Spatial Statistics 24 (2018) 54–65

Contents lists available at ScienceDirect

Spatial Statistics

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

On the formal specification of sum-zero constrained intrinsic conditional autoregressive models

Matthew J. Keefe, Marco A.R. Ferreira, Christopher T. Franck *

Department of Statistics, Virginia Tech, Blacksburg, VA 24061, United States

ARTICLE INFO

Article history:
Received 26 May 2017
Accepted 28 March 2018
Available online 8 April 2018

Keywords:
Areal data
Conditional autoregressive
Intrinsic autoregressive
Model selection
Spatial statistics

ABSTRACT

We propose a formal specification for sum-zero constrained intrinsic conditional autoregressive (ICAR) models. Our specification first projects a vector of proper conditional autoregressive spatial random effects onto a subspace where the projected vector is constrained to sum to zero, and after that takes the limit when the proper conditional autoregressive model approaches the ICAR model. As a result, we show that the sum-zero constrained ICAR model has a singular Gaussian distribution with zero mean vector and a unique covariance matrix. Previously, sum-zero constraints have typically been imposed on the vector of spatial random effects in ICAR models within a Markov chain Monte Carlo (MCMC) algorithm in what is known as centering-on-the-fly. This mathematically informal way to impose the sum-zero constraint obscures the actual joint density of the spatial random effects. By contrast, the present work elucidates a unique distribution for ICAR random effects. The explicit expressions for the resulting unique covariance matrix and density function are useful for the development of Bayesian methodology in spatial statistics which will be useful to practitioners. We illustrate the practical relevance of our results by using Bayesian model selection to jointly assess both spatial dependence and fixed effects.

---

* Corresponding author.
E-mail address: chfranck@vt.edu (C.T. Franck).

https://doi.org/10.1016/j.spasta.2018.03.007

1. Introduction

Intrinsic conditional autoregressive (ICAR) models are often used as prior distributions for spatial random effects in Bayesian hierarchical models. These autoregressive models are useful in applications where a neighbor-based notion of proximity is considered, such as disease mapping (Clayton and Kaldor, 1987; Bell and Broemeling, 2000; Moraga and Lawson, 2012; Goicoa et al., 2016), image restoration (Besag et al., 1991), complex survey data (Mercer et al., 2015), and neuroimaging (Liu et al., 2016). Because the ICAR model is not a proper distribution, practitioners typically constrain the sum of the spatial effects to equal zero via "centering-on-the-fly", or re-centering the vector of sampled spatial random effects around its mean after each iteration of the Markov chain Monte Carlo (MCMC) algorithm to make the ICAR model proper. While this approach works well in practice, centering-on-the-fly obscures the actual joint density of the spatial random effects. Unfortunately, the lack of knowledge about the exact joint density of the spatial random effects has inhibited the development of formal methodology, such as objective Bayesian analysis and Bayesian model selection for ICAR-based hierarchical models. Our main contribution here is three-fold. First, we propose a procedure to formally specify sum-zero constrained ICAR models by first imposing the sum-zero constraint on a proper conditional autoregressive model and then by taking the limit to obtain a unique ICAR model. Second, we demonstrate that the distribution of the resulting sum-zero constrained ICAR model is a unique singular Gaussian distribution. Third, the practical contribution of this work opens the door to the development of formal Bayesian procedures for hierarchical models with sum-zero constrained ICAR spatial random effects.

To accomplish these goals, we propose to project the vector of proper conditional autoregressive spatial random effects onto a subspace where the projected effects are constrained to sum to zero and, after that, to take the limit when the proper conditional autoregressive model approaches the ICAR. As a consequence, the prior distribution of the spatial random effects is shown to be in the class of singular Gaussian distributions. Moreover, singular Gaussian distributions have well-established theoretical results (Siotani et al., 1985; Anderson, 2003; Ferreira et al., 2011). In particular, the mean vector and covariance matrices of these distributions implicitly encode constraints. Hence, the consequences of imposing the sum-zero constraint can be elucidated through careful linear algebra. We note that singular Gaussian distributions have been successfully used in practice, for example in Bayesian multiscale multiple imputation (Holan et al., 2010) and dynamic multiscale spatiotemporal models (Ferreira et al., 2010, 2011). For completeness, we state the results of the singular Gaussian distribution used in this paper in the Appendix. We note that these are not conventions but instead are mathematical results. We use these results for the singular Gaussian distribution to formally specify the sum-zero constrained ICAR model.

There are numerous practical implications of the results we present. For example, in a related paper we have used our formal specification to develop objective Bayesian methodology for Bayesian hierarchical models with ICAR spatial random effects (Keefe et al., 2018). As another example, in this paper we use our formal ICAR specification to develop Bayesian model selection for models that account for spatial dependence.

Finally, we note the consequences of not imposing the sum-zero constraint. The case when the sum-zero constraint is not imposed has been thoroughly studied by Lavine and Hodges (2012). In that work, Lavine and Hodges (2012) have demonstrated that different approaches to obtain an ICAR model as the limit of a proper conditional autoregressive model lead to differing limiting likelihood ratios. We show that by first projecting the spatial random effects onto the sum-zero constrained subspace prior to taking the limit, the resulting sum-zero constrained ICAR model is unique for a broad class of covariance matrices of the proper conditional autoregressive model. Our uniqueness result is fundamentally useful to advance statistical research for a class of models widely used in practice.

The remainder of the paper is organized as follows. In Section 2, we review the details of ICAR models. In Section 3, we describe a formal procedure to define sum-zero constrained ICAR models. Then, in Section 4 we provide a few examples used by Lavine and Hodges (2012) to illustrate our result. To demonstrate the practical usefulness of our result, we describe an application in Bayesian model selection for spatial dependence and covariates with a case study in Section 5. Finally, we provide conclusions and discussion of additional practical implications of our work in Section 6.