When are linear predictions of random fields using wrong mean and covariance functions asymptotically optimal?

Kristin Kirchner

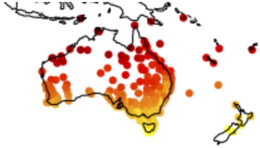
Joint work with David Bolin (KAUST)

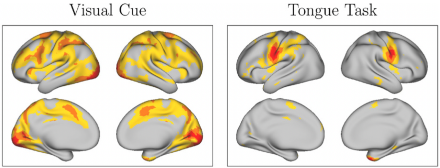
One World Seminar:

Stochastic Numerics and Inverse Problems

March 2, 2022

Linear spatial prediction





Consider a random field $\{Z(x) : x \in \mathcal{X}\}$ on a compact metric space $\mathcal{X}$.

Aim: Predict its value $Z(x^*)$ at $x^* \in \mathcal{X}$ based on a set of observations $\{Z(x_j)\}_{j=1}^n$ for locations $x_1, \ldots, x_n \in \mathcal{X}$ all distinct from x^*.

The **kriging predictor** is the linear predictor

$$Z_n(x^*) = \alpha_0 + \sum_{j=1}^{n} \alpha_j Z(x_j)$$

based on the observations, where $\alpha_0, \ldots, \alpha_n \in \mathbb{R}$ are chosen such that the variance of the error $Z_n(x^*) - Z(x^*)$ is minimized.

Model misspecification

If $m(\cdot)$ and $\varrho(\cdot,\cdot)$ are the mean and the covariance function of Z, then

$$Z_n(x^*) = m(x^*) + \mathbf{c}^\top \boldsymbol{\Sigma}^{-1}(\mathbf{Z} - \mathbf{m}), \qquad (*)$$

where

$$\mathbf{c} := (\varrho(x^*, x_1), \ldots, \varrho(x^*, x_n))^\top, \qquad \boldsymbol{\Sigma}_{ij} := \varrho(x_i, x_j),$$
$$\mathbf{Z} := (Z(x_1), \ldots, Z(x_n))^\top, \qquad \mathbf{m} := (m(x_1), \ldots, m(x_n))^\top.$$

$\Rightarrow$ the *kriging predictor* depends only on (m, ϱ).

Therefore, from now on we assume that $Z \sim \mu = \mathrm{N}(m, \varrho)$.

We are interested in the asymptotic behavior of

$$\frac{\mathrm{E}\left[(\widetilde{Z}_n(x^*) - Z(x^*))^2\right]}{\mathrm{E}\left[(Z_n(x^*) - Z(x^*))^2\right]} \quad \text{as} \quad n \to \infty,$$

where the linear predictor $\widetilde{Z}_n(x^*)$ is computed using $(*)$ with $(\widetilde{m}, \widetilde{\varrho})$.

What was known?

- M. L. Stein[1] showed that the best linear predictor based on $(\widetilde{m}, \widetilde{\varrho})$ is asymptotically optimal, as $n \to \infty$, provided that

$$\mu = \mathrm{N}(m, \varrho) \quad \text{and} \quad \widetilde{\mu} = \mathrm{N}(\widetilde{m}, \widetilde{\varrho}) \quad \text{are equivalent.}$$

- This result in fact holds *uniformly* with respect to x^* and, moreover, uniformly for each linear functional φ such that $\varphi(Z)$ has finite variance[2].
- Less restrictive conditions have been derived for some specific cases, such as *periodic* fields[3] on $[0,1]^d$ and *stationary* fields[4] on $\mathbb{R}^d$, i.e. $\varrho(x,y) = \varrho(x-y)$, observed on a lattice.

---

[1] M. L. Stein (1988). "Asymptotically efficient prediction of a random field with a misspecified covariance function". In: *Ann. Stat.* 16.1, pp. 55–63

[2] M. L. Stein (1990). "Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure". In: *Ann. Stat.* 18.2, pp. 850–872

[3] M. L. Stein (1997). "Efficiency of linear predictors for periodic processes using an incorrect covariance function". In: *J. Statist. Plann. Inference* 58.2, pp. 321–331

[4] M. L. Stein (1999). "Predicting random fields with increasing dense observations". In: *Ann. Appl. Probab.* 9.1, pp. 242–273

Goals of our work

For any constant $c \in (0, \infty)$, the linear predictor based on $(m, c\varrho)$ is equal to that based on (m, ϱ), whereas

$$\forall c \neq 1 : \quad \mu = \mathrm{N}(m, \varrho) \ \perp \ \widetilde{\mu} = \mathrm{N}(m, c\varrho).$$

⇒ Equivalence of the measures μ and $\widetilde{\mu}$ is a *sufficient, but not necessary* assumption for asymptotic optimality of linear predictions.

Topics of this talk

1. *necessary and sufficient* conditions on $(\widetilde{m}, \widetilde{\varrho})$ for uniform asymptotic optimality of linear predictions;
2. explicit conditions for a large class of non-stationary models and
   - equivalence of Gaussian measures;
   - asymptotic optimality of linear predictions.

Outline

Outline

Setting

- $Z : \mathcal{X} \times \Omega \to \mathbb{R}$ is a square-integrable Gaussian stochastic process indexed by a connected, compact metric space $(\mathcal{X}, d_{\mathcal{X}})$, with strictly positive and finite Borel measure $\nu_{\mathcal{X}} : \mathcal{B}(\mathcal{X}) \to [0, \infty)$.

$$\text{Notation :} \qquad L_2 := L_2(\mathcal{X}, \mathcal{B}(\mathcal{X}), \nu_{\mathcal{X}}).$$

- Z has mean $m \in L_2$, strictly positive definite and continuous covariance function $\varrho : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, and covariance operator

$$\mathcal{C} : L_2 \to L_2, \qquad (\mathcal{C}w)(x) := \int_{\mathcal{X}} \varrho(x, x') w(x') \, \mathrm{d}\nu_{\mathcal{X}}(x').$$

  $\mathcal{C}$ is self-adjoint, positive definite, and trace-class on L_2.

- We write $\mu = \mathrm{N}(m, \mathcal{C})$ for the Gaussian measure on L_2 induced by the process Z, i.e., for every Borel set $A \in \mathcal{B}(L_2)$,

$$\mu(A) = \mathbb{P}(\{\omega \in \Omega : Z(\cdot, \omega) \in A\}).$$

Kriging predictor as orthogonal projection

- $Z^0 := Z - m$ is a *centered* Gaussian process.
- The vector space $\mathcal{Z}^0 \subset L_2(\Omega, \mathbb{P})$ contains finite linear combinations

$$\mathcal{Z}^0 := \left\{ \sum_{j=1}^K \alpha_j Z^0(x_j) : K \in \mathbb{N},\ \alpha_j \in \mathbb{R},\ x_j \in \mathcal{X} \right\}, \quad \mathcal{H}^0 := \overline{\mathcal{Z}^0}^{\|\cdot\|_{L_2(\Omega,\mathbb{P})}}.$$

- Every $h = \sum_{j \le K} \alpha_j Z(x_j)$ has a representation

$$h = c + h^0, \qquad \text{with} \quad c \in \mathbb{R}, \quad h^0 \in \mathcal{Z}^0 \subset \mathcal{H}^0.$$

  We thus define the Hilbert space $\mathcal{H} := \mathbb{R} \oplus \mathcal{H}^0$,

$$(g, h)_{\mathcal{H}} = \mathsf{E}[g]\mathsf{E}[h] + (g - \mathsf{E}[g], h - \mathsf{E}[h])_{\mathcal{H}^0} = \mathsf{E}[gh].$$

- The kriging predictor h_n of $h \in \mathcal{H}$ based on $\mathcal{H}_n := \mathbb{R} \oplus \mathcal{H}_n^0 \subset \mathbb{R} \oplus \mathcal{H}^0$ is the $\mathcal{H}$-orthogonal projection of h onto $\mathcal{H}_n$:

$$h_n \in \mathcal{H}_n : \quad (h_n - h, g_n)_{\mathcal{H}} = \mathsf{E}\left[(h_n - h) g_n\right] = 0 \qquad \forall g_n \in \mathcal{H}_n,$$

$$h_n \in \mathcal{H}_n : \quad \mathsf{E}\left[(h_n - h)^2\right] = \inf_{g_n \in \mathcal{H}_n} \mathsf{E}\left[(g_n - h)^2\right].$$

Problem formulation

$$h_n \in \mathcal{H}_n : \qquad (h_n - h, g_n)_{\mathcal{H}} = \mathrm{E}\left[(h_n - h) g_n\right] = 0 \qquad \forall g_n \in \mathcal{H}_n,$$

Main question

What happens if, instead of h_n, we use the linear predictor $\widetilde{h}_n$ which is the kriging predictor if $\widetilde{\mu} = \mathrm{N}(\widetilde{m}, \widetilde{\mathcal{C}})$ was the correct model?

"Computing orthogonal projections with the wrong inner product"

Consistent kriging prediction

- We require that, for every $h \in \mathcal{H}$, the corresponding kriging predictors $\{h_n\}_{n\in\mathbb{N}}$ are consistent in the sense that

$$\lim_{n\to\infty} \mathrm{E}\big[(h_n - h)^2\big] = \lim_{n\to\infty} \|h_n - h\|_{\mathcal{H}}^2 = 0. \qquad (\textit{density})$$

- Let $\mathcal{S}^{\mu}_{adm}$ contain all *admissible* sequences $\{\mathcal{H}_n\}_{n\in\mathbb{N}}$ of subspaces $\mathcal{H}_n \subset \mathcal{H}$ which provide μ-consistent kriging prediction,

$$\mathcal{S}^{\mu}_{adm} := \Big\{\{\mathcal{H}_n\}_{n\in\mathbb{N}} \,\Big|\, \forall n \in \mathbb{N} : \mathcal{H}_n = \mathbb{R} \oplus \mathcal{H}_n^0 \text{ with } \dim(\mathcal{H}_n^0) = n, \\ \forall h \in \mathcal{H} : \{h_n\}_{n\in\mathbb{N}} \text{ satisfy } (\textit{density})\Big\}.$$

Example

Suppose that m, ϱ are continuous and that every $\mathcal{H}_n$ is generated by point observations $Z(x_1), Z(x_2), \ldots$ Then, $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{adm}$ if $\{x_j\}_{j\in\mathbb{N}}$ is a sequence in $(\mathcal{X}, d_{\mathcal{X}})$ which *accumulates at any* $x^* \in \mathcal{X}$.

Outline

Main assumptions

The Assumptions

Let $\varrho, \widetilde{\varrho} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be two continuous, (strictly) positive definite covariance functions. Assume that the corresponding covariance operators $\mathcal{C}, \widetilde{\mathcal{C}} : L_2 \to L_2$, and $m, \widetilde{m} \in L_2$ are such that:

I. The vector spaces

$$\mathcal{C}^{1/2}(L_2), \left(\mathcal{C}^{-1/2}\cdot, \mathcal{C}^{-1/2}\cdot\right)_{L_2} \quad \text{and} \quad \widetilde{\mathcal{C}}^{1/2}(L_2), \left(\widetilde{\mathcal{C}}^{-1/2}\cdot, \widetilde{\mathcal{C}}^{-1/2}\cdot\right)_{L_2}$$

are *norm equivalent* Hilbert spaces.

II. The difference of the means satisfies $m - \widetilde{m} \in \mathcal{C}^{1/2}(L_2)$.

III. There exists a positive constant $a \in (0, \infty)$ such that the operator

$$T_a : L_2 \to L_2, \qquad T_a := \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - a\mathcal{I}$$

is *compact* on L_2. Here $\mathcal{I}$ denotes the identity on L_2.

Theorem (Asymptotic optimality, K. and Bolin, 2020)

Let $\mu = \mathsf{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathsf{N}(\widetilde{m}, \widetilde{\mathcal{C}})$. Let $h_n, \widetilde{h}_n$ denote the best linear predictors of h based on $\mathcal{H}_n$ and μ resp. $\widetilde{\mu}$. Then, any of the assertions,

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathsf{E}\big[(\widetilde{h}_n - h)^2\big]}{\mathsf{E}\big[(h_n - h)^2\big]} = 1,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathsf{E}}\big[(h_n - h)^2\big]}{\widetilde{\mathsf{E}}\big[(\widetilde{h}_n - h)^2\big]} = 1,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left|\frac{\widetilde{\mathsf{E}}\big[(h_n - h)^2\big]}{\mathsf{E}\big[(h_n - h)^2\big]} - a\right| = 0,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left|\frac{\mathsf{E}\big[(\widetilde{h}_n - h)^2\big]}{\widetilde{\mathsf{E}}\big[(\widetilde{h}_n - h)^2\big]} - \frac{1}{a}\right| = 0,$$

holds for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{adm}$ if and only if the Assumptions I–III *are satisfied. The constant $a \in (0,\infty)$ is the same as that in* Assumption III.

Comparison with the Feldman–Hájek theorem

Equivalence of Gaussian measures

Let μ and $\widetilde{\mu}$ be two measures on $(L_2, \mathcal{B}(L_2))$. Then, μ and $\widetilde{\mu}$ are called

- equivalent if $\mu(A) = 0 \Leftrightarrow \widetilde{\mu}(A) = 0$;
- orthogonal if there exists $B \in \mathcal{B}(L_2)$ with $\mu(B) = 0$ and $\widetilde{\mu}(B) = 1$.

Theorem (Feldman–Hájek)

Two Gaussian measures $\mu = \mathrm{N}(m, \mathcal{C})$ and $\widetilde{\mu} = \mathrm{N}(\widetilde{m}, \widetilde{\mathcal{C}})$ on a separable Hilbert space $(E, (\cdot, \cdot)_E)$ are either orthogonal or equivalent. They are equivalent if and only if the following conditions are satisfied:

- *The Cameron–Martin spaces $\mathcal{C}^{1/2}(E)$, $\widetilde{\mathcal{C}}^{1/2}(E)$ are norm equivalent.*
- *The difference of the means satisfies $m - \widetilde{m} \in \mathcal{C}^{1/2}(E)$.*
- *The operator $T_1 := \mathcal{C}^{-1/2}\widetilde{\mathcal{C}}\mathcal{C}^{-1/2} - \mathrm{Id}_E$ is Hilbert–Schmidt on E.*

Proposition (Role of Assumption I)

Let $\mu = \mathrm{N}(m, \mathcal{C})$, $\widetilde{\mu} = \mathrm{N}(\widetilde{m}, \widetilde{\mathcal{C}})$, *and define* $\mathcal{H}^0, \widetilde{\mathcal{H}}^0$ *with respect to the measures* μ *and* $\widetilde{\mu}$, *respectively. The following are equivalent:*

(i) Assumption I *is satisfied.*

(ii) *The linear operator* $\widetilde{\mathcal{C}}^{1/2}\mathcal{C}^{-1/2} : L_2 \to L_2$ *is an isomorphism.*

(iii) *The Hilbert spaces* $\mathcal{H}^0, \widetilde{\mathcal{H}}^0$ *are* norm equivalent. *In particular,*

$$\exists k_0, k_1 > 0 : \quad k_0 \mathrm{Var}[h] \le \widetilde{\mathrm{Var}}[h] \le k_1 \mathrm{Var}[h] \quad \forall h \in \mathcal{H}.$$

(iv) *There exist* $0 < k \le K < \infty$ *such that, for every* $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{adm}$, *for all* $n \in \mathbb{N}$, *and every* $h \in \mathcal{H}_{-n}$,

$$\frac{\widetilde{\mathrm{Var}}[h_n - h]}{\mathrm{Var}[h_n - h]}, \ \frac{\mathrm{Var}[\widetilde{h}_n - h]}{\widetilde{\mathrm{Var}}[\widetilde{h}_n - h]}, \ \frac{\mathrm{Var}[\widetilde{h}_n - h]}{\mathrm{Var}[h_n - h]}, \ \frac{\widetilde{\mathrm{Var}}[h_n - h]}{\widetilde{\mathrm{Var}}[\widetilde{h}_n - h]} \in [k, K].$$

$h_n, \widetilde{h}_n$ *are the best linear predictors of h based on* $\mathcal{H}_n$ *and* μ *resp.* $\widetilde{\mu}$.

Proposition (Role of Assumptions I and III)

Let $\mu = \mathrm{N}(m, \mathcal{C})$ *and* $\widetilde{\mu} = \mathrm{N}(\widetilde{m}, \widetilde{\mathcal{C}})$*. Let* $h_n, \widetilde{h}_n$ *denote the best linear predictors of* h *based on* $\mathcal{H}_n$ *and* μ *resp.* $\widetilde{\mu}$*. Then, any of the assertions,*

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\mathrm{Var}\left[\widetilde{h}_n - h\right]}{\mathrm{Var}\left[h_n - h\right]} = 1,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \frac{\widetilde{\mathrm{Var}}\left[h_n - h\right]}{\widetilde{\mathrm{Var}}\left[\widetilde{h}_n - h\right]} = 1,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\widetilde{\mathrm{Var}}\left[h_n - h\right]}{\mathrm{Var}\left[h_n - h\right]} - a \right| = 0,$$

$$\lim_{n\to\infty} \sup_{h\in\mathcal{H}_{-n}} \left| \frac{\mathrm{Var}\left[\widetilde{h}_n - h\right]}{\widetilde{\mathrm{Var}}\left[\widetilde{h}_n - h\right]} - \frac{1}{a} \right| = 0,$$

holds for all $\{\mathcal{H}_n\}_{n\in\mathbb{N}} \in \mathcal{S}^{\mu}_{adm}$ *if and only if* Assumptions I and III *are fulfilled. The constant* $a \in (0, \infty)$ *is the same as that in* Assumption III.

Outline

Examples of valid covariance models and metric spaces

Valid covariance functions are given by the Matérn class,

$$\varrho(x, x') := \varrho_0\big(d_{\mathcal{X}}(x, x')\big), \qquad \varrho_0(r) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa r)^{\nu} K_{\nu}(\kappa r), \quad r \geq 0,$$

on the compact metric space $(\mathcal{X}, d_{\mathcal{X}})$, where

- $\mathcal{X} \subset \mathbb{R}^d$ is a *connected, compact Euclidean domain* for $\nu, \kappa, \sigma^2 > 0$.
- $\mathcal{X} := \mathbb{S}^d$ is the *d-sphere* equipped with the great circle distance, for every $\nu \in (0, 1/2]$ and all $\kappa, \sigma^2 > 0$[5].
- $\mathcal{X}$ is a *graph with Euclidean edges* equipped with the resistance metric for every for every $\nu \in (0, 1/2]$ and all $\kappa, \sigma^2 > 0$[6].

[5] T. Gneiting (2013). "Strictly and non-strictly positive definite functions on spheres". In: *Bernoulli* 19.4, pp. 1327–1349

[6] E. Anderes, J. Møller, and J. G. Rasmussen (2020). "Isotropic covariance functions on graphs and their edges". In: *Ann. Statist.* 48.4, pp. 2478–2503

Weakly stationary random fields on $\mathbb{R}^d$

Setting: $\mathcal{X} \subset \mathbb{R}^d$ and $\varrho|_{\mathcal{X}\times\mathcal{X}}$, $\widetilde{\varrho}|_{\mathcal{X}\times\mathcal{X}}$ are restrictions of continuous, positive definite, translation invariant functions $\varrho, \widetilde{\varrho} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$.

$$\exists \varrho_0, \widetilde{\varrho}_0 : \mathbb{R}^d \to \mathbb{R} \text{ even}: \quad \varrho(x,x') = \varrho_0(x-x'), \quad \widetilde{\varrho}(x,x') = \widetilde{\varrho}_0(x-x').$$

The spectral density f and ϱ_0 relate via the inversion formula:

$$\forall \omega \in \mathbb{R}^d : \quad f(\omega) = \tfrac{1}{(2\pi)^d}(\mathcal{F}\varrho_0)(\omega), \;\; (\mathcal{F}\varrho_0)(\omega) := \int_{\mathbb{R}^d} e^{-i\omega\cdot x}\varrho_0(x)\,\mathrm{d}x.$$

We define $\mathcal{F}_{\mathcal{X}} := \mathcal{F} \circ E^0_{\mathcal{X}}$, where $E^0_{\mathcal{X}}$ is the zero extension, and

$$\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X})) = \left\{ \hat{w} : \mathbb{R}^d \to \mathbb{C} \;\middle|\; \exists w \in L_2(\mathcal{X}) : \hat{w} = \mathcal{F}_{\mathcal{X}} w \right\} \subset L_2(\mathbb{R}^d;\mathbb{C}).$$

The Hilbert space H_f (over $\mathbb{R}$) is the closure of $\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))$ with respect to norm induced by the inner product

$$(\hat{v}_1, \hat{v}_2)_{H_f} := \int_{\mathbb{R}^d} f(\omega)\hat{v}_1(\omega)\overline{\hat{v}_2(\omega)}\,\mathrm{d}\omega, \qquad H_f := \overline{\mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X}))}^{\|\cdot\|_{H_f}}.$$

Proposition (Assumptions I and III in terms of spectral densities)

Suppose that $\mathcal{C}, \widetilde{\mathcal{C}} : L_2(\mathcal{X}) \to L_2(\mathcal{X})$ *pertain to restrictions (to* $\mathcal{X} \times \mathcal{X}$*) of translation invariant covariance functions* $\varrho, \widetilde{\varrho} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, *which have spectral densities* $f, \widetilde{f} : \mathbb{R}^d \to [0, \infty)$.

Then, Assumptions I and III *are satisfied if and only if:*

- I' *The spaces* H_f *and* $H_{\widetilde{f}}$ *are isomorphic with equivalent norms, i.e., there exist constants* $0 < k \leq K < \infty$ *such that*

$$k \|\hat{v}\|_{H_f}^2 \leq \int_{\mathbb{R}^d} \widetilde{f}(\omega) |\hat{v}(\omega)|^2 \, \mathrm{d}\omega \leq K \|\hat{v}\|_{H_f}^2 \quad \forall \hat{v} \in \mathcal{F}_{\mathcal{X}}(L_2(\mathcal{X})).$$

- III' *There exists* $a \in (0, \infty)$ *such that the linear operator* $\widehat{T}_a := S - a\mathcal{I}_{H_f}$ is compact *on* H_f, *where* $\mathcal{I}_{H_f}$ *denotes the identity on* H_f *and* $S : H_f \to H_f$ *is defined by*

$$(S\hat{v}_1, \hat{v}_2)_{H_f} = \int_{\mathbb{R}^d} \widetilde{f}(\omega) \hat{v}_1(\omega) \overline{\hat{v}_2(\omega)} \, \mathrm{d}\omega \qquad \forall \hat{v}_1, \hat{v}_2 \in H_f.$$

$$\varrho(x, x') := \varrho_0\big(\|x - x'\|_{\mathbb{R}^d}\big), \quad \varrho_0(r) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa r)^\nu K_\nu(\kappa r), \quad r \geq 0,$$

Parameters: $\nu, \kappa, \sigma^2 > 0$.

Example (Matérn covariance family)

Assumptions I and III are satisfied if and only if $\nu = \widetilde{\nu}$. In this case:

$$a = \frac{\widetilde{\sigma}^2 \widetilde{\kappa}^{2\nu}}{\sigma^2 \kappa^{2\nu}} \in (0, \infty).$$

For equivalence of the corresponding Gaussian measures, $a = 1$ is necessary. Indeed, Zhang[7] and Anderes[8] showed that, for $\nu = \widetilde{\nu}$,

$$\mu \sim \widetilde{\mu} \quad \Longleftrightarrow \quad \begin{cases} \sigma^2 \kappa^{2\nu} = \widetilde{\sigma}^2 \widetilde{\kappa}^{2\nu} & \text{for } d \leq 3, \\ \kappa = \widetilde{\kappa} \text{ and } \sigma^2 = \widetilde{\sigma}^2 & \text{for } d \geq 5. \end{cases}$$

[7] H. Zhang (2004). "Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics". In: *J. Amer. Statist. Assoc.* 99.465, pp. 250–261

[8] E. Anderes (2010). "On the consistent separation of scale and variance for Gaussian random fields". In: *Ann. Statist.* 38.2, pp. 870–893

Outline

Covariance operators with the same eigenbasis

Corollary

Suppose that $\mathcal{C}, \widetilde{\mathcal{C}}$ *are self-adjoint, positive definite, trace-class operators on* $L_2(\mathcal{X}, \nu_{\mathcal{X}})$ *which diagonalize with respect to the same orthonormal basis* $\{e_j\}_{j\in\mathbb{N}}$ *for* $L_2(\mathcal{X}, \nu_{\mathcal{X}})$*, i.e., there are* $\{\gamma_j\}_{j\in\mathbb{N}}, \{\widetilde{\gamma}_j\}_{j\in\mathbb{N}} \subset (0, \infty)$ *s.t.*

$$\mathcal{C} e_j = \gamma_j e_j \qquad \text{and} \qquad \widetilde{\mathcal{C}} e_j = \widetilde{\gamma}_j e_j \qquad \forall j \in \mathbb{N}.$$

Consider $\mu := \mathsf{N}(0, \mathcal{C})$ *and* $\widetilde{\mu} := \mathsf{N}(0, \widetilde{\mathcal{C}})$.

- *The Cameron–Martin spaces for* μ *and* $\widetilde{\mu}$ *are isomorphic if and only if there exist* $c_-, c_+ \in \mathbb{R}_+$ *such that* $\widetilde{\gamma}_j/\gamma_j \in [c_-, c_+]$ *for all* $j \in \mathbb{N}$.
- μ *and* $\widetilde{\mu}$ *are equivalent if and only if* $\sum_{j\in\mathbb{N}} (\widetilde{\gamma}_j/\gamma_j - 1)^2 < \infty$.
- Assumptions I and III *are satisfied if and only if there exists a constant* $a \in (0, \infty)$ *such that* $\lim_{j\to\infty} \widetilde{\gamma}_j/\gamma_j = a$.

Consider the shifted negative Dirichlet Laplacian:

$$Lv := \left(-\Delta + \kappa^2\right)v, \qquad v \in \mathscr{D}(L) := H^2(\mathcal{D}) \cap H^1_0(\mathcal{D}),$$

where $\mathcal{D} \subsetneq \mathbb{R}^d$ is a bounded open Lipschitz domain.

Corollary ("Classical" Whittle–Matérn fields)

Let $d \in \mathbb{N}$, $\beta, \widetilde{\beta} > d/4$, $\tau, \widetilde{\tau} > 0$, and let $L, \widetilde{L}$ have shift parameters $\kappa^2 \geq 0$ and $\widetilde{\kappa}^2 \geq 0$, respectively. Consider on $L_2(\mathcal{D})$ the Gaussian measures

$$\mu = \mathrm{N}\left(0, \tau^{-2} L^{-2\beta}\right) \quad \textit{and} \quad \widetilde{\mu} = \mathrm{N}\left(0, \widetilde{\tau}^{-2} \widetilde{L}^{-2\widetilde{\beta}}\right).$$

- *The Cameron–Martin spaces for μ and $\widetilde{\mu}$ are* isomorphic, *with* equivalent norms, *if and only if $\beta = \widetilde{\beta}$.*
- *μ and $\widetilde{\mu}$ are equivalent if and only if*

$$\begin{cases} \beta = \widetilde{\beta} \ \text{and} \ \tau = \widetilde{\tau} & \text{if } d \leq 3, \\ \beta = \widetilde{\beta}, \ \tau = \widetilde{\tau} \ \text{and} \ \kappa^2 = \widetilde{\kappa}^2 & \text{if } d \geq 4. \end{cases}$$

- *For any $d \in \mathbb{N}$,* Assumptions I & III *are fulfilled if and only if $\beta = \widetilde{\beta}$.*

Outline

Generalized Whittle–Matérn fields on $\mathcal{D} \subsetneq \mathbb{R}^d$

Next, we consider Gaussian measures on $L_2(\mathcal{D})$, $\mathcal{D} \subsetneq \mathbb{R}^d$, of the form

$$\mu_d(m; \beta, \boldsymbol{a}, \kappa) := \mathrm{N}\big(m, L^{-2\beta}\big),$$

where $m \in L_2(\mathcal{D})$, $\beta > d/4$ and

$$Lv = -\nabla \cdot (\boldsymbol{a} \nabla v) + \kappa^2 v, \qquad v \in \mathscr{D}(L) \subseteq L_2(\mathcal{D}) \cap H_0^1(\mathcal{D}).$$

We suppose that $\boldsymbol{a}$ and κ and the domain $\mathcal{D} \subset \mathbb{R}^d$ satisfy the following.

Setting WM

i. $\boldsymbol{a} : \overline{\mathcal{D}} \to \mathbb{R}^{d \times d}$ is symmetric and uniformly positive definite, i.e.,

$$\exists a_0 > 0 : \quad \forall \xi \in \mathbb{R}^d : \quad \operatorname{ess\,inf}_{s \in \mathcal{D}} \xi^\top \boldsymbol{a}(s) \xi \geq a_0 \|\xi\|^2_{\mathbb{R}^d},$$

and $\boldsymbol{a} = (\boldsymbol{a}_{jk})_{j,k=1}^d$ is smooth, $\boldsymbol{a}_{jk} \in C^\infty(\overline{\mathcal{D}})$ for all $j, k \in \{1, \ldots, d\}$.

ii. $\kappa : \overline{\mathcal{D}} \to \mathbb{R}$ is smooth, $\kappa \in C^\infty(\overline{\mathcal{D}})$.

iii. The domain $\mathcal{D} \subset \mathbb{R}^d$ has a smooth boundary $\partial\mathcal{D}$ of class C^∞.

Outline

Lemma (Cameron–Martin space for $\mu_d(m; \beta, \boldsymbol{a}, \kappa)$)

Suppose Setting WM. For every $\beta > d/4$, *the* Cameron–Martin space *of the Gaussian measure* $\mu_d(m; \beta, \boldsymbol{a}, \kappa)$ *is*

$$\mathcal{C}^{1/2}(L_2(\mathcal{D})) = \dot{H}_L^{2\beta} := \mathscr{D}\big(L^\beta\big) \subseteq L_2(\mathcal{D}), \qquad \|v\|_{2\beta,L} := \|L^\beta v\|_{L_2(\mathcal{D})}.$$

It is a subspace of the Sobolev space $H^{2\beta}(\mathcal{D})$ *and*

$$\big(\dot{H}_L^{2\beta}, \|\cdot\|_{2\beta,L}\big) \hookrightarrow \big(H^{2\beta}(\mathcal{D}), \|\cdot\|_{H^{2\beta}(\mathcal{D})}\big) \hookrightarrow \big(C^0(\overline{\mathcal{D}}), \|\cdot\|_{C^0(\overline{\mathcal{D}})}\big)$$

Furthermore, provided that $2\beta \notin \mathscr{E}$, *where*

$$\mathscr{E} := \{2k + 1/2 : k \in \mathbb{N}_0\},$$

we have the identification

$$\dot{H}_L^{2\beta} = \Big\{v \in H^{2\beta}(\mathcal{D}) : \big(\kappa^2 - \nabla\cdot(\boldsymbol{a}\nabla)\big)^j v = 0 \text{ in } L_2(\partial\mathcal{D}) \;\; \forall j = 0, \dots, \big\lfloor \beta - \tfrac{1}{4} \big\rfloor\Big\},$$

and on $\dot{H}_L^{2\beta}$ *the Sobolev norm* $\|\cdot\|_{H^{2\beta}(\mathcal{D})}$ *and* $\|\cdot\|_{2\beta,L}$ *are* equivalent.

Theorem (Isomorphic Cameron–Martin spaces)

Suppose Setting WM for each of the parameter tuples $(\mathbf{a}, \kappa)$, $(\widetilde{\mathbf{a}}, \widetilde{\kappa})$ *and for* $\mathcal{D} \subsetneq \mathbb{R}^d$. *Let* $\beta, \widetilde{\beta} > d/4$ *be such that* $2\beta \notin \mathscr{E}$.

The Cameron–Martin spaces *of two Gaussian measures*

$$\mu_d(0; \beta, \mathbf{a}, \kappa) \quad \text{and} \quad \mu_d(0; \widetilde{\beta}, \widetilde{\mathbf{a}}, \widetilde{\kappa})$$

are isomorphic with equivalent norms *if and only if* $\beta = \widetilde{\beta}$ *and for every* $j \in \mathbb{N}_0$ *with* $j \leq \lfloor \beta - 5/4 \rfloor$ *the following hold:*

$$\begin{aligned} &\forall v \in \dot{H}_L^{2\beta}: \quad \left(\kappa^2 - \nabla \cdot (\mathbf{a}\nabla)\right)^j \left(\delta_{\kappa^2} - \nabla \cdot (\delta_{\mathbf{a}} \nabla)\right) v = 0 \ \text{ in } \ L_2(\partial\mathcal{D}), \\ &\forall \widetilde{v} \in \dot{H}_{\widetilde{L}}^{2\beta}: \quad \left(\widetilde{\kappa}^2 - \nabla \cdot (\widetilde{\mathbf{a}}\nabla)\right)^j \left(\delta_{\kappa^2} - \nabla \cdot (\delta_{\mathbf{a}} \nabla)\right) \widetilde{v} = 0 \ \text{ in } \ L_2(\partial\mathcal{D}). \end{aligned} \qquad \text{(BCs)}$$

Here, $\delta_{\kappa^2}(s) := \widetilde{\kappa}^2(s) - \kappa^2(s)$ *and* $\delta_{\mathbf{a}}(s) := \widetilde{\mathbf{a}}(s) - \mathbf{a}(s)$ *for all* $s \in \overline{\mathcal{D}}$.

⇒ **The behavior of δ_{κ^2} and $\delta_{\mathbf{a}}$ on the boundary $\partial\mathcal{D}$ matters!**

Outline

Conditions on the parameters for equivalent measures

Theorem (Characterizing equivalence of the measures μ and $\widetilde{\mu}$)

Suppose Setting WM for each of the parameter tuples $(\boldsymbol{a}, \kappa)$, $(\widetilde{\boldsymbol{a}}, \widetilde{\kappa})$ and for $\mathcal{D} \subsetneq \mathbb{R}^d$. Let $m, \widetilde{m} \in L_2(\mathcal{D})$ and $\beta, \widetilde{\beta} > d/4$ be such that $2\beta \notin \mathscr{E}$.

- *In dimension $d \leq 3$, the Gaussian measures $\mu_d(m; \beta, \boldsymbol{a}, \kappa)$ and $\mu_d(\widetilde{m}; \widetilde{\beta}, \widetilde{\boldsymbol{a}}, \widetilde{\kappa})$ are* equivalent *if and only if*
  - $\beta = \widetilde{\beta}$,
  - *the* boundary conditions (BCs) *hold for every $j = 0, 1, \ldots, \lfloor \beta - 5/4 \rfloor$,*
  - $m - \widetilde{m} \in \dot{H}_L^{2\beta}$,
  - $\boldsymbol{a} = \widetilde{\boldsymbol{a}}$.
- *In dimension $d \geq 4$, the Gaussian measures $\mu_d(m; \beta, \boldsymbol{a}, \kappa)$ and $\mu_d(\widetilde{m}; \widetilde{\beta}, \widetilde{\boldsymbol{a}}, \widetilde{\kappa})$ are* equivalent *if and only if*
  - $\beta = \widetilde{\beta}$,
  - $\kappa^2 = \widetilde{\kappa}^2$,
  - $m - \widetilde{m} \in \dot{H}_L^{2\beta}$,
  - $\boldsymbol{a} = \widetilde{\boldsymbol{a}}$.

Conditions on the parameters for optimal linear prediction

Theorem (Uniformly asymptotically optimal linear prediction)

Suppose Setting WM for each of the parameter tuples $(\boldsymbol{a}, \kappa)$, $(\widetilde{\boldsymbol{a}}, \widetilde{\kappa})$ *and for* $\mathcal{D} \subsetneq \mathbb{R}^d$*. Let* $m, \widetilde{m} \in L_2(\mathcal{D})$ *and* $\beta, \widetilde{\beta} > d/4$ *be such that* $2\beta \notin \mathscr{E}$*.*

Consider the Gaussian measures

$$\mu_d(m; \beta, \boldsymbol{a}, \kappa) \quad \text{and} \quad \mu_d(\widetilde{m}; \widetilde{\beta}, \widetilde{\boldsymbol{a}}, \widetilde{\kappa})$$

Then, the Assumptions I–III *(for uniformly asymptotically optimal linear prediction) are satisfied if and only if*

- $\beta = \widetilde{\beta}$,
- *the* boundary conditions (BCs) *hold for every* $j = 0, 1, \ldots, \lfloor \beta - 5/4 \rfloor$,
- $m - \widetilde{m} \in \dot{H}_L^{2\beta}$,
- *here exists a constant* $c > 0$ *such that* $c\boldsymbol{a} = \widetilde{\boldsymbol{a}}$*.*

Outline

Summary

Let $\mu := \mu_d(0; \beta, \mathbf{a}, \kappa)$, $\widetilde{\mu} := \mu_d(0; \widetilde{\beta}, \widetilde{\mathbf{a}}, \widetilde{\kappa})$ be Gaussian measures for generalized Whittle–Matérn fields with parameters $(\beta, \mathbf{a}, \kappa)$ resp. $(\widetilde{\beta}, \widetilde{\mathbf{a}}, \widetilde{\kappa})$.

	Interval for β, assuming that $\beta \notin \{k + 1/4 : k \in \mathbb{N}\}$		
Conditions for	$(d/4, 9/4)$	$(9/4, 13/4)$	$(13/4, \infty)$
Asymptotically optimal linear prediction	$\beta = \widetilde{\beta}$, $c\mathbf{a} = \widetilde{\mathbf{a}}$ for some $c \in (0, \infty)$	$\beta = \widetilde{\beta}$, $c\mathbf{a} = \widetilde{\mathbf{a}}$, $\left(\mathbf{a}\nabla\delta_{c,\kappa^2}\right)\big\vert_{\partial\mathcal{D}} \cdot \mathbf{n} = 0$	$\beta = \widetilde{\beta}$, $c\mathbf{a} = \widetilde{\mathbf{a}}$ + b.c. on δ_{c,κ^2}
Equivalence of measures in dimension $d \leq 3$	$\beta = \widetilde{\beta}$, $\mathbf{a} = \widetilde{\mathbf{a}}$	$\beta = \widetilde{\beta}$, $\mathbf{a} = \widetilde{\mathbf{a}}$, $\left(\mathbf{a}\nabla\delta_{1,\kappa^2}\right)\big\vert_{\partial\mathcal{D}} \cdot \mathbf{n} = 0$	$\beta = \widetilde{\beta}$, $\mathbf{a} = \widetilde{\mathbf{a}}$ + b.c. on δ_{1,κ^2}
Equivalence of measures in dimension $d \geq 4$	$\beta = \widetilde{\beta}$, $\mathbf{a} = \widetilde{\mathbf{a}}$, $\kappa^2 = \widetilde{\kappa}^2$		

$\delta_{c,\kappa^2}(s) := \widetilde{\kappa}^2(s) - c\kappa^2(s)$, and $\mathbf{n}$ is the outward pointing normal on $\partial\mathcal{D}$.

Simulation 1: The difference between κ^2 and a

$$\left(-\nabla\cdot(\boldsymbol{a}\nabla)+\kappa^2\right)^{\beta}(\tau Z)=\mathcal{W} \quad \text{in } \mathcal{D}=(0,1).$$

True model: $\beta = 1$, $\boldsymbol{a} \equiv 1$, $\kappa^2 \equiv 1200$, and $\tau = \frac{1}{2}\kappa^{-3/2}$.

Two misspecified models: correct values of β, τ, and we set

$$\left(\kappa^2(s), \boldsymbol{a}(s)\right) = \begin{cases} \left(1200 f(s)^{-1}, 1\right) & \text{for model 1,} \\ \left(1200, f(s)\right) & \text{for model 2,} \end{cases} \qquad s \in \overline{\mathcal{D}} = [0,1],$$

where $f(s) := 1 + \frac{1}{2}\,\mathrm{erf}\left(\frac{\delta(s-0.5)}{\sqrt{2}}\right)$.

As a measure of accuracy, we use

$$\mathcal{E}_n(h) := \frac{\mathrm{E}\left[(\widetilde{h}_n - h)^2\right]}{\mathrm{E}\left[(h_n - h)^2\right]} - 1.$$

We predict $Z(x_0)$ as well as the integrals $I_\ell := (Z, e_\ell)_{L_2(\mathcal{D})}$ and set

$$\mathcal{E}_{I,n}^{\max} := \max\left\{\mathcal{E}_{I,n}^{\ell} : n+1 \le \ell \le N\right\}, \qquad \mathcal{E}_{I,n}^{\ell} := \mathcal{E}_n(I_\ell), \quad \ell \in \{n+1, \dots, N\}.$$

Results of Simulation 1

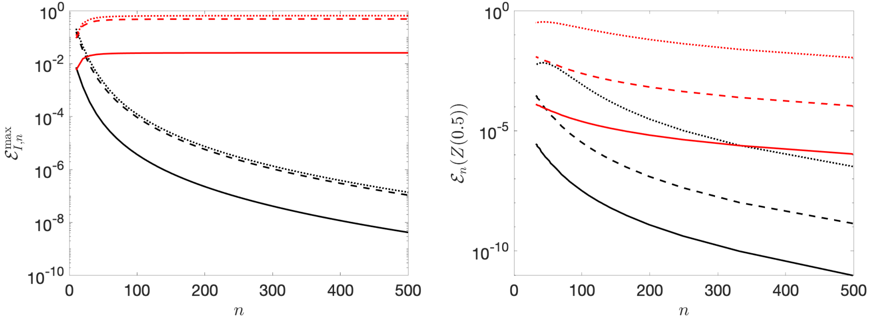


Figure: The results for model 1 (black) and model 2 (red) for the first example with integral observations (left) and point observations (right).

Solid lines correspond to $\delta = 1$, dashed to $\delta = 10$, and dotted to $\delta = 100$.

Simulation 2: The effect of the smoothness parameter β

For $\beta \in \{1, 2, 3\}$,we consider

$$\left(-\nabla \cdot (\boldsymbol{a}\nabla) + \kappa^2\right)^{\beta}(\tau Z) = \mathcal{W} \quad \text{in } \mathcal{D} = (0,1).$$

True model: $\boldsymbol{a} \equiv 1$, $\kappa^2 \equiv 100(4\beta - 1)$, and $\tau > 0$.

Two misspecified models: correct values of $\beta, \boldsymbol{a}, \tau$, and we set

$$\kappa^2(s) = 100(4\beta - 1) \cdot \begin{cases} 1 - 1.5s^2 + s^3 & \text{for model 1,} \\ 1 + s - 1.5s^3 & \text{for model 2,} \end{cases} \qquad s \in \overline{\mathcal{D}} = [0,1].$$

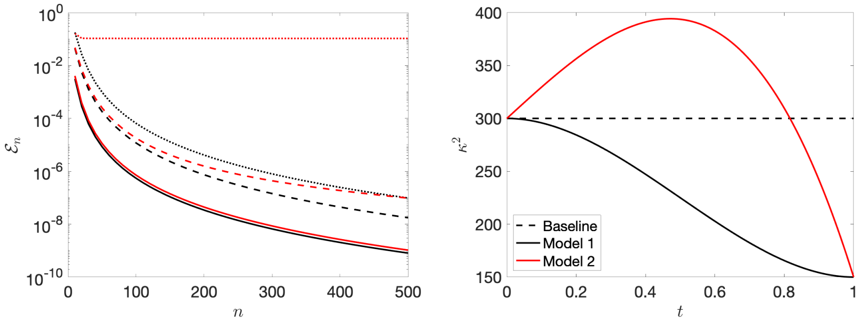
We again predict the integrals $I_\ell := (Z, e_\ell)_{L_2(\mathcal{D})}$ and consider

$$\mathcal{E}_{I,n}^{\max} := \max\left\{\mathcal{E}_{I,n}^{\ell} : n+1 \leq \ell \leq N\right\}, \qquad \mathcal{E}_{I,n}^{\ell} := \mathcal{E}_n(I_\ell), \quad \ell \in \{n+1, \ldots, N\},$$

where

$$\mathcal{E}_n(h) := \frac{\mathrm{E}\left[(\widetilde{h}_n - h)^2\right]}{\mathrm{E}\left[(h_n - h)^2\right]} - 1.$$

Results of Simulation 2



Left: The results for model 1 (black) and model 2 (red) in the second example, with $\beta = 1$ (solid), $\beta = 2$ (dashed), and $\beta = 3$ (dotted).

Right: κ^2 for the two models when $\beta = 1$.

References

K. Kirchner and D. Bolin (2020). Necessary and sufficient conditions for asymptotically optimal linear prediction of random fields on compact metric spaces. To appear in: *Annals of Statistics*. Preprint: arXiv:2005.08904.

D. Bolin and K. Kirchner (2021). Equivalence of measures and asymptotically optimal linear prediction for Gaussian random fields with fractional-order covariance operators. Preprint: arXiv:2101.07860.

Thank you for your attention!