On Matérn-type Gaussian processes

Didier Girard

LJK, CNRS

Closing ECODEP conference
IHP, Amphitheater Hermite: September 30- October 1, 2024

Matérn correlations, context

Matérn correlations are very popular, their uses date at least to Matheron 1960, maybe Krige ... A Matérn isotropic zero-mean stationary Gaussian field model for (possibly noisy) observations at sites $\vec{s_i} \in \mathbb{R}^d, i = 1, \cdots, n$, with (known) regularity (an index $\nu > 0$), is determined by, $|| \cdot ||$ being the Euclidean norm in $\mathbb{R}^d$:

$$\mathbb{E}(Z(\vec{s})Z(\vec{t})) = b\ \rho_{\nu,\theta}(||\vec{s} - \vec{t}||), \quad \mathbb{E}(Z(\vec{s})^2) \equiv b > 0$$

where

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

where K_ν is the Bessel function of the 2nd kind of order $\nu > 0$.

- $\nu = 1/2$: $\rho_{\nu,\theta}(x) = e^{-\theta x}$ ("Exponential" case. In 1D: O.U.)
- $\nu = 3/2$: $\rho_{\nu,\theta}(x) = (1 + \theta x)e^{-\theta x}$
- $\nu = +\infty$: $\lim_{\nu \to \infty} \rho_{\nu, \frac{2\nu^{\frac{1}{2}}}{\ell}}(x) = e^{-\frac{x^2}{\ell^2}}$ ("squared exponential")

Spectral densities

Models more succinctly defined by their family of spectral densities over $(-\infty, +\infty)^d$:

$$f^*_{\nu,b,\theta}(\overrightarrow{\omega}) = b\, g^*_{\nu,\theta}(\overrightarrow{\omega}), \quad \text{with} \quad g^*_{\nu,\theta}(\overrightarrow{\omega}) := \frac{C(d,\nu)\,\theta^{2\nu}}{(\theta^2 + ||\overrightarrow{\omega}||^2)^{\nu+\frac{d}{2}}}$$

where the constant $C(d,\nu)$ is chosen s.t. $\int_{\mathbb{R}^d} g^*_{\nu,\theta}(\overrightarrow{\omega})\mathrm{d}\overrightarrow{\omega} = 1$

Here one mainly considers gridded data sites, oberved on the hypercube (i.e. data sites are $\delta\{1,\cdots,n\}^{\otimes d}$)

From the well known aliasing formula, the spectral density on $(-\pi,\pi]^d$ of Z_δ is

$$f^\delta_{\nu,\theta}(\overrightarrow{\lambda}) = b\, g^\delta_{\nu,\theta}(\overrightarrow{\lambda}) \quad \text{with} \quad g^\delta_{\nu,\theta}(\overrightarrow{\lambda}) := \frac{1}{\delta^d} \sum_{\overrightarrow{k}\in\mathbb{Z}^d} g^*_{\nu,\theta}\left(\frac{\overrightarrow{\lambda} + 2\overrightarrow{k}\pi}{\delta}\right).$$

Spectral densities, and quasi-Matérn variant

Simple expressions for $g^{\delta}_{\nu,\theta}(\vec{\lambda})$ are available only for $d = 1$, when $\nu - 1/2$ is a small integer, says $0, 1$ or 2 : they then coincide with particular ARMA spectral densities with a *constrained* vector of parameters.

The quasi-Matérn spectral density model one considers here, for a process defined on $\mathbb{Z}^d$, is defined by $f_{\mathcal{Q},\nu,\beta}(\vec{\lambda}) = b\, g_{\mathcal{Q},\nu,\beta}(\vec{\lambda})$ with

$$g_{\mathcal{Q},\nu,\beta}(\vec{\lambda}) := C_{\mathcal{Q}}(d,\nu)\left(d + \beta - \sum_{k=1}^{d} \cos(\lambda_k)\right)^{-\nu - d/2}, \quad \beta > 0,$$

$C_{\mathcal{Q}}(d,\nu)$ being again chosen s.t. $\int_{[\pi,\pi]^d} g_{\mathcal{Q},\nu,\beta}(\vec{\lambda})\mathrm{d}\vec{\lambda} = 1$

- the case $d = 2$, $\nu = 1$ gives the SAR model of order 1

Motivation : As $\delta \downarrow 0$, $g^{\delta}_{\nu,\theta}(\vec{\lambda}) \simeq g_{\mathcal{Q},\nu,\beta}(\delta\vec{\lambda})$ (with $\beta = (\delta\theta)^2/2$).

A not so “well-known” expression

Let ϑ_3 denote the Jacobi 3rd theta function. Recall that $\sqrt{\frac{\pi}{\alpha^2}}\sum_{j=-\infty}^{\infty} e^{-\frac{(2\pi j+\lambda)^2}{4\alpha^2}} = \vartheta_3\left(\frac{\lambda}{2}, e^{-\alpha^2}\right)$

Proposition 1

With $\alpha = \delta\theta$ and $w_\nu(\kappa) := \frac{e^{-\frac{1}{4\kappa}}\kappa^{-\nu-1}}{2^{2\nu}\Gamma(\nu)}$ (it is a pdf on $[0, \infty[$), it holds

$$g^{\delta}_{\nu,\theta}(\lambda_1, \cdots, \lambda_d) = \int_0^{\infty} g_{\infty,\alpha\phi}(\lambda_1, \cdots, \lambda_d) \times 2\phi w_\nu(\phi^2)\, \mathrm{d}\phi$$

where

$$g_{\infty,\alpha}(\vec{\lambda}) = \frac{1}{(2\pi)^d}\prod_{k=1}^{d} \vartheta_3\left(\frac{\lambda_k}{2}, e^{-\alpha^2}\right).$$

Two "monotonicity" results (or conjectures)

Let $d = 1, 2, \cdots$ and $\nu > 0$ fixed. In the following, $g_\alpha(\cdot)$ stands either for $g_{\nu,\alpha}^{\delta=1}(\cdot)$ (which coincides with $g_{\nu,\frac{\alpha}{\delta}}^{\delta}(\cdot)$) or for $g_{\mathcal{Q},\nu,\alpha}(\cdot)$.

Monotonicity of the "entropy-rate"

$$\alpha \in \mathbb{R}_+ \longrightarrow \int_{[-\pi,\pi]^d} \log\left(g_\alpha(\vec{\lambda})\right) \mathrm{d}\vec{\lambda}$$

is strictly increasing.

'Unimodality" property. Let α_0 fixed (> 0)

$$\alpha \in \mathbb{R}_+ \longrightarrow \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{g_{\alpha_0}(\vec{\lambda})}{g_\alpha(\vec{\lambda})} \mathrm{d}\vec{\lambda} \quad - \quad 1 \quad =: \quad F(\alpha_0, \alpha)$$

has a unique finite root $\alpha = \alpha_0$.

Two “monotonicity” results (or conjectures)

Their proofs will be (hopefully) published soon.

- They are easy for the quasi-Matérn spectral densities.
- For pure Matérn spectral densities, the proofs are relatively easy for $d = 1$.

Consequence of this “entropy monotonicity” property

As is well known, in 1D, this entropy rate is proportional to the prediction error variance of $Z(t = 0)$ based on the past (Szego’s formula).
Generalization of the Szego’s formula is also known in 2D when the “past” is a half-plane.
This "entropy" monotonicity thus indicates that the ability to extrapolate at unobserved sites (in geostatistic it’s called the ability to interpolate), of such processes, whose the underlying true range parameter is θ, is strenghened when θ increases.

Consequences of this "Unimodality" property

The so-called "(Gaussian) energy-variance matching" (GE-EV) method for estimating α_0 is an empirical version of the above equation $F(\alpha_0, \alpha) = 0$ in α.
Recall that the (maybe simplest version of) GE–EV estimating equation is defined by replacing, in this integrated ratio, $g_{\alpha_0}(\cdot)$ by a tapered periodogram normalized by $\frac{1}{n^d} \sum_{s \in \delta\{1,\cdots,n\}^{\otimes d}} Z(s)^2$.
The previously published consistency for GE–EV was only a weak result ("there exists a sequence of roots ...") and required the assumption : α_0 (or δ) must be small..
This "unimodality" property implies a global consistency result, even without this assumption.

It also entails some "robustness" properties for simple algorithms for computing the GE–EV estimate of α_0

A third "monotonicity" result (or conjecture...)

Let $h_\alpha(\overrightarrow{\lambda}) = \frac{\partial}{\partial\alpha}\log(g_\alpha(\overrightarrow{\lambda}))$. Define the "coefficient of variation" of f (where "$\int$" denotes integrals over $[-\pi,\pi]^d$) by:

$$\mathrm{CV}(f) := \left\{ \frac{1}{(2\pi)^d} \int \left(\left| f - \left(\frac{1}{(2\pi)^d} \int f \right) \right|^2 \right) \right\} \Bigg/ \left(\frac{1}{(2\pi)^d} \int f \right)^2$$

Monotonicity property of the "inefficiency"

$$\alpha \in \mathbb{R}_+ \longrightarrow \mathrm{CV}(h_\alpha)\,\mathrm{CV}(g_\alpha)$$

is a decreasing function.

Consequences of this "CVs product monotonicity"

Assume the process is Gaussian Matérn (or Gaussian quasi-Matérn) with sdf $b_0\, g_{\alpha_0}$ (recall that $\alpha_0 = \delta\theta_0$ for pure Matérn).
$\mathrm{inefficiency}_{\nu,\alpha_0}$:= "asymptotic variance" (a.v.) of the GE–EV estimator divided by the a.v. of the maximum likelihood estimator .

Recall

Theorem (slight extension of Girard 2016).

For $d = 1$

$$\mathrm{inefficiency}_{\nu,\alpha_0} = \mathrm{CV}(h_{\alpha_0})\,\mathrm{CV}(g_{\alpha_0}) \to \mathit{ineff}_1(\nu) \quad \mathit{as} \quad \alpha_0 \downarrow 0,$$

where $\mathit{ineff}_1(\nu) := \frac{\sqrt{\pi}}{2}\left(\frac{\Gamma(\nu+3/2)}{\Gamma(\nu+1)}\right)^2 \frac{\Gamma(2\nu+1/2)}{\Gamma(2\nu+1)}$.

The above approximates values of $\text{inefficiency}_{\nu,\alpha_0}$ for the cases when the true α_0 is small are satisfactory since they are close to 1 for typical values of ν. For instance

$$\mathit{ineff}_1(1/2) = 1, \quad \mathit{ineff}_1(3/2) = 1.054^2, \quad \mathit{ineff}_1(5/2) = 1.122^2$$

Conjecture (extension of the near-efficiency result of Girard 2016 to any α_0 (or to any δ in the case of pure Matérn)).

$$\forall \alpha_0 > 0, \quad \text{inefficiency}_{\nu,\alpha_0} \leq \mathit{ineff}_d(\nu)$$

where $\mathit{ineff}_d(\nu) := \lim_{\alpha \downarrow 0} \mathrm{CV}(h_\alpha)\,\mathrm{CV}(g_\alpha)$.

Theorem.

For $d = 1$, the above conjecture holds true for $\nu - 1/2$ a "small" integers; and thus, in particular, for any true α_0

- $\mathrm{inefficiency}_{\nu,\alpha_0} \equiv 1$ in the case $\nu = 1/2$,
- $\mathrm{inefficiency}_{\nu,\alpha_0} \leq 10/9$ in the case $\nu = 3/2$,
- $\mathrm{inefficiency}_{\nu,\alpha_0} \leq 63/50$ in the case $\nu = 5/2$,
- $\mathrm{inefficiency}_{\nu,\alpha_0} \leq \frac{1716}{1225} \simeq 1.18356^2$ in the case $\nu = 7/2$.

In 2D, a numerical verification for quasi-Matérn$_{\frac{1}{2}}$

$$\nu = \frac{1}{2}$$

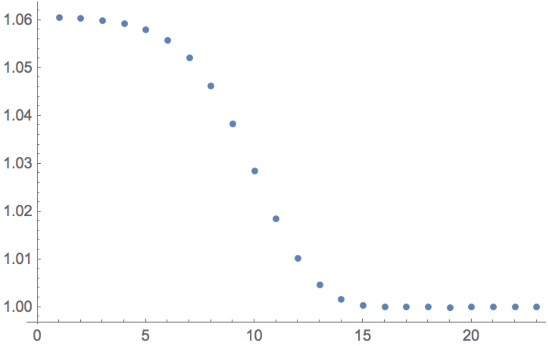

Figure 1 : $\alpha_0 \to \sqrt{(\text{inefficiency}_{\nu,\alpha_0})}$

In 2D, for quasi-Matérn$_{\frac{3}{2}}$

$\nu = \frac{3}{2}$

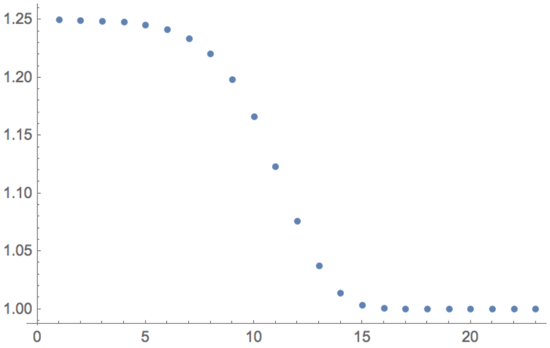

Figure 2 : $\alpha_0 \to \sqrt{(\mathrm{inefficiency}_{\nu,\alpha_0})}$

A finite size analog of "monotonicity of the entropy-rate"

Let $d = 1, 2, \cdots$ and $\nu > 0$ fixed. Let $S = \{\vec{s_i}, i = 1, \cdots, n\}$ fixed. In the following, $R_{\nu,\theta}$ denotes the $n \times n$ matrix with $(i,j)\text{th term} := \rho_{\nu,\theta}(||\vec{s_i} - \vec{s_j}||)$

Monotonicity of the determinant

$$\theta \in \mathbb{R}_+ \longrightarrow \det\left(R_{\nu,\theta}\right)$$

is strictly increasing.

Theorem. If $\nu = \frac{1}{2}$ then

this Monotonicity of the determinant holds true.

Proof. It stems on $R_{\nu,\theta_1+\theta_2} \equiv R_{\nu,\theta_1} \circ R_{\nu,\theta_2}$ when $\nu = 1/2$, and a known determinant inequality for Hadamard products (Oppenheim 1930).

Consequence of this determinant's monotonicity

Let $F_n(\theta_0, \theta) := \frac{1}{n}\mathrm{trace}\left(R_{\nu,\theta}^{-1} R_{\nu,\theta_0}\right)$.
It is known that

$$F_n(\theta_0, \theta) \leq \left(\det\left(R_{\nu,\theta}^{-1} R_{\nu,\theta_0}\right)\right)^{\frac{1}{n}}$$

Theorem. If $\nu = \frac{1}{2}$ then

the function

$$\theta \longrightarrow F_n(\theta_0, \theta) - 1$$

has no root in $\theta \in]0, \theta_0[$.

NB: The GE-EV method for estimating θ_0 is an empirical version of the above equation $F_n(\theta_0, \theta) = 1$ in θ.

Some References

- Girard, D.A., 2015. "Estimating a Centered Matérn-1 Process: Three Alternatives to Maximum Likelihood via Conjugate Gradient Linear Solvers". Wolfram Demonstrations Project.
- Girard, D.A., 2016. Asymptotic near-efficiency of the "Gibbs-energy and empirical-variance" estimating functions for fitting Matérn models - I: Densely sampled processes. Statistics & Probability Letters, 110, 191-197.
- Girard, D.A., 2017. Efficiently estimating some common geostatistical models by 'energy-variance matching' or its randomized 'conditional-mean' versions. Spatial Statistics, 21(Part A), 1-26.
- Drouilhet R., Girard, D.A., 2017. R-package CGEM–EV. https://github.com/didiergirard/CGEMEV

THANK YOU FOR YOUR ATTENTION