MASCOT-NUM 2022
Clermont Ferrand, France
7-9 June 2022

Physics-informed random fields. Application to Kriging.

I. Henderson
Université Toulouse III-Paul Sabatier

Supervisor(s): P. Noble (Univ. Paul Sabatier), O. Roustant (Univ. Paul Sabatier), R. Baraille (SHOM)

PhD expected duration: Oct. 2020 - Sep. 2023

Address: 135 avenue de Rangueil, 31000 Toulouse, France

E-mail: henderso@insa-toulouse.fr

Abstract: The use of "physics-informed" Gaussian process regression (GPR) models has become more and more popular since their introduction in the early 2000'. An important part of these models are designed to deal with homogeneous linear partial differential equations (PDEs)

$$L(u) := \sum_{|\alpha|\leq n} a_\alpha(x)\partial^\alpha u = 0 \tag{1}$$

Above, u is the unknown function to be approximated, defined over an open set $D \subset \mathbb{R}^d$, and L is a linear partial differential operator. In (1), for a multi-index $\alpha = (\alpha_1, ..., \alpha_d)^T \in \mathbb{N}^d$, we used the notations $|\alpha| = \alpha_1 + ... + \alpha_d$ and $\partial^\alpha = (\partial_{x_1})^{\alpha_1}...(\partial_{x_d})^{\alpha_d}$. Starting from (1), one models u as a sample path of a Gaussian process (GP) $U = (U(x))_{x\in D} \sim GP(0, k_u)$ and draws the consequences of (1) on the covariance structure of U. For general linear operators, the converse is expected to hold : enforcing the linear constraints on the sample paths of U can be done by enforcing the linear constraints on the functions $k_u(x, \cdot)$. Under the assumptions that U is a GP with n times differentiable sample paths, [1] proves this property for some classes of differential operators.

In the standard PDE approach though, equation (1) is reinterpreted by weakening the definition of the derivatives of u, thereby weakening the required regularity assumptions over u. It can indeed happen in practice, such as with hyperbolic PDEs, that the sought solutions of the PDE $L(u) = 0$ are not n times differentiable; they are only solutions of some weakened formulation of equation (1). We focus here on the distributional formulation of the PDE (1), where the regularity assumptions over u are relaxed to the maximum. Consider equation (1), multiply it by a compactly supported, smooth test function $\varphi \in C_c^\infty(D)$ and integrate over D. For each integral term $\int_D \varphi(x)a_\alpha(x)\partial^\alpha u(x)dx$, perform $|\alpha|$ successive integrations by parts to transfer the derivatives from u to φ. Since $\varphi \in C_c^\infty(D)$, we have that

$$\forall\varphi \in C_c^\infty(D), \ \int_D u(x) \sum_{|\alpha|\leq n} (-1)^{|\alpha|}\partial^\alpha(a_\alpha\varphi)(x)dx = 0 \tag{2}$$

One only requires that $u \in L^1_{loc}(D)$, i.e. $\int_K |u(x)|dx < +\infty$ for all compact set $K \subset D$, to make sense of equation (2). We then say that $u \in L^1_{loc}(D)$ is a solution to $L(u) = 0$ in the distributional sense if u verifies (2). Under the weak assumptions that U is a measurable centered second order random field and that $\sigma : x \longmapsto \sqrt{k_u(x,x)} \in L^1_{loc}(D)$, we are able to show that [2]

$$\mathbb{P}(L(U) = 0 \text{ in the distrib. sense}) = 1 \iff \forall x \in D, L(k_u(x,\cdot)) = 0 \text{ in the distrib. sense} \tag{3}$$

This extends results from [1], and comes in handy for understanding GP models for PDEs. As a prototype for hyperbolic PDEs, we examine the following wave equation in $\mathbb{R}^3$. Note $\Delta = \partial^2_{xx} + \partial^2_{yy} + \partial^2_{zz}$:

$$\begin{cases} \Big(\dfrac{1}{c^2}\partial^2_{tt} - \Delta\Big)u = 0 \quad \forall(x,t) \in \mathbb{R}^3 \times \mathbb{R}^*_+ & (4)\\ u(x,0) = u_0(x) \quad \text{and} \quad (\partial_t u)(x,0) = v_0(x) \quad \forall x \in \mathbb{R}^3 & (5)\end{cases}$$

Its distributional solution u is represented as

$$u(x,t) = (F_t * v_0)(x) + (\dot{F}_t * u_0)(x) \qquad \forall(x,t) \in \mathbb{R}^3 \times \mathbb{R}_+ \tag{6}$$

Above, F_t is a multiple of the Lebesgue measure over the sphere of radius ct and $\dot{F}_t$ is its time derivative; it is only a generalized function in the sense of L. Schwartz. Equation (6) can be expressed explicitly thanks to convolutions over the unit sphere $S(0,1)$ involving v_0, u_0 and ∇u_0 ; this is the Kirschoff formula. From there, one can see that when u_0 and v_0 are not smooth enough (e.g. no more than $v_0 \in C^0(\mathbb{R}^3)$ or $u_0 \in C^1(\mathbb{R}^3)$), u is not of class C^2 and u does not verify the PDE (4) pointwise.

Suppose now that u_0 and v_0 are sample paths of two centered independent GPs $U_0 \sim GP(0, k_u^0)$ and $V_0 \sim GP(0, k_v^0)$. We have shown in ([2]) that u in equation (6) is a sample path of a centered GP $U \sim GP(0, k_u)$, whose covariance kernel can be expressed in a compact way with tensor products and convolutions:

$$k_u((x,t),(x',t')) = [(F_t \otimes F_{t'}) * k_v^0](x,x') + [(\dot{F}_t \otimes \dot{F}_{t'}) * k_u^0](x,x') \tag{7}$$

An explicit spherical convolution expression can also be derived for (7). One is then able to show that the right-hand side of (3) is verified for k_u and thus the sample paths of U verify the wave equation in the sense of distributions, though not pointwise in general. The kernel (7) can then be used for GPR on pointwise observations of a solution u of (4). In particular, evaluating the corresponding Kriging mean $m_K(x,t)$ or its time-derivative at $t = 0$ provides a reconstruction of u_0 and/or v_0. We show in Figure 1 an example of such a reconstruction (described in [2]). Given v_0, we numerically simulate the corresponding solution u on the domain $[0,1]^3$. Having scattered 25 artificial sensors in $[0,1]^3$, we obtain a database comprised of 25 time series, one for each captor, on which we perform Kriging with (7). The images in Figure 1 correspond to the 3D functions $v_0(x)$ and $m_K(x,0)$ evaluated on a slice $z = Cst$.

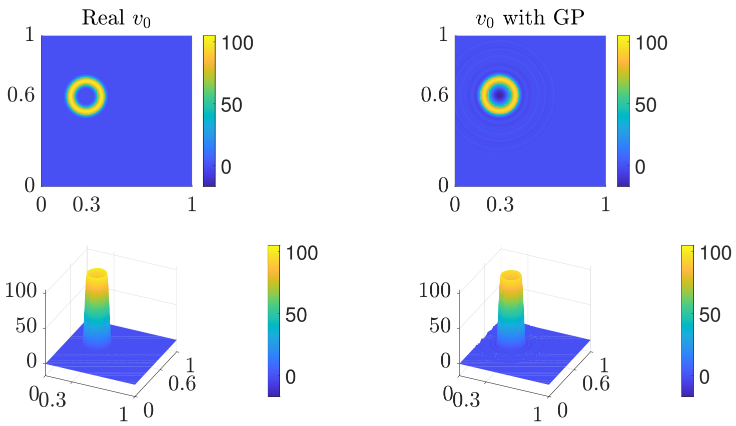

Figure 1: Reconstruction of an initial speed v_0 using kernel (7)

References

[1] David Ginsbourger, Olivier Roustant, and Nicolas Durrande. On degeneracy and invariances of random fields paths with applications in Gaussian process modelling. *Journal of Statistical Planning and Inference*, 170:117–128, 2016.

[2] Iain Henderson, Pascal Noble, and Olivier Roustant. Stochastic processes under linear differential constraints : Application to Gaussian process regression for the 3 dimensional free space wave equation. *arXiv*, 2021.

Short biography – Iain Henderson holds an engineering degree from CentraleSupélec and a Master degree in PDEs from the Université d'Orsay (Master AMS). The PhD is funded by the SHOM (Service Hydrographique et Océanographique de la Marine) and is motivated by using machine learning techniques and/or surrogate models for enhancing tidal wave forecasting.