Journal of the Korean Statistical Society 45 (2016) 276–284

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

An approximate likelihood function of spatial correlation parameters

Yongku Kim, Dal Ho Kim *

Department of Statistics, Kyungpook National University, South Korea

A R T I C L E   I N F O

Article history:
Received 25 January 2015
Accepted 5 November 2015
Available online 4 December 2015

AMS 2000 subject classifications:
primary 62F15
secondary 62B30

Keywords:
Approximated likelihood
Correlogram
Spatial likelihood
Spatial statistics

A B S T R A C T

Even under assumption of normality, likelihood-based inferences are often difficult for large and irregularly spaced spatial datasets. Exact calculations of the likelihood for a Gaussian spatial process observed in n locations require $O(n^3)$ operations. Instead of Whittle's approximation to the Gaussian log likelihood for large spatial datasets, this paper introduces an approximated likelihood function of spatial parameters based on the correlogram, which involves no calculation of determinants and is computationally feasible. The proposed likelihood approximation method for spatial parameter is applied to the estimation of the spatial structure of changes in the average summer temperature based on 30 years of data by using an regional climate model (RCM) with a particular global climate model (GCM) boundary condition. The results verify the benefits and the performance of the proposed method.

1. Introduction

Estimating the spatial covariance structure of huge datasets is fundamental to geostatistics but is also difficult because of heavy computational burdens. Under the assumption of normality, calculating the exact likelihood function requires $O(n^3)$ operations, where n is the number of observations. Based on a complete regular lattice, the likelihood function can be computed with fewer calculations by using spectral methods (Dahlhaus & Küsch, 1987; Guyon, 1982; Stein, 1995, 1999; Whittle, 1954). However, the methods based on the likelihood approximation proposed by Whittle (1954) cannot be applied to data that are irregularly spaced or are not on a complete regular lattice. Because any joint density can be written as the product of conditional densities based on some ordering of observations, another way to reduce computational burdens is to condition only particular neighbors of observations when computing the conditional density. Stein, Chi, and Welty (2004) adopt this approach to approximate the restricted likelihood function and demonstrate how the estimation of equations allows for the evaluation of the efficacy of the resulting approximation.

Recently, Fuentes (2007) considered a version of Whittles approximation to the Gaussian log likelihood for spatial regular lattices with missing values and for irregularly spaced datasets. Alternatively, Kaufman, Schervish, and Nychika (2008) proposed the method of covariance tapering to approximate the likelihood in large spatial dataset. In their approach, covariance matrices are "tapered", or multiplied element wise by a sparse correlation matrix. The resulting matrices can then be manipulated using efficient sparse matrix algorithms. Sang and Huang (2012) introduced a full scale approximation to the covariance function of a spatial process that facilitates efficient computation with very large spatial datasets. Another

* Corresponding author.
E-mail address: dalkim@knu.ac.kr (D.H. Kim).

http://dx.doi.org/10.1016/j.jkss.2015.11.003
1226-3192/

possible approach is to consider a multi-resolution Gaussian Markov random field and take advantage of the fact that the basis is organized as a lattice.

In general, Whittle's approximation to the Gaussian process can be applied to lattice data. This paper introduces the approximated likelihood function of spatial parameters based on a correlogram. The proposed method does not involve the calculation of determinants with huge dimensions and is attractive in that it is very simple, computationally efficient, and fast in comparison to the exact likelihood approach. Compared to other approximation approaches (e.g. Stein et al., 2004, Kaufman et al., 2008, Fuentes, 2007 and Sang & Huang, 2012), the proposed method does not require tapering (gives less weight to boundary observations) before approximation and it is straightforward, flexible and provides enormous computational benefits for both complete or incomplete and regularly or irregularly spaced observations. The proposed likelihood approximation method for spatial parameter is applied to the estimation of the spatial structure of changes in the average summer temperature based on 30 years of data by using an regional climate model (RCM) with a particular global climate model (GCM) boundary condition.

2. Correlogram and its properties

A correlogram is a function showing correlations between sample points separated by the distance d (see Friendly, 2002). The correlation usually decreases with an increase in the distance until it reaches zero. A correlogram is estimated using the equation:

$$\rho(d) = \frac{\sum_{i<j} Z_i Z_j - N_d M_d^2}{N_d s_d^2}, \tag{1}$$

where Z_i and Z_j are stochastic variables in two locations separated by the lag distance d; the summation is performed over all pairs of samples separated by the distance d; N_d is the number of pairs of samples separated by the distance d; and M_d and s_d are the mean and standard deviation of samples separated by the distance d (each sample is weighted by the number of pairs of samples in which it is included). This is an omnidirectional correlogram, and here the word "omnidirectional" implies that there is no need to address the direction of the lag d. If a correlogram depends on the direction, then the spatial pattern is called anisotropic. If no anisotropy is detected, then it is possible to use the omnidirectional correlogram. Other measures of spatial dependence are the covariance function and the variogram (= semivariogram). The correlogram, the covariance function, and the variogram are related.

Assume that $Z(s)$, $s \in \Re^k$ follows a Gaussian process with the correlation function $r_{\boldsymbol{\theta}}(\cdot)$. For any two variables $Z(s)$ and $Z(s')$ with the distance $d = |s - s'|$ for $s, s' \in \Re^k$, there exists a function $\mathcal{C}$ satisfying

$$r_{\boldsymbol{\theta}}(d) = \mathcal{C}_{F_d}(Z(s), Z(s'); d),$$

where F_d is the bivariate normal distribution function induced by the Gaussian random field. In general, it is assumed that the function $\mathcal{C}$ can be parametrically modeled. For example,

$$r_{\xi,\nu}(s, s') = \frac{(d(s, s')/\xi)^{\nu}}{2^{\nu-1}\Gamma(\nu)} \kappa_{\nu}(d(s, s')/\xi), \tag{2}$$

which is the Matérn correlation function (Matérn, 1986) with parameters $\boldsymbol{\theta} = (\xi, \nu)$, evaluated at $d(s, s')$. Because the Matérn correlation function is positive definite in $\Re^2$, $r_{\xi,\nu}(d)$ is a valid periodic correlation function on $[0, 1]$ (Yaglom, 1987). Then the corresponding covariance matrix of $Z(s)$ can be expressed as $\boldsymbol{\Sigma}(s, s') = \sigma^2 r_{\xi,\nu}(s, s')$. Note that the distance space $\mathcal{D}$ is finite in general and that this approach is computationally efficient, particularly when the spatial domain is defined on grid points. Also note that $r_{\boldsymbol{\theta}}(d)$ is a correlation function and $\rho(d)$ is an empirical correlogram evaluated at d.

Let $\mathcal{P}_d$ be a set of pairs of $Z(s)$ and $Z(s')$ with $d(=|s - s'|)$. By the stationarity of the Gaussian random field Z, any element of $\mathcal{P}_d$ can be considered as a random sample from F_d. Let $\hat{F}_d$ be the corresponding empirical distribution of samples of size n_d from F_d. For a given value of σ^2,

$$\rho(d) = \mathcal{C}_{\hat{F}_d}(Z(s), Z(s'); d).$$

Then it can be easily checked for given σ^2 that

$$\rho(d) = \mathcal{C}_{\hat{F}_d}(Z(s), Z(s'); d) \to \mathcal{C}_{F_d}(Z(s), Z(s'); d) = r_{\boldsymbol{\theta}}(d) \quad \text{as } n_d \to \infty.$$

It is well known that $\hat{F}_d \to F_d$ almost surely and uniformly in $(Z(s), Z(s'))$. In particular, the central limit theorem of $\rho(d)$ makes the bootstrap approximation more accurate under the weak condition. Therefore, this leads to

$$\rho(d) \sim N(r_{\boldsymbol{\theta}}(d), \sigma_d^2) \tag{3}$$

or

$$\log\sqrt{\frac{1+\rho(d)}{1-\rho(d)}} \sim N\left(\log\sqrt{\frac{1+r_{\boldsymbol{\theta}}(d)}{1-r_{\boldsymbol{\theta}}(d)}}, \frac{1}{n_d - 3}\right), \tag{4}$$