بِسْمِ اللهِ الرَّحْمٰنِ الرَّحِيمِ

For my parents.

Abstract

Regression analysis is undoubtedly an important tool to understand the relationship between one or more explanatory and independent variables of interest. In this thesis, we explore a novel methodology for fitting a wide-range of parametric and non-parametric regression models, called the I-prior methodology (Bergsma, 2018).

We assume that the regression function belongs to a reproducing kernel Hilbert or Kreĭn space of functions, and by doing so, it allows us to utilise the convenient topologies of these vector spaces. This is important for the derivation of the Fisher information of the regression function, which might be infinite-dimensional. Based on the principle of maximum entropy, an I-prior is an objective Gaussian process prior for the regression function with covariance function proportional to its Fisher information.

Our work focusses on the statistical methodology and computational aspects of fitting I-priors models. We examine a likelihood-based approach (direct optimisation and EM algorithm) for fitting I-prior models with normally distributed errors. The culmination of this work is the R package **iprior** (Jamil, 2017) which has been made publicly available on CRAN. The normal I-prior methodology is subsequently extended to fit categorical response models, achieved by "squashing" the regression functions through a probit sigmoid function. Estimation of I-probit models, as we call it, proves challenging due to the intractable integral involved in computing the likelihood. We overcome this difficulty by way of variational approximations. Finally, we turn to a fully Bayesian approach of variable selection using I-priors for linear models to tackle multicollinearity.

We illustrate the use of I-priors in various simulated and real-data examples. Our study advocates the I-prior methodology as being a simple, intuitive, and comparable alternative to similar leading state-of-the-art models.

Keywords: regression, model, reproducing kernel, Hilbert space, Kreĭn space, Fréchet derivative, Gâteaux derivative, Fisher information, binary, multinomial, Gaussian process, variational inference, empirical Bayes, EM algorithm, MCMC, truncated normal

http://phd.haziqj.ml • http://myphdcode.haziqj.ml

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent.

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 65,899 words.

I confirm that Chapters 2 and 3 were jointly co-authored with Dr. Wicher Bergsma, and I contributed 60% of these works. Chapter 4 was also jointly co-authored with Dr. Wicher Bergsma, in which I contributed 70% of this work, and we plan to submit this chapter for publication.

Acknowledgements

The fable of Aesop of the crow and the pitcher, as translated by George Fyler Townsend, reads as follows

> A crow perishing with thirst saw a pitcher and flew to it to discover, to his grief, that it contained so little water that he could not possibly get at it. He tried everything he could think of to reach the water, but all his efforts were in vain. At last, he collected as many stones as he could carry and dropped them one by one with his beak into the pitcher, until he brought the water within his reach so he could drink.

I am this fabled crow, but was most certainly not left to my own devices to reach the water. Wicher Bergsma, my teacher and friend, inspired and imparted all that he knew to me so I might collect the stones. I could not have asked for a more thoughtful, patient, and encouraging academic advisor. I do not recall an instance where he was unavailable when I needed advice. For this and much more, I am infinitely grateful.

Irini Moustaki, for all her wits and humour, was the person I turned to for affirmation and comfort. I remember most our walk to the Royal Statistical Society to attend my first Ordinary Meeting, and her saying to me "This is what my supervisor used to do with me, and now I am doing it with you". In that moment, I felt acceptance.

Quite literally, my PhD could not have been a success without our dedicated and ultra-supportive research administrators, Ian Marshall and Penny Montague. I am also thankful to the Statistics Department for providing a pleasurable and stimulating environment that allowed me to thrive. I have shared many enjoyable conversations with eager minds of the academic staff and students during my time there.

To my fellow inhabitants of the second floor office of Columbia House—Andy T. Y. Ho, Phoenix H. Feng, Ragvir Sabharwal, Tayfun Terzi, Alice Pignatelli-di-Cerchiara, Yan Qu, and Xioalin Zhu—I will miss our time together. Perhaps "inhabitant" might be too strong a word given that I'm rarely in, but rest assured, the sentiment is genuine.

To my parents, for whom this thesis is dedicated, and my family— I may have been able to see further by standing on the shoulders of giants, but were it not for you I would not have stood at all. For all you have given me, and continue to give, I will never be able to repay.

To my beloved wife Ireve Ameera and darling daughter Naqiyyah Afrah who was born halfway through my PhD—you both make me whole and my days are filled with such joy by your presence. Thank you especially for linguistic and aesthetic help, and for being a forbearing audience for scores of my practice presentations. Toddlers make for a tougher crowd than accomplished statisticians, for sure.