Statistics and Probability Letters 129 (2017) 182–188

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Numerical instability of calculating inverse of spatial covariance matrices

Chae Young Lim [a], Chien-Hung Chen [b], Wei-Ying Wu [c,*]

[a] Department of Statistics, Seoul National University, Republic of Korea
[b] Risk Consulting, KPMG in Taiwan, Taiwan
[c] Department of Applied Mathematics, National Dong Hwa University, Taiwan

ARTICLE INFO

Article history:
Received 27 January 2017
Received in revised form 23 May 2017
Accepted 25 May 2017
Available online 7 June 2017

Keywords:
MLE
Matérn covariance models
Ill-conditioned

ABSTRACT

Computing an inverse of a covariance matrix is a common computational component in Statistics. For example, Gaussian likelihood function includes the inverse of a covariance matrix. For the computation of the inverse of a spatial covariance matrix, numerically unstable results can arise when the observation locations are getting denser. In this paper, we investigate when computational instability in calculating the inverse of a spatial covariance matrix makes maximum likelihood estimator unreasonable for a Matérn covariance model. Also, some possible approaches to relax such computational instability are discussed.

1. Introduction

Computation of the inverse of a covariance matrix is required very often in Statistics. For example, Gaussian likelihood function includes the inverse of a covariance matrix. Spatial prediction called Kriging requires computation of the inverse of a spatial covariance matrix. For a (stationary) spatial covariance matrix which depends on the distance between neighboring observations, computation of an inverse of a spatial covariance matrix can be unstable when the observed locations are dense. Such instability could result in a negative eigenvalue or numerical singularity for a spatial covariance matrix.

In numerical analysis or computer science, the condition number of a matrix is used as a measure of numerical stability for a matrix. For a symmetric matrix M, the L_2-norm condition number is defined by $\kappa_2(M) = \frac{\text{maximum eigenvalue of } M}{\text{minimum eigenvalue of } M}$. The matrix is numerically invertible if its condition number is finite. It is *ill-conditioned* if its condition number is finite but high (greater than 10^{12} in Andrianakis and Challenor, 2012 or 10^3 in Won et al., 2014) and *well-conditioned* if its condition number is moderate. In the LAPACK, a standard software library for numerical linear algebra used in MATLAB (The MathWorks), a reciprocal condition number using L_1-norm is used to determine whether a matrix is numerically singular. The L_1-norm condition number, $\kappa_1(M)$, of a matrix M is defined as $\kappa_1(M) = \|M\|_1\|M^{-1}\|_1$. Specifically, if $(\kappa_1(M))^{-1} < 10\,\varepsilon$, where $\varepsilon\,(= 2^{-52} \approx 2.22 \times 10^{-16})$ is the floating-point relative accuracy, the matrix M is numerically singular. The same accuracy is used in R and GSTAT computer code (Peng and Wu, 2014). The condition number depends on the matrix norm used, but all condition numbers are equivalent in the sense that one can be bounded below/above by a constant multiple of the other (e.g. Horn and Johnson, 1990).

In this paper, we investigate what components or situations bring computational instability in calculating the inverse of a spatial covariance matrix, in particular, for a Matérn covariance model and how computational instability affects a maximum likelihood estimator (MLE) for spatial data. Such a numerical problem could also happen when the spatial prediction called Kriging is applied. The results regarding spatial prediction are not presented in the paper due to the limited space.

* Corresponding author.
E-mail address: wuweiying1011@mail.ndhu.edu.tw (W.-Y. Wu).

http://dx.doi.org/10.1016/j.spl.2017.05.019
0167-7152/

Table 1
The condition number (κ_2), bias (Bias) and standard deviation (STD) of $\hat{\sigma}^2_{\alpha_*}\alpha_*^{2\nu}$, where $\hat{\sigma}^2_{\alpha_*}$ given in (1) from 20 datasets at various δ and ν. The observation domain is [0,1] and α^* is set to 0.1.

δ	0.002			0.001			0.0005		
ν	κ_2	Bias	STD	κ_2	Bias	STD	κ_2	Bias	STD
0.3	1.09E+05	0.00	0.16	3.29E+05	−0.03	0.13	9.97E+05	−0.02	0.10
0.5	4.85E+06	−0.04	0.19	1.94E+07	−0.01	0.15	7.74E+07	0.02	0.09
1.2	2.56E+12	−0.19	0.27	2.70E+13	−0.16	0.22	2.85E+14	−0.07	0.10
1.5	7.55E+14	−0.21	0.32	2.85E+16	0.15	0.27	2.78E+19	−159.35	283.85

All numerical study was conducted by MATLAB 2014 with the double floating-point precision on Intel(R) Core(TM) i7-3770@3.40 GHz and 16.0 GB memory. The paper is organized as follows. In Section 2, we present how the inverse of a covariance matrix with a large condition number leads to unstable computation on estimation. In Section 3, we discuss factors that influence the condition number of a spatial covariance matrix under the Matérn covariance model. Simulation study is provided to support our findings. Some approaches used to reduce such computation instability are discussed in Section 4. Section 5 summarizes our work.

2. Computational instability on MLE for spatial data

We first investigate impact of a large condition number of a spatial covariance matrix on estimation via numerical study. Also, we investigate when a spatial covariance matrix has a large condition number. We consider a mean zero Gaussian random field with a Matérn covariance model, which is defined as $K(|\boldsymbol{s}|) = \frac{\sigma^2 2^{1-\nu}}{\Gamma(\nu)}(\alpha|\boldsymbol{s}|)^\nu \mathcal{K}_\nu(\alpha|\boldsymbol{s}|)$, where $\boldsymbol{s} \in \mathbf{R^d}$, $\mathcal{K}_\nu$ is a modified Bessel function of the second kind with order ν, which is called a smoothing parameter, α is the range parameter and $|\cdot|$ is the d-dimensional Euclidean norm.

Suppose that the data are observed at $\{\boldsymbol{s}_1, \dots, \boldsymbol{s}_n\}$ so that the data vector is $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \dots, Z(\boldsymbol{s}_n))^t$, where t is the transpose of a matrix. When ν is known, for any given $\alpha_* > 0$, which is not necessarily the true value, the MLE of the variance parameter, σ^2, can be written as

$$\hat{\sigma}^2_{\alpha_*} = \boldsymbol{Z}^t \Sigma^{-1}_{\alpha_*,\nu}\boldsymbol{Z}/n, \tag{1}$$

where $\Sigma_{\alpha_*,\nu}$ is the covariance matrix constructed by the Matérn covariance model with the parameters $\sigma^2 = 1, \alpha = \alpha^*$ and ν.

The asymptotic properties of $\hat{\sigma}^2_{\alpha_*}$ can be found in the literature under infill asymptotics. Infill asymptotics or fixed domain asymptotics is the asymptotic framework when the observation domain is bounded while the number of data points within the observation domain is increasing (see e.g. Stein, 1999; Zhang, 2004; Kaufman et al., 2008; Du et al., 2009). The estimator $\hat{\sigma}^2_{\alpha_*}\alpha_*^{2\nu}$ of the microergodic parameter, $\sigma^2\alpha^{2\nu}$, is asymptotically consistent for any given $\alpha^* > 0$ but obvious bias could be observed for the finite samples (e.g. Kaufman and Shaby, 2013) while one cannot estimate σ^2, α and ν, separately.

To see how numerical instability affects estimation of spatial data, numerical investigation is done first. We generated the data on a grid within [0, 1] with the parameter values $(\sigma^2, \alpha) = (2, 1.5)$ and $\nu \in \{0.3, 0.5, 1.2, 1.5\}$. 20 datasets were generated for each parameter setting under different grid sizes, δ. We then estimate $\sigma^2\alpha^{2\nu}$ for each dataset. Bias and standard deviation were obtained.

Table 1 shows the condition number (κ_2), the bias (Bias) and standard deviation (STD) of the estimates of $\sigma^2\alpha^{2\nu}$ for various ν values when $\alpha_* = 0.1$. When $\nu = 0.3, 0.5$ and 1.2, $\hat{\sigma}^2_{\alpha_*}\alpha_*^{2\nu}$ performs well. However, when ν is 1.5, we observe a larger bias and it increases as the grid size (δ) is getting smaller (i.e. the sample size is getting larger). Some numerically negative eigenvalues for the covariance matrix were found which contradicts to the positive definiteness of the covariance matrix. These results on estimation are likely due to the numerical instability of a corresponding covariance matrix as we see that κ_2 is large where bias is large. Particularly, bad estimation results or negative eigenvalues happen when the condition number calculated by MATLAB function "*cond*" is greater than 10^{17}. This happens when ν gets larger or distance between neighboring observations, δ, becomes smaller. These findings indicate that numerical instability depends on several factors such as covariance parameters (σ^2, α, ν) and the grid size (δ). Thus, we further investigate how these components affect the condition number of the spatial covariance matrix, in turn, the estimation.

The difficulty to study the condition number of a Matérn covariance matrix is that a Matérn covariance function involves the modified Bessel function of the second kind which cannot be expressed in a closed form except when ν is half integer (Gelfand et al., 2012). In Section 3, we show that the condition number of a Matérn covariance matrix depends on δ, ν and α by studying some theoretical bounds of the extreme values of the spectral density and some simulation study.

3. Condition number of a Matérn covariance matrix

For simplicity, we assume that a stationary Gaussian random field $Z(\boldsymbol{s})$ on $\mathbf{R^d}$ with a Matérn covariance function is observed on a regular grid. The corresponding spectral density is

$$f(\boldsymbol{\lambda}) = \frac{\Gamma(\nu + d/2)\alpha^{2\nu}}{\pi^{d/2}\Gamma(\nu)} \frac{\sigma^2}{(\alpha^2 + |\boldsymbol{\lambda}|^2)^{(\nu+d/2)}} \quad \text{for } \boldsymbol{\lambda} \in \mathbf{R^d}. \tag{2}$$