Bayesian Regularisation for Tail Index Regression

M.W. Lee[1], M. de Carvalho[1], D. Paulin[1], S. Pereira[2], R. Trigo[2]
C. Da Camara[2],

[1] School of Mathematics, University of Edinburgh, United Kingdom
[2] Faculdade de Ciências and CEAUL, Universidade de Lisboa, Portugal

E-mail for correspondence: `johnny.myungwon.lee@ed.ac.uk`

Abstract: We propose a novel approach for modelling the extreme values via a Bayesian regularisation that learns about a tail index regression framework. Our method is based on a conditional Pareto-type specification which is regularised with a shrinkage prior. To validate the performance of the proposed approach a battery of numerical experiments was conducted, and an illustration is given on extreme wildfires in Portugal.

Keywords: Bayesian Regularisation; ℓ_p-penalty; Tail Index Regression; Heavy-Tailed Response; Conditional Pareto-type Distribution.

1 Introduction

In this paper, we extend the Tail Index Regression (Wang & Tsai, 2009) to a Bayesian regularisation framework that characterises the extreme behaviour of a response variable that follows a conditional Pareto-type tail specification.

From a Bayesian perspective, each regression coefficient follows an independent and identically distributed shrinkage prior that behaves equivalently to the ℓ_p-type penalty regularisation. This aligns with the structure of the heavy-tailed distribution where certain covariates are determined as key factors of the extremeness. As a result, our approach entails a regularisation on a fully semiparametric framework by concentrating on learning about the regression coefficients that achieve a relatively sparse structure. Our contribution has important implications, particularly to the research in modelling extreme wildfires—such as the devastating 2017 Portugal wildfire (Turco et al., 2019)—and on the identification of their underlying drivers.

This paper was published as a part of the proceedings of the 37th International Workshop on Statistical Modelling (IWSM), Dortmund, Germany, 16–21 July 2023.

2 Bayesian Regularisation for Tail Index Regression

Our starting point for modelling is the following conditional Pareto-type tail specification that stems from Beirlant et al. (2004, Ch. 9):

$$P(Y > y \mid \boldsymbol{X} = \boldsymbol{x}) \equiv 1 - F(y \mid \boldsymbol{x}) = y^{-\alpha(\boldsymbol{x})}\mathcal{L}(y \mid \boldsymbol{x}). \tag{1}$$

Here, $\alpha(\boldsymbol{x}) = \exp(\boldsymbol{x}^T\boldsymbol{\beta})$, is a covariate-adjusted tail index, with the observation $\boldsymbol{x} = (x_1, \ldots, x_p)^T$ and regression coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$; in addition, $\mathcal{L}(y \mid \boldsymbol{x})$ is a covariate-adjusted slowly varying function, that is, $\mathcal{L}(yt \mid \boldsymbol{x})/\mathcal{L}(y \mid \boldsymbol{x}) \to 1$, as $y \to \infty$, for all $t > 0$. The specification in (1), allows for the heavy tail behaviour—as captured by the tail index—to depend on covariates. We follow Wang & Tsai (2009) and consider the Hall's (1982) class of covariate-adjusted slowly-varying functions given by,

$$\mathcal{L}(y \mid \boldsymbol{x}) = c_0(\boldsymbol{x}) + c_1(\boldsymbol{x})y^{-\theta(\boldsymbol{x})} + O(y^{\theta(\boldsymbol{x})}), \tag{2}$$

where $c_0(\boldsymbol{x}), c_1(\boldsymbol{x})$ and $\theta(\boldsymbol{x}) > 0$. Hence, $\mathcal{L}(y \mid \boldsymbol{x}) \to c_0(\boldsymbol{x})$ and $\partial\mathcal{L}(y \mid \boldsymbol{x})/\partial y \to 0$, as $y \to \infty$.

To regularise the above described tail index regression model, we resort to shrinkage priors. Given the space constraint we will focus on the Laplace prior, but other variants of the approach can be readily constructed by considering other shrinkage penalties, as illustrated in Fig. 1.

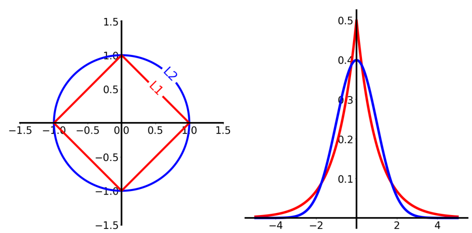


FIGURE 1. Comparison of the geometry of a unit ball induced by Laplace (red) and Normal (blue) priors depicting ℓ_1 and ℓ_2 penalty regularisation, respectively.

Concretely, in terms of the Bayesian Lasso (Park and Casella, 2008) version of our approach for (1), we learn about $\boldsymbol{\beta}$ from a random sample $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^n \sim F(\boldsymbol{x}, y)$. In Bayesian context, the posterior mode from a Laplace prior corresponds to maximising the constraints of the objective function with a ℓ_1 penalty. Thus, the sparsity-inducing regularisation shrinks the coefficients of less influential variables and results in the sparse structure distribution. Then, the resulting posterior density is,

$$p(\boldsymbol{\beta} \mid \{(\boldsymbol{x}_i, y_i)\}_{i=1}^n) \propto \exp\left\{1/\lambda \sum_{j=1}^p |\beta_j|\right\} L(\boldsymbol{\beta}) \tag{3}$$

where L is the approximated likelihood that follows from (2),

$$L(\boldsymbol{\beta}) \approx \prod_{i=1}^{n} f(y_i \mid \boldsymbol{x}_i) \approx \prod_{i=1}^{n} \alpha(\boldsymbol{x}_i)(y_i/u)^{-\alpha(\boldsymbol{x}_i)} y_i^{-1}, \tag{4}$$

for some large threshold u, with $y_i > u$, and where $f = \mathrm{d}F/\mathrm{d}y$. The priors for the regression parameters, $\boldsymbol{\beta}$ of tail index, $\alpha(\boldsymbol{x})$ are then defined as

$$\beta_j \mid \lambda \sim \text{Laplace}(\lambda), \quad \lambda \sim \text{Gamma}(a, b), \tag{5}$$

with $a, b > 0$, where an uninformative Gamma prior was chosen as the hyperprior for λ. Since the posterior has no closed-form expression, we resort to Markov Chain Monte Carlo (MCMC) methods for sampling.

3 Simulation Study

To assess the performance of the proposed method, we consider:
Scenario A: Conditional Pareto, i.e., $\mathcal{L}(y|\boldsymbol{x}) \propto 1$, with

$$\boldsymbol{\beta} = c(0.2, 0, 0.8, 0, 0, -0.1, 0, 0, 0, -0.4)^T.$$

Scenario B: Conditional Burr, i.e., $\mathcal{L}(y|\boldsymbol{x}) \propto (y^{-c(\boldsymbol{x})}+1)^{-2}$, where $\alpha(\boldsymbol{x}) = c(\boldsymbol{x})$ with

$$\boldsymbol{\beta} = c(0.1, 0.5, 0, 0, -0.9, -0.5, 0, 0.4, 0, 0)^T.$$

Scenario C: Conditional F, i.e., $\mathcal{L}(y|\boldsymbol{x}) \propto (y^{k_1/2-1}(k_1+k_2(\boldsymbol{x})y)^{-(k_1+k_2(\boldsymbol{x}))2})$, where $\alpha(\boldsymbol{x}) = k_2(\boldsymbol{x})/2$ with

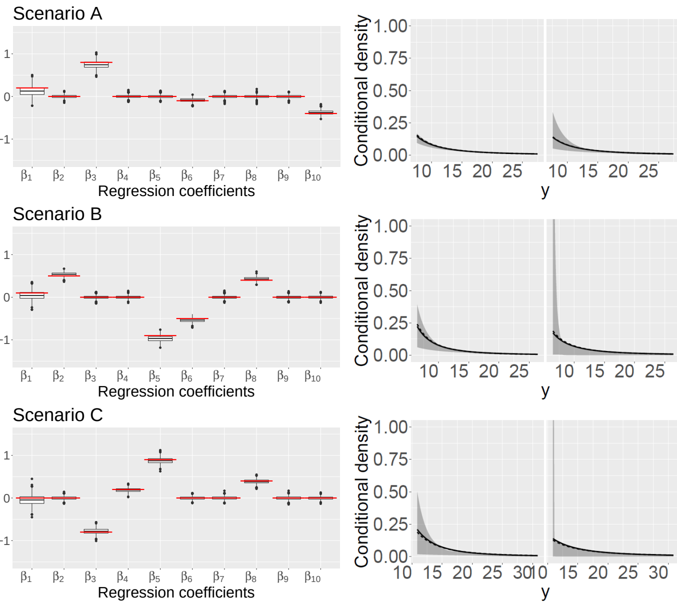
$$\boldsymbol{\beta} = c(0, 0, -0.8, 0.2, 0.9, 0, 0, 0.4, 0, 0)^T.$$

We retrieved the 90% quantile of 5000 random samples from (1) and repeated the study 250 times for Monte Carlo simulation. An uninformative Gamma prior, $\Gamma(0.1, 0.1)$ was employed, and in terms of MCMC we took $10,000$ burn-in iterations and collected $20,000$ samples. From Fig. 2 (top), it can be observed that each posterior mean approximates the true value well, hence suggesting a good performance of our method.

4 Real Data Application

We illustrate the proposed method on data of *Instituto Dom Luiz* that consists the daily burn area of forest fires between 1980 and 2019 in Portugal. We examine the following potential drivers for the same period of time: southerly flow (SF), westerly flow (WF), total flow (F), southerly shear vorticity (ZS), westerly shear vorticity (ZW), total shear vorticity (Z), and direction of flow (DF). We filtered 731 observations out of 14609 and their covariates were standardised. We used a Normal prior, $N(0, 100^2)$ for the intercepts and the same setup as in § 3. Fig. 2 (bottom left) suggests WF, DF and SF are the drivers for extreme forest fires. Fig. 2 (bottom right) depicts the corresponding randomised quantile residuals against the theoretical standard normal quantiles and it evidences a good fit.

Simulation



Real Data

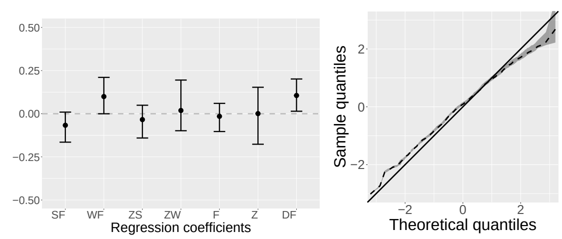


FIGURE 2. Results from simulation study (Top) and data illustration (Bottom).

References

Beirlant, J. , Goegebeur, Y., Segers, J. & Teugels, J. L. (2004). *Statistics of Extremes: Theory and Applications*. John Wiley & Sons.

Hall, P. (1982). On Some Simple Estimates of an Exponent of Regular Variation. *Journal of the Royal Statistical Society: Series B*, **44(1)**, 37 – 42.

Park, T. & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, **103(482)**, 681 – 686.

Turco, M., Jerez, S., Augusto, S., Tarín-Carrasco, P., Ratola, N., Jiménez-Guerrero, P., & Trigo, R. M. (2019). Climate Drivers of the 2017 Devastating Fires in Portugal. *Scientific reports*, **9(1)**, 13886.

Wang, H. & Tsai, C.L. (2009). Tail Index Regression. *Journal of the American Statistical Association*, **104(487)**, 1233 – 1240.