Bayesian smoothing with second order random walk model: An detailed overview and comparison

Ziang Zhang

aguero.zhang@mail.utoronto.ca

Supervisor(s): James Stafford, Patrick Brown

Department of Statistical Sciences
University of Toronto

June 2021

Smoothing methods are often used when there is little information on the functional structures of some covariate effects. Common smoothing methods are mostly based on the application of spline functions, such as regression splines, smoothing splines and penalizing regression splines (P-splines). The main challenging of smoothing is to provide enough flexibility so that the functional form of covariate effect can be accurately inferred without over-fitting the observed data. Typically this trade off is controlled by a smoothing parameter σ, which penalizes the wiggliness of inferred function. In typical frequentist method, the smoothing parameter σ is either taken as fixed value input by the users, or substituted by an optimal value selected from procedure such as cross validation. Therefore, how to take into account the uncertainty with the unknown hyper-parameter increases the difficulty of frequentist smoothing methods. On the other hand, the hyper-parameter σ will be assigned with a prior distribution in Bayesian smoothing methods, and hence any uncertainty involved with that parameter will be taken into account for the inference. Furthermore, the development of approximate Bayesian inference methods such as Rue et al. (2009) enables Bayesian smoothing to be implemented in a computationally convenient way. Hence, application of Bayesian smoothing method can be advantageous in a lot of settings.

Based on the well known connection between smoothing splines and integrated Wiener processes (Wahba, 1978), Lindgren and Rue (2008) developed a Bayesian smoothing method by assigning second order random walk priors (RW2) to the unknown true effect functions. Their method can be viewed as a discrete approximation to the second integrated Wiener process, using a finite element method called Galerkin approximation. The hyper-parameter σ represents the standard deviation parameter of the second derivative of the covariate effect function, and will be assigned with a proper prior distribution. Because of the use Galerkin approximation, the resulting prior distribution for the effect function will have a sparse precision matrix, and hence will be computationally efficient if used together with approximate Bayesian inference method such as Integrated Nested Laplace Approximation (INLA) (Rue et al., 2009). Both theoretical results and simulation results have been demonstrated for their Galerkin approximation methods in their original paper (Lindgren and Rue, 2008).

In this report, we will give a thorough overview for Bayesian smoothing methods, and a detailed description of the RW2 method proposed in Lindgren and Rue (2008). Through extensive simulation studies, we will demonstrate both the advantages and the disadvantages of the RW2 method compared to existing methods such as Bayesian regression splines and general Bayesian P-splines (Lang and Brezger, 2004). Finally, we will discuss the potential extensions and generalizations of the RW2 method.

References

Lang, S. and Brezger, A. (2004). Bayesian p-splines. *Journal of Computational and Graphical Statistics*, 13(1):183–212.

Lindgren, F. and Rue, H. (2008). On the second-order random walk model for irregular locations. *Scandinavian Journal of Statistics*, 35(4):691–700.

Rue, H., Martino, S., and Chopin, N. (2009). Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.

Wahba, G. (1978). Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(3):364–372.