Journal of Statistical Planning and Inference 165 (2015) 78–90

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Bayesian semiparametric hierarchical empirical likelihood spatial models

Aaron T. Porter [a,*], Scott H. Holan [b], Christopher K. Wikle [b]

[a] *Department of Applied Mathematics and Statistics, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, United States*

[b] *Department of Statistics, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65211-6100, United States*

ARTICLE INFO

Article history:
Received 6 May 2014
Received in revised form 5 December 2014
Accepted 10 April 2015
Available online 24 April 2015

Keywords:
Conditional autoregressive model
Fay–Herriot model
Kriging
Random field
Small area estimation

ABSTRACT

We introduce a general hierarchical Bayesian framework that incorporates a flexible nonparametric data model specification through the use of empirical likelihood methodology, which we term semiparametric hierarchical empirical likelihood (SHEL) models. Although general dependence structures can be readily accommodated, we focus on spatial modeling, a relatively underdeveloped area in the empirical likelihood literature. Importantly, the models we develop naturally accommodate spatial association on irregular lattices and irregularly spaced point-referenced data. We illustrate our proposed framework by means of a simulation study and through three real data examples. First, we develop a spatial Fay–Herriot model in the SHEL framework and apply it to the problem of small area estimation in the American Community Survey. Next, we illustrate the SHEL model in the context of areal data (on an irregular lattice) through the North Carolina sudden infant death syndrome (SIDS) dataset. Finally, we analyze a point-referenced dataset from the North American Breeding Bird Survey that considers dove counts for the state of Missouri. In all cases, we demonstrate superior performance of our model, in terms of mean squared prediction error, over standard parametric analyses.

1. Introduction

The empirical likelihood (EL) dates back to the seminal work of Owen (1988) and has become increasingly popular in recent years, as a result of Owen (2001), which placed many of the fundamental concepts in a single text. Early work by Qin and Lawless (1994) greatly expanded the use of EL by placing it in the context of estimating equations. Kolaczyk (1994) derived general conditions for the use of estimating equations for the EL that are applicable to many types of linear, nonlinear, and semiparametric models. Many EL-type estimators have since been derived, known as Generalized Empirical Likelihood (GEL) estimators. Newey and Smith (2004) provides an excellent overview of these estimators and their higher order properties.

Lazar (2003) provides evidence, by means of a simulation study, that the EL framework is appropriate for Bayesian inference. Making use of a result from Monahan and Boos (1992) that yields conditions by which a likelihood can be determined suitable for Bayesian inference, this paper initiated Bayesian research on EL and GEL estimators. Schennach (2005) derived a Bayesian GEL estimator by means of nonparametric priors and further extended their approach in Schennach (2007). Fang and Mukerjee (2006) derived the asymptotic frequentist coverage properties of the Bayesian

* Corresponding author.
E-mail address: aporter@mines.edu (A.T. Porter).

http://dx.doi.org/10.1016/j.jspi.2015.04.002

credible intervals for the mean parameters of a wide class of EL-type likelihoods, and demonstrated undercoverage for credible intervals for parametric means generated by GEL estimators. Additional work comparing the properties of credible intervals for specific types of EL-type likelihoods can be found in Chang and Mukerjee (2008). In particular, this work demonstrates favorable coverage rates for the traditional EL of Owen (1988).

Bayesian hierarchical modeling (BHM) has become an expansive field. When modeling complex stochastic phenomena within the BHM framework, typically at least three levels of model hierarchy are considered, which are the data model, process model, and parameter model (Berliner, 1996; Wikle, 2003). Subsequently, modeling typically proceeds by selecting parametric distributions for each stage of the hierarchy. As demonstrated in Cressie and Wikle (2011), this framework advantageously also allows for scientifically motivated process models to be utilized at the latent stage. One aspect of this approach is that model implementation typically requires selection of an appropriate data distribution (likelihood) for the observations.

Our approach extends the general applicability of BHMs by broadly placing them in the context of the empirical likelihood. The model we propose can be viewed as a semiparametric hierarchical empirical likelihood (SHEL) model and utilizes either EL estimators or GEL estimators at the data stage of the model hierarchy. Parametric process models can then be utilized to handle the potentially complex underlying dependence structures. By placing the EL in the context of Bayesian hierarchical modeling, we alleviate the issues of modeling the dependency in the observations, which is often difficult to handle in the usual observation-driven EL framework and generally utilizes restrictive blocking arguments. Specifically, we expand the BHM framework to allow empirical data models, rather than requiring the user to select a parametric structure for the data.

Hierarchical approaches to empirical likelihood have been recently considered, but still remain largely underdeveloped, with no general framework to date. Chaudhuri and Ghosh (2011) proposed using the EL in a semiparametric hierarchical nested error regression model for small area estimations (SAE). The model they developed extends the traditional Fay–Herriot (FH) model (Fay and Herriot, 1979) to the EL framework. Although Chaudhuri and Ghosh (2011) demonstrate good model performance, their implementation utilized informative priors for some of the model parameters, and they noted sensitivity to these specifications. The general approach they propose allows for both semiparametric and nonparametric specifications of the model for the superpopulation mean, with the nonparametric specification relying on a Bayesian nonparametric formulation (i.e., a Dirichlet process mixture with Gaussian base measure). We pursue a more complete development of EL in the context of BHMs. The model we propose here is of independent interest and readily allows for various other hierarchical and/or dependence structures, such as temporal and/or spatio-temporal dependencies. However, for the sake of brevity, subsequent exposition focuses on spatially correlated data.

Based on blocking arguments originally developed for time series by Kitamura (1997), Nordman and Caragea (2008) developed a point referenced spatial model in the frequentist EL framework that considers variogram fitting for data collected on a regular grid, and assumes stationarity. Utilizing a similar blocking argument, Nordman (2008) considered an observation-driven model for spatial data on a regular lattice using the EL framework that does not require stationarity. To the best of our knowledge, hierarchical models for spatial data on an irregular lattice that explicitly account for the underlying spatial structure in the data do not exist in the current literature. A recent advancement in the spatial EL literature is Bandyopadhyay et al. (2012), in which irregularly spaced spatial data is modeled using frequency domain techniques. Their framework greatly expands EL methodology for point referenced spatial data but is based on different assumptions than those presented herein and does not immediately extend to the lattice case, where distances are not uniquely defined.

The structure of this paper is as follows. Section 2 develops methodology that will be needed for the general specification of the SHEL model. Section 3 discusses technical details related to the Bayesian estimation of the SHEL model, and provides the Markov chain Monte Carlo (MCMC) algorithm that we propose. Section 4 presents three case studies: the FH model for SAE in the context of the American Community Survey (ACS), the North Carolina SIDS data (areal data), and a point referenced dataset from the North American Breeding Bird Survey that considers dove counts for the state of Missouri. Section 5 provides concluding discussion. For ease of exposition, the results of two simulation studies, illustrating the effectiveness of our approach, are left to an Appendix.

2. Spatial SHEL models

2.1. The SHEL framework

Let $\mathbf{Z}$ be an n_Z-dimensional vector of observations, $\mathbf{Y}$ be an n_Y-dimensional vector corresponding to an unobserved process, and $\boldsymbol{\xi}$ be a set of parameters related to both the data model and process model. Here, $\mathbf{Z}$ and $\mathbf{Y}$ do not need to be of the same dimension. For example, the observations could be mapped to the unobserved process through a matrix that accounts for change-of-support or aggregation (Wikle and Berliner, 2005). However, for ease of notation, we assume $n_Z = n_Y \equiv n$, unless specified otherwise. Further, let $[\mathbf{Z}|\mathbf{Y}]$ denote the conditional distribution of $\mathbf{Z}$ given $\mathbf{Y}$ and $[\mathbf{Y}]$ denote the marginal distribution of $\mathbf{Y}$. We propose a general setup for the SHEL framework that considers a data model $[\mathbf{Z}|\mathbf{Y}, \boldsymbol{\xi}_D]$, process model $[\mathbf{Y}|\boldsymbol{\xi}_P]$, and parameter model $[\boldsymbol{\xi}] = [\boldsymbol{\xi}_D, \boldsymbol{\xi}_P]$, with $[\boldsymbol{\xi}_D]$ being the joint prior distribution of the data model parameters and $[\boldsymbol{\xi}_P]$ being the joint prior distribution of the process model parameters. The framework we propose here is not unique to spatial data, and any process model in which $[\mathbf{Y}, \boldsymbol{\xi}]$ is proper can be utilized.