Journal of Multivariate Analysis 146 (2016) 84–94

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Kriging for Hilbert-space valued random fields: The operatorial point of view

Alessandra Menafoglio [a,*], Giovanni Petris [b]

[a] MOX, Department of Mathematics, Politecnico di Milano, Italy

[b] Department of Mathematical Sciences, University of Arkansas, USA

ARTICLE INFO

Article history:
Received 30 December 2014
Available online 25 June 2015

AMS subject classifications:
60G15
60G25
60G60
62F10
62H11
62H20
62M20
62M30
62M40

Keywords:
Geostatistics
Gaussian processes
Conditional expectations
Measurable linear transformations

ABSTRACT

We develop a comprehensive framework for linear spatial prediction in Hilbert spaces. We explore the problem of Best Linear Unbiased (BLU) prediction in Hilbert spaces through an original point of view, based on a new Operatorial definition of Kriging. We ground our developments on the theory of Gaussian processes in function spaces and on the associated notion of measurable linear transformation. We prove that our new setting allows (a) to derive an explicit solution to the problem of Operatorial Ordinary Kriging, and (b) to establish the relation of our novel predictor with the key concept of conditional expectation of a Gaussian measure. Our new theory is posed as a unifying theory for Kriging, which is shown to include the Kriging predictors proposed in the literature on Functional Data through the notion of finite-dimensional approximations. Our original viewpoint to Kriging offers new relevant insights for the geostatistical analysis of either finite- or infinite-dimensional georeferenced dataset.

1. Introduction

In recent years, the increasing availability of complex and high-dimensional data has motivated a fast and extensive growth of Functional Data Analysis (FDA, e.g., [28]) and Object Oriented Data Analysis (OODA, e.g., [21], and references therein). These new branches of statistics share the same abstract approach in interpreting each datum as a realization of a random element in a finite- or infinite-dimensional space. Properties of the space to which data are assumed to belong directly reflect on the methodologies that one can employ for the statistical analysis. For instance, the geometry of a Hilbert space allows for a class of methods based on the notions of inner product and norm (e.g., [4], and references therein), whereas methods suitable for data in general metric spaces need to rely on the notion of distance only.

A rich body of literature has been devoted to the statistical analysis of functional data. Most works in this field rely upon the embedding of the data into a Hilbert space, particularly L^2, to develop parametric or non-parametric methods for their treatment (e.g., [28,14,17]). The Hilbert space embedding allows for the generalization to the functional setting of several well-known multivariate methods, such as principal component analysis (e.g., [28]), K-mean clustering (e.g., [31,29]), or hypothesis testing (e.g., via T^2-Hotelling statistics, [27]). However, new issues emerged with the advent of FDA, such as

* Corresponding author.
E-mail addresses: alessandra.menafoglio@polimi.it (A. Menafoglio), GPetris@uark.edu (G. Petris).

http://dx.doi.org/10.1016/j.jmva.2015.06.012
0047-259X/

the problem of data smoothing (e.g., [28]) or curve alignment (i.e., registration, e.g., [33,30], and references therein). For an overview on FDA and its most recent advances we refer to [8,3].

In this framework, a relatively large body of literature addresses the problem of the geostatistical characterization and prediction of spatially dependent functional data. Early works in this field focused on L^2 data to develop linear spatial predictors (i.e., Kriging predictors) in the form of optimal linear combinations of the data (e.g., [12,16,5]). Even though the L^2 embedding is commonly employed in FDA, several environmental applications deal with constrained or manifold data, for which the L^2 geometry may be inappropriate. For instance, Menafoglio et al. [22,24] deal with a set of constrained functional data in the form of particle-size densities, i.e., probability density functions describing the distribution of grains sizes within a given soil sample. In this case, the usual L^2 geometry is not appropriate, as it completely neglects the data constraints (see, e.g., [10,11]).

These elements motivate the adoption of an abstract viewpoint, along the line of OODA. In this setting, Menafoglio et al. [23] establish a Kriging theory for random fields valued in any separable Hilbert space, allowing for the analysis of a broad range of object data, such as curves, surfaces or images. The present work stands in continuity with the approach of Menafoglio et al. [23], with whom we share the geometric viewpoint to the treatment of either finite- or infinite-dimensional data as *atoms* of the geostatistical analysis. However, we here explore the problem of linear spatial prediction in Hilbert spaces through an original point of view, based on a new operatorial definition of Kriging. In this setting, the theory of Operatorial Kriging is posed as a unifying framework for Kriging, with the scope of including either the formulations of Kriging for curves in L^2 (e.g., [12,25]) or that for Hilbert data [23].

The remaining part of this work is organized as follows. Section 2 introduces the problem and highlights the main contributions of this work. Section 3 recalls the theory of Gaussian measures on Hilbert spaces, upon which we ground the developments of Sections 4 and 5. Section 6 investigates discretizations of the Operatorial Kriging predictor, and the relation of our new theory with the existing literature works of Nerini et al. [25] and Menafoglio et al. [23]. Section 7 provides a discussion on the impact of our results from the application viewpoint and Section 8 concludes the work.

2. Kriging for Hilbert data: state of the art and main contributions

We denote by D a d-dimensional spatial domain, and by $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ the locations of the available data $x_{\boldsymbol{s}_1}, \ldots, x_{\boldsymbol{s}_n}$. As in classical geostatistics, we assume that the latter are a partial observation of a random field $\{\mathcal{X}_{\boldsymbol{s}}, \boldsymbol{s} \in D\}$. Throughout this work, we assume that $\{\mathcal{X}_{\boldsymbol{s}}, \boldsymbol{s} \in D\}$ is valued in a separable Hilbert space $\mathcal{H}$, and that it is Gaussian and stationary (in the sense that will be clarified in Sections 3–5). Our aim is the prediction of the element $\mathcal{X}_{\boldsymbol{s}_0}$ at an unobserved location $\boldsymbol{s}_0$ in D.

In this setting, if the Hilbert space $\mathcal{H}$ was the one-dimensional Euclidean space $\mathbb{R}$, classical geostatistics literature would advocate the use of a Kriging predictor, that is the Best Linear Unbiased Predictor (BLUP) $\mathcal{X}^*_{\boldsymbol{s}_0} = \sum_{i=1}^n \lambda_i \mathcal{X}_{\boldsymbol{s}_i}$, whose weights minimize the variance of prediction error under the unbiasedness constraint (e.g., [7]). This can also be interpreted – in the Gaussian setting – in terms of the conditional expectation of $\mathcal{X}_{\boldsymbol{s}_0}$ given $\mathcal{X}_{\boldsymbol{s}_1}, \ldots, \mathcal{X}_{\boldsymbol{s}_n}$.

We note that, in the scalar case, the notion of linear predictor is equivalently understood either as a linear combination of the observations or as a linear transformation of the vector of observations, i.e., any linear transformation applied to the vector of observations $(\mathcal{X}_{\boldsymbol{s}_1}, \ldots, \mathcal{X}_{\boldsymbol{s}_n})^T \in \mathbb{R}^n$ and valued in $\mathbb{R}$ acts as a linear combination of $\mathcal{X}_{\boldsymbol{s}_1}, \ldots, \mathcal{X}_{\boldsymbol{s}_n}$. Instead, when $\mathcal{H}$ is an infinite-dimensional Hilbert space, an ambiguity exists in the definition of a Kriging predictor. For instance, Giraldo et al. [16] and Menafoglio et al. [23] interpret the Kriging problem in terms of finding the BLUP among the predictors of the form

$$\mathcal{X}^{\boldsymbol{\lambda}}_{\boldsymbol{s}_0} = \sum_{i=1}^n \lambda_i \mathcal{X}_{\boldsymbol{s}_i}, \tag{1}$$

with λ_i scalar weights in $\mathbb{R}$, for $i = 1, \ldots, n$.

Despite its simplicity, predictor (1) does not provide, in general, the best linear unbiased transformation of the vector of observations, that is the *Operatorial* Kriging predictor $\mathcal{X}^{\Lambda}_{\boldsymbol{s}_0} = \Lambda(\mathcal{X}_{\boldsymbol{s}_1}, \ldots, \mathcal{X}_{\boldsymbol{s}_n})$, for some linear operator $\Lambda : \mathcal{H} \times \cdots \times \mathcal{H} \to \mathcal{H}$. The operatorial viewpoint has been first considered by Nerini et al. [25] in Reproducing Kernel Hilbert Spaces (RKHSs). These authors address the problem of finding the best predictor over the class of linear unbiased Hilbert–Schmidt transformations of the observations, i.e., of the form

$$\mathcal{X}^{\boldsymbol{B}}_{\boldsymbol{s}_0} = \sum_{i=1}^n B_i \mathcal{X}_{\boldsymbol{s}_i}, \tag{2}$$

where $B_i : \mathcal{H} \to \mathcal{H}$ are Hilbert–Schmidt linear operators and $\mathcal{X}_{\boldsymbol{s}_i}$ observations in a RKHS. Even though this class of predictors is more general than that of (1), the RKHS-embedding – which is key to the well-posedness of the problem – still appears a too restrictive setting, as, for instance, the Hilbert space L^2 is not a RKHS, even though it is commonly employed in FDA.

In this work we establish an Operatorial Kriging theory valid for any separable Hilbert space, which relies upon the key notion of measurable linear transformation associated with a Gaussian measure [20,19] (Section 3). This broad class of operators includes linear Hilbert–Schmidt operators, and is here shown to allow for the Operatorial Kriging prediction in any finite- or infinite-dimensional separable Hilbert space.