Spatial Statistics 21 (2017) 1–26

Contents lists available at ScienceDirect

Spatial Statistics

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

Efficiently estimating some common geostatistical models by ‘energy–variance matching’ or its randomized ‘conditional–mean’ versions

Didier A. Girard

CNRS and Univ. Grenoble Alpes, Lab. LJK, Grenoble, France

ARTICLE INFO

Article history:
Received 30 June 2016
Accepted 5 January 2017
Available online 23 March 2017

Keywords:
Gaussian random fields
Covariance estimation
Maximum likelihood
Randomized estimating functions
Matérn autocorrelation
Preconditioned conjugate gradient

ABSTRACT

We consider the problem of fitting an isotropic zero-mean stationary Gaussian field model to (possibly noisy) observations, when the model belongs to the Matérn family with known regularity index $\nu > 0$, or to the spherical family. For estimating the correlation range (also called "decorrelation length") and the variance of the field, two simple estimating functions based on the so-called "conditional Gaussian Gibbs-energy mean" (CGEM) and the empirical variance (EV) were recently introduced. This article presents an extensive Monte Carlo simulation study for problems with around a thousand observations and settings including large, moderate, and even "small", correlation ranges. The known observation sites are either on a 2D grid (including a case of "very non-uniform" grid spacings) or randomly uniformly distributed on a simple 2D region. Some experiments for a 256×256 grid with missing values are also analyzed.

This study empirically demonstrates that, for all the (possibly random) uniform designs, the statistical efficiency of CGEM-EV compared to exact maximum likelihood (ML) is globally very satisfactory (except a degradation for the very extremal ranges in some contexts) provided the signal-to-noise ratio (SNR) is strong enough and ν is not too large, this SNR restriction being alleviated as ν decreases. For the "very non-uniform" design, a simple weighting of EV restores this efficiency. In the less favorable cases, the statistical loss remains in fact acceptable: e.g. for the largest considered index ($\nu = 3/2$) and a "not strong enough" SNR, it may happen (in

E-mail addresses: didier.girard@imag.fr, didier.girard@univ-grenoble-alpes.fr.

http://dx.doi.org/10.1016/j.spasta.2017.01.001

fact only for large ranges) that CGEM-EV almost doubles the mean squared error for the range parameter or for the widely used combination of the two parameters known as microergodic-parameter. Furthermore an important conclusion for computational efficiency is that the use of the natural fast randomized-trace version of CGEM-EV does not significantly degrade this statistical efficiency.

1. Introduction

We mainly consider the following statistical model which arises e.g. in remote sensing image analysis: let $Z(\mathbf{s})$, $\mathbf{s} \in \mathbb{R}^2$, be a zero mean stationary Gaussian stochastic process whose autocorrelation function is assumed to belong to the popular isotropic Matérn family. One realization of this process is observed at $n = n_1 \times n_2$ regularly spaced (with step-size δ_1 in abscissa, δ_2 in ordinate) sites $\mathbf{s}_k, k = 1, \ldots, n$, of $[0, (n_1 - 1)\delta_1] \times [0, (n_2 - 1)\delta_2]$, with an additive Gaussian white noise whose variance is σ_N^2 (this noise can model either suspected homoscedastic measurement errors or an additional nugget effect in Z, see e.g. Zhang and Zimmerman, 2007 and references therein). In this article, we restrict ourselves to the case where σ_N is known, e.g. from previous calibration experiments (as it is common when dealing with satellite data, see Tzeng et al., 2005). Using a standard lexicographic ordering, the observations thus form a vector $\boldsymbol{y}$ of size n whose law is Gaussian:

$$\boldsymbol{y} \sim \mathcal{N}(\mathbf{0}, \tau_0^2 \boldsymbol{R}_{\theta_0} + \sigma_N^2 \boldsymbol{I}_n) \tag{1.1}$$

with $\boldsymbol{I}_n$ denoting the identity matrix and $\boldsymbol{R}_\theta$ the autocorrelation matrix of the gridded process i.e. the block Toeplitz matrix (with n_1^2 Toeplitz square blocks, each of size $n_2 \times n_2$) whose coefficients are given by

$$[\boldsymbol{R}_\theta]_{j,k} := \rho_{\nu,\theta}(\|\mathbf{s}_j - \mathbf{s}_k\|), \quad j, k = 1, \ldots, n,$$

$\|\cdot\|$ being the Euclidean norm and $\rho_{\nu,\theta}$ the Matérn autocorrelation function

$$\rho_{\nu,\theta}(x) = \frac{(\theta x)^\nu}{\Gamma(\nu)2^{\nu-1}} K_\nu(\theta x), \quad x > 0, \ \theta > 0,$$

where K_ν is the modified Bessel function of the second kind of order $\nu > 0$. For more details on these widely used autocorrelation functions see Guttorp and Gneiting (2006). Note that

$$\tau_0^2 = \mathbb{E}\left((Z(\mathbf{s}))^2\right) \equiv \mathrm{E}(y_k^2) - \sigma_N^2$$

will be called the process (or signal) variance. When mentioned, we will also consider another well known autocorrelation function, namely the spherical model ρ_θ^S. See e.g. Zhang and Zimmerman (2007) for these definitions. Notice that a significant variant of the above uniform grid, that we call "a very nonuniform Cartesian grid", will be analyzed with some details. We also study a case of $n = 1000$ observation sites randomly but uniformly distributed on a simple 2D region. And, to illustrate the "scalability" offered by the proposed parametric estimation method, we will also consider, albeit with less extensive simulations, a much larger 256×256 grid with a few missing regions (see Section 2.4).

The order ν, often called the regularity (or differentiability) index, is assumed to be known in this paper. Recall that $\rho_{1/2,\theta}(x) = \exp(-\theta x)$ is the very popular exponential model, and that simple expressions also exist for $\rho_{\nu,\theta}(x)$ for $\nu = 3/2$ and $5/2$: these ν's correspond to models also often used (see e.g. Stein, 1999; Rasmussen and Williams, 2006). In the Monte Carlo simulation study of this paper, we only consider three contexts: the order ν will be either 1/6, 1/2 or 3/2.

The parameter θ^{-1} is often called the "decorrelation length" or "the range parameter".

Estimation of the variance and range parameters in such autocovariance models is needed for various tasks, for example for establishing confidence bands for the autocovariance function, for