Journal of Statistical Planning and Inference 170 (2016) 117–128

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

On degeneracy and invariances of random fields paths with applications in Gaussian process modelling

David Ginsbourger [a,b,*], Olivier Roustant [c], Nicolas Durrande [c]

[a] Idiap Research Institute, Martigny, Switzerland
[b] IMSV, Department of Mathematics and Statistics, University of Bern, Switzerland
[c] Mines Saint-Etienne, UMR CNRS 6158, LIMOS, F-42023 Saint-Etienne, France

ARTICLE INFO

Article history:
Received 12 May 2015
Received in revised form 24 August 2015
Accepted 4 October 2015
Available online 23 October 2015

MSC:
60G60
60G17
62J02

Keywords:
Covariance kernels
Linear operators
RKHS
Structural priors

ABSTRACT

We study pathwise invariances and degeneracies of random fields with motivating applications in Gaussian process modelling. The key idea is that a number of structural properties one may wish to impose a priori on functions boil down to degeneracy properties under well-chosen linear operators. We first show in a second order set-up that almost sure degeneracy of random field paths under some class of linear operators defined in terms of signed measures can be controlled through the two first moments. A special focus is then put on the Gaussian case, where these results are revisited and extended to further linear operators thanks to state-of-the-art representations. Several degeneracy properties are tackled, including random fields with symmetric paths, centred paths, harmonic paths, or sparse paths. The proposed approach delivers a number of promising results and perspectives in Gaussian process modelling. In a first numerical experiment, it is shown that dedicated kernels can be used to infer an axis of symmetry. Our second numerical experiment deals with conditional simulations of a solution to the heat equation, and it is found that adapted kernels notably enable improved predictions of non-linear functionals of the field such as its maximum.

1. Introduction

Whether for function approximation, classification, or density estimation, probabilistic models relying on random fields have been increasingly used in recent works from various research communities. Finding their applied roots in geostatistics and spatial statistics with optimal linear prediction and Kriging (Matheron, 1963; Stein, 1999), random field models for prediction have become a main stream topic in machine learning (under the *Gaussian Process Regression* terminology, see, e.g., Rasmussen and Williams, 2006), with a spectrum ranging from metamodeling and adaptive design approaches in science and engineering Welch et al. (1992), Jones (2001), O'Hagan (2006) to theoretical Bayesian statistics in function spaces (see Van der Vaart and Van Zanten, 2008a, Van der Vaart and Van Zanten, 2008b, Van der Vaart and van Zanten, 2011 and references therein).

Often, a Gaussian random field model is assumed for some function f of interest, and so all prior assumptions on f are accounted for by the corresponding mean function m and covariance kernel k. The choice of m and k should thus reflect as much as possible any prior belief the modeller wishes to incorporate in the model. Such prior belief on f

* Corresponding author at: IMSV, Department of Mathematics and Statistics, University of Bern, Switzerland.
E-mail addresses: ginsbourger@stat.unibe.ch (D. Ginsbourger), roustant@emse.fr (O. Roustant), durrande@emse.fr (N. Durrande).

http://dx.doi.org/10.1016/j.jspi.2015.10.002
0378-3758/

may of course include classical regularity properties in the first place (continuity, differentiability, Hölder regularity, etc.), but also more specific properties such as symmetries (Haasdonk and Burkhardt, 2007; Ginsbourger et al., 2012), sparse functional ANOVA decompositions (Duvenaud et al., 2011; Durrande et al., 2012; Ginsbourger et al., 2014), or degeneracy under multivariate differential operators in the case of vector-valued random fields. To take a concrete example, covariance structures characterizing divergence-free and curl-free random vector fields have been recently presented and illustrated in Scheuerer and Schlather (2012). Besides that, the idea of expressing structure with kernels has been explored in Duvenaud (2014), where a number of practical aspects regarding positive-semidefiniteness-preserving operations are addressed.

Here we shall discuss how the two first moments influence mathematical properties of associated realizations (or *paths*), both in a general second order set-up and in the Gaussian case. A number of well-known random field properties driven by the covariance kernel are in the mean square sense, e.g. L^2 continuity and differentiability (Cressie, 1993). However, such results generally are not informative about the pathwise behaviour of underlying random fields. On the other hand, much can be said about path regularity properties of random field paths (see, e.g., classical results in Cramér and Leadbetter, 1967, Adler, 1990), based in particular on the behaviour of the covariance kernel in the neighbourhood of the diagonal in the second order case. In the stationary case, it is then sufficient to look at the covariance function in the neighbourhood of the origin (with similar results for the variogram in the intrinsic stationary – but not necessarily second order – case). More recently, Scheuerer (2010) has taken a new look at path regularity of second-order random fields, and drew conclusions about a.s. continuous differentiability in non-Gaussian settings. Also, we refer to Scheuerer (2009) for an enlightening exposition of state-of-the-art results concerning regularity properties of random field sample paths in various frameworks.

Our focus in the present work is on *pathwise* mathematical properties of second order random fields and statistical applications thereof in the context of Gaussian process modelling. Motivated by several practical situations, we pay a particular attention to random fields $Z = (Z_{\mathbf{x}})_{\mathbf{x}\in D}$ that are supported by the null space of some linear operator T, i.e. for which

$$T(Z) = \mathbf{0} \quad \text{(a.s.)}. \tag{1}$$

As we first develop in general second-order settings, an impressive diversity of path properties including invariances under group actions or sparse ANOVA decompositions of multivariate paths can be encapsulated in the framework of Eq. (1). Furthermore, in the particular case of Gaussian random fields, a more general class of path properties (notably some degeneracy properties involving differential operators) can be covered through the link between operators on the paths and operators on the *reproducing kernel Hilbert space* (Berlinet and Thomas-Agnan, 2004) associated with the random field, and also through an additional representation of Z in terms of Gaussian measures on Banach spaces.

While Section 2 is dedicated to the exposition of the main results, proofs are presented in the Appendix to ease the reading. Applications in the context of random field modelling, and especially for Gaussian process modelling, are then investigated throughout Section 3. In particular, we tackle zero-integral random processes, random fields with paths invariant under group actions, random fields with additive paths, random fields with harmonic paths, and discuss further potential applications.

In Section 4, we present two original numerical experiments where the notions of degeneracy and invariance appear very useful in Gaussian process modelling under two types of structural prior information. In the first case, the objective function possesses an unknown axis of symmetry, which is inferred by maximum likelihood, relying on a family of argumentwise invariant covariance kernels. In the second case, we obtain an improved interpolation of a solution to the heat equation thanks to a bi-harmonic kernel. The proposed model enables performing harmonic conditional simulations, which has very beneficial consequences in terms of estimation of the maximum. Section 5 is dedicated to conclusions and perspectives. The main results are finally proven in the Appendix.

2. Main results

Let $(D, \mathcal{D})$ be a measurable space, $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space, and $Z = (Z_{\mathbf{x}})_{\mathbf{x}\in D}$ be a measurable real-valued stochastic process over $(\Omega, \mathcal{A}, \mathbb{P})$. Let us further assume that the paths of Z belong with probability 1 to some function space $\mathcal{F} \subset \mathcal{M}(D, \mathbb{R})$, where $\mathcal{M}(D, \mathbb{R})$ is the set of $(\mathcal{D}, \mathcal{B}(\mathbb{R}))$-measurable functions, and consider a linear operator $T : \mathcal{F} \longrightarrow \mathcal{F}$. Here both Z and $T(Z)$ are assumed second order, in the sense that their marginals possess a variance, and we aim at giving necessary and sufficient conditions in terms of the two first moments of Z for the following *degeneracy* to hold:

$$\mathbb{P}(T(Z) = \mathbf{0}) = \mathbb{P}(\forall \mathbf{x} \in D\; T(Z)_{\mathbf{x}} = 0) = 1. \tag{2}$$

We prove that in a variety of settings on T and Z, this is equivalent to having that both m and k are in the null space of T in a sense to be discussed next. In Section 2.1 we discuss equivalent conditions that do not involve any distributional assumption, and we obtain a characterization of degeneracy under a specific class of operators that prove useful for applications in Sections 3 and 4. In Section 2.2 we generalize the results to a wider class of operators T in the specific framework of Gaussian processes and Gaussian measures.