

GEOSTATISTICS WITH INFINITE DIMENSIONAL DATA: A GENERALIZATION OF COKRIGING AND MULTIVARIABLE SPATIAL PREDICTION

¹Giraldo Ramón, ²Delicado Pedro, ³Mateu Jorge

Abstract. *We extend cokriging analysis and multivariable spatial prediction to the case where the observations at each sampling location consist of samples of random functions, that is, we extend two classical multivariable geostatistical methods to the functional context. Our cokriging method predicts one variable at a time as in a classical multivariable sense, but considering as auxiliary information curves instead of vectors. We also propose an extension of multivariable kriging to the functional context by defining a predictor of a whole curve based on samples of curves located at a neighborhood of the prediction site. In both cases a non-parametric approach based on basis function expansion is used to estimate the parameters, and we prove that both proposals coincide when using such an approach. A linear model of coregionalization is used to define the spatial dependence among the coefficients of the basis functions, and therefore for estimating the functional parameters. As an illustration the methodological proposals are applied to analyze two real data sets corresponding to average daily temperatures measured at 35 weather stations located in the Canadian Maritime Provinces, and penetration resistance data collected at 32 sampling sites of an experimental plot.*

Keywords: Basis functions; Cross-validation; Functional linear model; LMC; Multivariable cokriging.

Resumen. *Ampliamos cokriging análisis y predicción espacial multivariable para el caso de que las observaciones en cada punto de muestreo consistirá en muestras de funciones aleatorias, es decir, que se extienden dos métodos clásicos de geoestadística multivariable al contexto funcional. Nuestro método cokriging predice una variable a la vez como en un sentido clásico multivariable, pero teniendo en cuenta como información auxiliar curvas en lugar de vectores. Se propone también una extensión de varias variables kriging al contexto funcional mediante la definición de un factor de predicción de una curva de conjunto, basada en muestras de curvas situado en un barrio del sitio predicción. En ambos casos, un enfoque no paramétrico basado en la expansión de funciones de base se utiliza para estimar los parámetros, y se demuestra que ambas propuestas coinciden cuando se utiliza este enfoque. Un modelo lineal de coregionalización se utiliza para definir la dependencia espacial entre los coeficientes de las funciones de base, y por lo tanto para estimar los parámetros funcionales. A modo de ejemplo las propuestas metodológicas se aplican a analizar dos conjuntos de datos reales correspondientes a las temperaturas medias diarias medido en 35 estaciones meteorológicas ubicadas en las provincias canadienses Marítima, y los datos de resistencia a la penetración recogidos en 32 sitios de muestreo de una parcela experimental.*

Palabras Claves: Funciones de base; validación cruzada; modelo lineal funcional; cokriging multivariable; LMC.

Recibido: Noviembre, 2010

Aceptado: Febrero, 2011

1. INTRODUCTION

In this paper we focus on spatially correlated functional data, and particularly in modeling curves collected in sites of a region with spatial continuity. In spatial statistics, and in particular in geostatistics, both cokriging analysis [8, 1] and multivariable kriging [12, 11] are used for modeling observations of vector-valued random fields. Here we adapt these methodologies to the functional context. We extend the multivariable kriging from random vectors to the functional context by defining a functional kriging predictor which allows to predict a whole curve at an unvisited site by using as information the curves sampled in the neighborhood of the prediction site. In both cases (cokriging based on orthogonality is not a required condition.

The problem of spatial prediction of curves in a curves and functional kriging) we give a non-parametric solution based on basis functions. The problem of functional kriging prediction is also studied by [9]. However, their proposal is based only on orthonormal basis functions. In our case, we propose a more general technique in which geostatistical context has been considered from several points of view. [5] is a pioneer work in this topic. They propose three geostatistical approaches to predict curves: a curve kriging approach and two multivariable approaches based on cokriging on either discrete data or coefficients of the parametric models that have been fitted to the observed curves. [4] propose a non-parametric approach for solving the first approach considered by [5]. The predictor in the first proposal of [5] as well as that considered by [4] has the same form as the classical ordinary kriging predictor [2], but considering curves instead of one-dimensional data, that is, each curve is weighted by a scalar parameter. [3] solve the problem of spatial prediction of functional data by

¹ Giraldo Ramón, M.Sc., Universidad Nacional de Colombia. Bogotá-Colombia. (e_mail: rgiraldoh@unal.edu.co).

² Delicado Pedro, M.Sc., Universidad Politécnica de Cataluña. Barcelona-España. (e_mail: pedro.delicado@upc.edu).

³ Mateu Jorge, M.Sc., Universidad Jume I. Castellón-España. (e_mail: pedro.delicado@upc.edu).

weighting each observed curve by a functional parameter. This approach is a hybrid between ordinary kriging and the functional linear concurrent (*point-wise*) model such as shown in [10]. The methodologies proposed in this paper follow the line of [3] in the sense that each observed curve is weighted by a functional parameter. However, here the flexibility increases because double indexed functional parameters are considered and estimated. Now, each curve is weighted by a functional parameter for the prediction at each time. This modeling approach follows the basic philosophy of the functional linear model for functional response, for which a bivariate regression coefficient function must be estimated [7]. A good knowledge of the spatial and temporal patterns of meteorological and climate variables is often required when dealing with environmental problems. Here we illustrate how the predictors proposed can be used in modeling a data set consisting of daily temperature measurements recorded at 35 weather stations of the Canadian Maritime Provinces. In addition we apply the methodology to an agronomic data set corresponding to penetration resistance data recorded at 35 sites of an experimental plot at the National University of Colombia. The remainder of the paper is organized as follows. Section 2 presents an overview of cokriging and multivariable kriging. In Section 3 an extension of multivariable kriging to the functional context is shown. The paper ends with a brief discussion and suggestions for further research.

2. COKRIGING AND MULTIVARIABLE SPATIAL PREDICTION

In this section we show the basics of cokriging [8, 1] and multivariable spatial prediction [12]. As in [11] we use the term cokriging to mean prediction of a single random variable, and the term multivariable spatial prediction when predicting a vector of random variables. Let $\{Z(s) = (Z_1(s), \dots, Z_m(s)) : s \in D\}$ be a multivariable spatial process defined over a domain $D \subset \mathbb{R}^d$. We now consider the model $Z(s) = \mu(s) + \varepsilon(s)$, where $\mu(s)$ is a mean vector and $\varepsilon(s)$ a random vector with expected value $E(\varepsilon(s)) = 0$. It is assumed that the process is stationary, that is, the mean vector is considered constant for all $s \in D$, and the variance (covariance), cross-covariance and crossvariogram

functions depend only on the separation vector h , and not on locations.

We use the following notation:

- $2\gamma_{lq}(s_i, s_j) = V(Z_l(s_i) - Z_q(s_j))$, where V stands for the variance.
- $\gamma_{lk}^T = (\gamma_{lk}(s_1, s_0), \dots, \gamma_{lk}(s_n, s_0))$
- $\Gamma_{lq} = \begin{pmatrix} \gamma_{lq}(s_1, s_1) & \cdots & \gamma_{lq}(s_1, s_n) \\ \vdots & \ddots & \vdots \\ \gamma_{lq}(s_n, s_1) & \cdots & \gamma_{lq}(s_n, s_n) \end{pmatrix}$

The cokriging predictor of the k -th variable, $k = 1, \dots, m$, at the location s_0 is given by

$$\begin{aligned} \hat{Z}_k(s_0) &= \sum_{j=1}^m \lambda_{1j}^k Z_j(s_1) + \dots + \sum_{j=1}^m \lambda_{nj}^k Z_j(s_n) \\ &= \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^k Z_j(s_i) \end{aligned} \quad (1)$$

The predictor (1) is unbiased if $\sum_{i=1}^n \lambda_{ik}^k = 1$ and

$$\sum_{i=1}^n \lambda_{ij}^k = 0 \text{ for all } j \neq k, j = 1, \dots, m.$$

A variant of cokriging uses only the condition

$$\sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^k = 1 \text{ [6]. Using the method}$$

of Lagrange multipliers to minimize the mean squared prediction error subject to the unbiasedness constraints gives the cokriging system of equations, which in matrix notation can be expressed by

$$C\lambda^k = c^k,$$

Geostatistics with infinite dimensional data with

$$C = \begin{pmatrix} \Gamma_{11} & \cdots & \Gamma_{1k} & \cdots & \Gamma_{1m} & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots & & \vdots \\ \Gamma_{k1} & & \Gamma_{kk} & & \Gamma_{km} & 0 & & 1 & & 0 \\ \vdots & & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ \Gamma_{m1} & \cdots & \Gamma_{m2} & \cdots & \Gamma_{mm} & 0 & \cdots & 0 & \cdots & 1 \\ 1^T & \cdots & 0^T & \cdots & 0^T & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0^T & & 1^T & & 0^T & 0 & & 0 & & 0 \\ \vdots & & \vdots & \ddots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0^T & \cdots & 0^T & \cdots & 1^T & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} \Gamma & X \\ X^T & 0^* \end{pmatrix} \quad (2)$$

$$\lambda^k = \begin{pmatrix} \lambda_1^k \\ \vdots \\ \lambda_k^k \\ \vdots \\ \lambda_{km}^k \\ \delta_1 \\ \vdots \\ \delta_k \\ \vdots \\ \delta_m \end{pmatrix}, \quad c^k = \begin{pmatrix} \gamma_1^k \\ \vdots \\ \gamma_k^k \\ \vdots \\ \gamma_{km}^k \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

where,

$$\begin{aligned} & (\Gamma_{ij})_{(n \times n)}, \mathbf{1} = (1, \dots, 1)_{(n \times n)}^T, \\ & \mathbf{0} = (0, \dots, 0)_{(n \times 1)}^T, (\Gamma)_{(m \times n) \times (m \times n)}, (\mathbf{X})_{(n \times m) \times m}, (\mathbf{0}^*)_{(m \times m)}, \lambda_j^k \\ & = (\lambda_{1j}^k, \dots, \lambda_{mj}^k)^T \end{aligned}$$

, and $\gamma_j^k = (\gamma_{1j}^k, \dots, \gamma_{mj}^k)^T$, for all $i, j = 1, \dots, m$.

In multivariable spatial prediction [12, 11] all the m variables are predicted simultaneously at s_0 . In this case the predictor kriging is given by

$$\begin{pmatrix} \hat{Z}_1(s_0) \\ \vdots \\ \hat{Z}_m(s_0) \end{pmatrix} = \begin{pmatrix} \lambda_{11}^1 & \dots & \lambda_{1m}^1 & \dots & \lambda_{n1}^1 & \dots & \lambda_{nm}^1 \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ \lambda_{11}^m & \dots & \lambda_{1m}^m & \dots & \lambda_{n1}^m & \dots & \lambda_{nm}^m \end{pmatrix} \begin{pmatrix} Z_1(s_1) \\ \vdots \\ Z_m(s_1) \\ \vdots \\ Z_1(s_n) \\ \vdots \\ Z_m(s_n) \end{pmatrix} \quad (3)$$

and the matrix of parameters is obtained by solving the system [11]

$$\begin{pmatrix} \Gamma & \mathbf{X} \\ \mathbf{X}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Lambda \\ \Delta \end{pmatrix} = \begin{pmatrix} \mathbf{G} \\ \mathbf{I} \end{pmatrix},$$

where Γ and \mathbf{X} are defined as in (2), Λ is the matrix of parameters, Δ is a diagonal matrix of Lagrange multipliers, \mathbf{I} is an identity matrix and

$$G = \begin{pmatrix} \gamma_1^1 & \gamma_1^2 & \dots & \gamma_1^m \\ \gamma_2^1 & \gamma_2^2 & \dots & \gamma_2^m \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_m^1 & \gamma_m^2 & \dots & \gamma_m^m \end{pmatrix}.$$

Cokriging could be used for predicting simultaneously all m variables by cokriging each variable, one at a time. The cokriging prediction for one variable at a time coincides with the prediction of that variable obtained by multivariable spatial

prediction [12]. The difference between both approaches is given by their prediction variances. With cokriging we obtain a prediction variance for each univariate variable. In multivariable spatial prediction, in addition to the univariate prediction variances, it is possible to estimate a multidimensional prediction region with its long axis oriented toward regions where the predicted variables tend to co-vary [12].

3. FUNCTIONAL KRIGING: TOTAL MODEL

To define the functional kriging (*total model*) predictor (FKTM) we assume the same stationarity and isotropy assumptions. Thus, the predictor of the whole curve is given by

$$\hat{\chi}_{s_0}(v) = \sum_{i=1}^n \int_T \lambda_i(t, v) \chi_{s_i}(t) dt, v \in T, \quad (4)$$

such that $\lambda_1(t, v), \dots, \lambda_n(t, v): T \times T \rightarrow \mathbb{R}$. The

functional parameter $\lambda_i(t, v)$ in (4) determines the impact of the i -th observed function at time t on an unobserved function at time v . This modeling approach is coherent with the functional linear model for functional responses (*total model*) shown in [10]. In that framework and assuming that we have $X \in L_2(T_1), Y \in L_2(T_2)$, with L_2 a space of square integrable functions, the functional response $Y(v)$ is modeled in terms of the functional covariates by

$$Y(v) = \alpha(v) + \sum_{j=1}^q \int_{T_1} X_j(t) \beta_j(t, v) dt + \varepsilon(v), \quad (5)$$

where $\beta_j \in L_2(T_1 \times T_2)$ is a parameter function,

$\alpha \in L_2(T_2)$ is an intercept function and

$\varepsilon \in L_2(T_2)$ is a random error process such that

$E(\varepsilon(v)) = 0$ for all v . Estimation of functional

parameters in (5) is carried out by solving [10]

$$\text{Min}_{\alpha(\cdot), \beta_1(\cdot), \dots, \beta_q(\cdot)} E \|\hat{Y}(v) - Y(v)\|^2$$

In our context the covariates are the observed curves in n sites of a region and the functional response is an unobserved function at an unvisited location. Consequently our objective function is

$$E \|X_{s_0}(v) - \hat{X}_{s_0}(v)\|^2,$$

depending on $\lambda_1(\cdot, \cdot), \dots, \lambda_n(\cdot, \cdot)$, or by using

Fubini's Theorem

$$\int_T E \left(\hat{X}_{s_0}(v) - X_{s_0}(v) \right)^2 dv$$

Considering stationarity, the objective function becomes

$$\int_T V \left(\hat{X}_{s_0}(v) - X_{s_0}(v) \right) dv$$

Again the functional parameters $\lambda_i(t, v)$ in (4) are estimated taking into account the constraints of unbiasedness and minimum prediction variance.

Thus, the optimization problem becomes

$$\begin{aligned} & \underset{\lambda_1(\cdot), \dots, \lambda_n(\cdot)}{\text{Min}} \int_T V \left(\hat{X}_{s_0}(v) - X_{s_0}(v) \right) dv \\ & \text{s.t. } E \left(\hat{X}_{s_0}(v) \right) \\ & = E \left(X_{s_0}(v) \right), \forall v \in T \end{aligned}$$

We solve this problem by using an approach based on basis functions. We expand the functional variables using a basis function such as B-splines or Fourier and the bivariate functional parameters by

$$\lambda_i(t, v) = \sum_{j=1}^K \sum_{l=1}^K c_{jl}^i B_j(t) B_l(v) = B^T(t) C_i B(v), \quad (6)$$

where

$$C_i = \begin{pmatrix} c_{11}^i & c_{12}^i & \dots & c_{1K}^i \\ c_{21}^i & c_{22}^i & \dots & c_{2K}^i \\ \vdots & \vdots & \ddots & \vdots \\ c_{K1}^i & c_{K2}^i & \dots & c_{KK}^i \end{pmatrix}_{(K \times K)}$$

Consequently the predictor (4) can be expressed as

$$\begin{aligned} \hat{X}_{s_0}(v) &= \sum_{i=1}^n \int_T a_i^T B(t) B^T(t) C_i B(v) dt \\ &= \sum_{i=1}^n a_i^T W C_i B(v) = B^T(v) \sum_{i=1}^n C_i^T W^T a_i, \\ &= B^T(v) \sum_{i=1}^n C_i^T W a_i = B^T(v) \hat{a}_0 \end{aligned} \quad (7)$$

The predictor (4) is also considered by [9]. These authors assume that W is the identity matrix because they consider a solution based on orthonormal basis expansions. In our solution this is not a necessary condition. Now we consider the unbiasedness and minimum variance properties of the proposed predictor. The expected value of the curve on an unvisited site s_0 is given by

$$\begin{aligned} E \left(X_{s_0}(v) \right) &= E \left(\sum_{j=1}^K a_{0j} B_j(v) \right) \\ &= E \left(B^T(v) a_0 \right) = B^T(v) E \left(a_0 \right) \\ &= B^T(v) \vartheta \end{aligned} \quad (8)$$

On the other hand taking expected values in (7) we have

$$\begin{aligned} E \left(X_{s_0}(v) \right) &= B^T(v) \sum_{i=1}^n C_i^T W E \left(a_i \right) \\ &= B^T(v) \sum_{i=1}^n C_i^T W \vartheta \end{aligned} \quad (9)$$

Consequently from equations (8) and (9) we note that the predictor (4) is unbiased if and only if

$$B^T(v) \sum_{i=1}^n C_i^T W \vartheta = B^T(v) \vartheta, \text{ for all } \vartheta \in T,$$

that is, if and only if,

$$\sum_{i=1}^n C_i^T W \vartheta = \vartheta, \text{ for all } \vartheta \in T,$$

Given that W is full rank, this condition is equivalent to

$$\sum_{i=1}^n C_i = W^{-1}$$

The n functional parameters in the predictor (4) are given by the solution of the following optimization problem

$$\underset{C_1, \dots, C_n}{\text{Min}} \int_T V \left(B^T(v) \hat{a}_0 - B^T(v) a_0 \right) dv \text{ s.t. } \sum_{i=1}^n C_i = W^{-1} \quad (10)$$

The integral in the objective function (10) can be rewritten as

$$\begin{aligned}
 & \int_T B^T(v) V(\hat{a}_0 - a_0) B(v) dv \\
 &= \int_T Tr \left(B^T \left| (v) V(\hat{a}_0 - a_0) B(v) \right. \right) dv \\
 &= Tr \left(V(\hat{a}_0 - a_0) \int_T B^T(v) B^T(v) dv \right) \\
 &= Tr \left(V(\hat{a}_0 - a_0) W \right)
 \end{aligned} \tag{11}$$

The variance in (11) is

$$\begin{aligned}
 V(\hat{a}_0 - a_0) &= \sum_{i=1}^n C_i^T W V(a_i) W C_i + 2 \sum_{i < j}^n C_i^T W C(j) W C_j \\
 &+ V(a_0) - 2 \sum_{i=1}^n C_i^T W C(a_i, a_0)
 \end{aligned} \tag{12}$$

From (12) and defining the following matrices

$$Q_{ii} = W V(a_i) W, Q_{ij} = W C(a_i, a_j) W, N_i = W C(a_i, a_0)$$

the optimization problem (10) can be expressed as

$$\begin{aligned}
 \text{Min}_{C_1, \dots, C_n, m} & \sum_{i=1}^n Tr \left(C_i^T Q_{ii} C_i W \right) + 2 \sum_{i < j}^n Tr \left(C_i^T Q_{ij} C_j W \right) \\
 &+ Tr \left(V(a_0) W \right) - 2 \sum_{i=1}^n Tr \left(C_i^T N_i W \right) + 2m \left(\sum_{i=1}^n C_i - W^{-1} \right)
 \end{aligned} \tag{13}$$

Derivatives with respect to $C_i, i = 1, \dots, n$ and m in (13) are given, respectively, by

$$2 \sum_{j=1}^n Q_{ij} C_j W - 2 N_i W + 2m \text{ and } \sum_{i=1}^n C_i - W^{-1}$$

The solution of the problem given in (13) is achieved by setting these derivatives equal to zero. This solution can be represented in matrix notation as

$$\begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1n} & 1 \\ Q_{21} & Q_{22} & \dots & Q_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Q_{n1} & Q_{n2} & \dots & Q_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \\ m^* \end{pmatrix} = \begin{pmatrix} N_1 \\ N_2 \\ \vdots \\ N_n \\ W^{-1} \end{pmatrix}, \tag{14}$$

Where $m^* = m W^{-1}$. From equations (11) and (12), an estimation of the integrated prediction variance

$$\sigma_{int}^2 = \int_T \sigma_{s_0}^2(v) dv = \int_T V(\hat{x}_{s_0}(v) - x_{s_0}(v)) dv$$

is given by

$$\begin{aligned}
 \hat{\sigma}_{int}^2 &= \sum_{i=1}^n Tr \left(\hat{C}_i^T Q_{ii} \hat{C}_i W \right) + 2 \sum_{i < j}^n Tr \left(\hat{C}_i^T Q_{ij} \hat{C}_j W \right) \\
 &+ Tr \left(V(a_0) W \right) - 2 \sum_{i=1}^n Tr \left(\hat{C}_i N_i W \right)
 \end{aligned} \tag{15}$$

where the matrices $\hat{C}_1, \dots, \hat{C}_n$ are obtained by solving the system (14).

REFERENCES AND ELECTRONIC

- [1]. **P. BOGAERT.** (1996). “*Comparison of kriging techniques in a space-time context*”. *Mathematical Geology*, 28(1):73{86.
- [2]. **N. CRESSIE.** (1993). “*Statistic for spatial data*”. Jhon Wiley & Sons, New York.
- [3]. **R. GIRALDO, P. DELICADO, AND J. MATEU.** (2010). “*Continuous time-varying kriging for spatial prediction of functional data: An environmental application*”. *Journal of Agricultural, Biological, and Environmental Statistics*, 15(1):66{82.
- [4]. **R. GIRALDO, P. DELICADO, AND J. MATEU.** (2010). “*Ordinary kriging for function-valued spatial data*”. *Environmental and Ecological Statistics*, DOI: 10.1007/s10651-010-0143-y.
- [5]. **M. GOULARD AND M. VOLTZ.** (1993). “*Geostatistical interpolation of curves: A case study in soil science*”. In A. Soares, editor, *Geostatistics Tróia '92*, volume 2, pages 805{816. Kluwer Academic Press.
- [6]. **E. ISAAKS AND M. SRIVASTAVA.** (1987). “*Applied geostatistics*”. Oxford University Press, New York.
- [7]. **N. MALFAIT AND J. RAMSAY.** (2003). “*The historical functional linear model*”. *The Canadian Journal of Statistics*, 31(2):115{128.
- [8]. **D. MYERS.** (1982). “*Matrix formulation of Co-kriging*”. *Mathematical Geology*, 14(3):249{257.
- [9]. **D. NERINI AND P. MONESTIEZ.** (2008). “*A cokriging method for spatial functional data with applications in oceanology*”. Long summary sent to the first international workshop on functional and operational statistics. Toulouse, France. <http://www.lsp.ups-tlse.fr/staph/IWFOS2008/>.
- [10]. **J. RAMSAY AND B. SILVERMAN.** (2005). “*Functional data analysis*”. Second edition. Springer, New York.
- [11]. **J. VER HOEF AND R. BARRY.** (1998). “*Constructing and fitting models for cokriging and multivariable spatial prediction*. *Journal of Statistical Planning and Inference*, 69:275{294.
- [12]. **J. VER HOEF AND N. CRESSIE.** (1993). “*Multivariable spatial prediction*”. *Mathematical Geology*, 25(2):219{240.