efficient NUTS Sampler for Bayesian Conditional Transformation Models

Student: Aaron Lee   Supervisor: Prof. Dr. Nadja Klein

Abstract

Distributional regression models have recently gained interest in the statistical literature to model and predict complete distributions as functions of covariates rather than focusing on modelling and predicting the mean of a quantity of interest (see e.g., Klein; 2024, for a recent review). Such models are of high relevance when it comes to e.g. predicting extreme events or when interest is in specific quantiles of the distribution. One class of distributional models that is the focus of this project are conditional transformation models (Hothorn et al.; 2018), which provide a model directly for the conditional distribution function of a response variable Y as a function of the covariates $X = x$ via $F_{Y|X=x} = P(Y \leq y|X = x) = P(h(Y|x) \leq h(y|x)) = F_Z(h(y|x))$, where F_z is a reference distribution not depending on x and h is an unknown transformation function to be learned from the data. A key challenge is to learn the transformation function flexibly but monotonically increasing. The latter is required to obtain a valid cumulative distribution function. In this project we will follow the recent work of Carlan et al. (2024), who propose a Bayesian version of conditional transformation models using monotonically increasing B-splines. Unfortunately, the proposed implementation via Markov chain Monte Carlo simulations is unreliable, slow and inefficient. It is therefore the aim of this project to provide an efficient solution by making usage of new results on efficient parameterizations of involved penalty matrices (Bach and Klein; 2022) and an improved version of the No-U-Turn-Sampler (NUTS; Hoffman and Gelman; 2014) that updates all parameters in a single step rather than using blocks as in Carlan et al. (2024). It is part of the project to implement all routines and create an R or Python package along with testing on simulated data to ensure reproducibility. The project will also compare and evaluate to the results of Carlan et al. (2024).

References

Bach, P. and Klein, N. (2022). Anisotropic multidimensional smoothing using Bayesian tensor product P-splines.
URL: *https://arxiv.org/abs/2211.16218*

Carlan, M., Kneib, T. and Klein, N. (2024). Bayesian conditional transformation models, *Journal of the American Statistical Association* . In press. doi:10.1080/01621459.2023.2191820.

Hoffman, M. D. and Gelman, A. (2014). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo, *Journal of Machine Learning Research* **15**(47): 1593–1623.
URL: *http://jmlr.org/papers/v15/hoffman14a.html*

Hothorn, T., Möst, L. and Bühlmann, P. (2018). Most likely transformations, *Scandinavian Journal of Statistics* **45**(1): 110–134.

Klein, N. (2024). Distributional regression for data analysis, *Annual Review of Statistics and its Application* **11**: 321–346.