Physics-informed Kriging with applications to inverse problems involving PDEs

I. Henderson[†,1], P. Noble[§,1], O. Roustant[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2020 – Sep. 2023

[1] Institut de Mathématiques de Toulouse (IMT); INSA Toulouse, France
{henderso, noble, roustant}@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'. Some of these models in particular aim to approximate functions $u : D \to \mathbb{R}$, where $D \subset \mathbb{R}^d$ is an open set, which are solutions of a given homogeneous linear partial differential equation (PDE), i.e. an equation of the form

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

Above, given $\alpha = (\alpha_1, \ldots, \alpha_d)^T \in \mathbb{N}^d$, we denoted $|\alpha| := \alpha_1 + \ldots + \alpha_d$ and $\partial^\alpha := (\partial_{x_1})^{\alpha_1} \ldots (\partial_{x_d})^{\alpha_d}$. Starting from (1), one models u as a realization of a Gaussian process (GP) $U = (U(x))_{x \in D} \sim GP(0, k_u)$ and draws the consequences of (1) on the kernel k_u. For general linear operators, it is expected that enforcing the linear constraints on the sample paths of U is ensured by enforcing the linear constraints on the functions $k_u(x, \cdot)$. When 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. 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), which relaxes the regularity assumptions over u 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 $\varphi \in C_c^\infty(D)$, this yields 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 prove in [2] that

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

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 $\mathbb{R}^3$. Denote

$\Delta = \partial^2_{xx} + \partial^2_{yy} + \partial^2_{zz}$ the three dimensional Laplacian and consider the following PDE

$$\Big(\frac{1}{c^2}\partial^2_{tt} - \Delta\Big)u = 0, \quad (x,t) \in \mathbb{R}^3 \times \mathbb{R}^*_+, \tag{4}$$

given the initial data $u(x,0) = u_0(x)$ and $(\partial_t u)(x,0) = v_0(x)$. Its distributional solution u is

$$u(x,t) = (F_t * v_0)(x) + (\dot{F}_t * u_0)(x), \quad \forall (x,t) \in \mathbb{R}^3 \times \mathbb{R}_+. \tag{5}$$

Above, F_t and $\dot{F}_t$ are not functions but distributions in the sense of L. Schwartz. Equation (5) can be expressed as convolutions over the unit sphere $S(0,1)$; this is the Kirschoff formula. From this formula, one sees that when u_0 and v_0 are not smooth enough (e.g. u_0 is of class C^1 and no more), u is not of class C^2 and thus does not verify the PDE (4) pointwise.

Suppose now that u_0 and v_0 are realizations of two independent GPs $U_0 \sim GP(0, k^0_u)$ and $V_0 \sim GP(0, k^0_v)$. We show in [2] that u in equation (5) is then a realization of a GP $U \sim GP(0, k_u)$, whose kernel can be expressed in a compact way with tensor products and convolutions:

$$k_u((x,t),(x',t')) = [(F_t \otimes F_{t'}) * k^0_v](x,x') + [(\dot{F}_t \otimes \dot{F}_{t'}) * k^0_u](x,x'). \tag{6}$$

We next show that the right-hand side of (3) is verified for k_u and thus the realizations of U verify the wave equation in the distributional sense, though not pointwise in general. The kernel (6) 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. Figure 1 shows an example of such a reconstruction. Additionally, physical parameters such as the velocity c or the source position parameters can be viewed as hyperparameters of the kernel (6) and estimated via marginal likelihood optimization. Interestingly, for point sources, this method for estimating the source position naturally reduces to the triangulation method used in GPS systems (Figure 2). We detail these examples in [2].

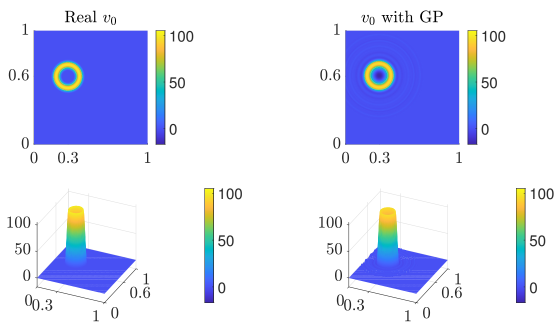

Figure 1: Reconstruction of an initial speed v_0 using kernel (6) on a slice $z = Cst$.

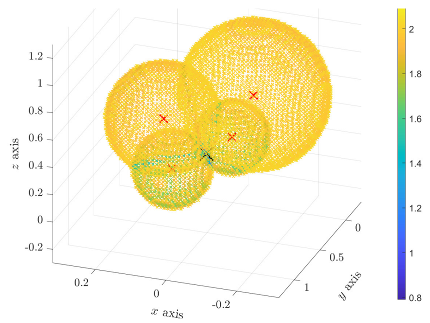

Figure 2: Level sets of the negative log marginal likelihood, as a function of the source position, in the point source limit.

Short biography (PhD student)

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

References

[1] D. Ginsbourger, O. Roustant, and N. Durrande. On degeneracy and invariances of random fields paths with applications in Gaussian process modelling. *J. Statist. Plann. Inference*, 170:117–128, 2016.

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