Discussiones Mathematicae
Probability and Statistics 25 (2005) 139–159

KRIGING AND MASUREMENT ERRORS

István Fazekas

Faculty of Informatics, University of Debrecen
P.O. Box 12, H–4010 Debrecen, Hungary
e-mail: fazekasi@inf.unideb.hu

and

Alexander G. Kukush

Department of Mathematics and Mechanics, Kiev University
Vladimirskaya st. 64, 252601 Kiev, Ukraine
e-mail: kuog@mechmat.univ.kiev.ua

Abstract

A linear geostatistical model is considered. Properties of a universal kriging are studied when the locations of observations are measured with errors. Alternative prediction procedures are introduced and their least squares errors are analyzed.

Keywords: universal kriging, least squares, errors-in-variables.

2000 Mathematics Subject Classification: 62M30.

1. Introduction

A random field $z(t)$, $t \in D$, of the form (1) is observed at locations $\xi_1, \ldots, \xi_n$. Kriging means optimal predicting the unobserved values of the random field z. The procedure of kriging is well-known (see e.g. Krige (1951), Cressie (1991), Yakowitz–Szidarovszky (1985), and Berke (1998)). Our aim is to study the result of kriging when locations $\xi_1, \ldots, \xi_n$, where the field is observed, are not known precisely.

A similar problem was studied in Fazekas–Kukush (1999) but another mathematical setting was used there. In that paper, the location ξ_i was

considered as a non-random point, but we observed ξ_i with a random error. Therefore the appropriate tool was the deconvolution method often used for estimation in functional errors-in-variables models (see e.g. Fazekas–Baran–Kukush–Lauridsen (1999)).

In this paper, we consider that $\xi_1, \dots, \xi_n$ are random and we do not know their actual value but we have certain knowledge about their distribution. In Section 2, we introduce a measurement error model for random fields prediction problems. In Section 3, a modification of the universal kriging is given. Our model and kriging method is similar to those proposed in Gabrosek–Cressie (2002). In Gabrosek–Cressie (2002) the objective function is the first term of our objective function $Q(\lambda)$ in (8). In Theorem 3, we prove that under some conditions the L_2-error of our predictor is asymptotically 0. Another possibility for prediction in our case is to use the universal kriging procedure without any modification for data with error. This is called naive kriging and is studied in Section 4. In Section 5, we present a small set of simulation results. It shows that in the case of location errors neither naive kriging nor modified kriging can reproduce the precision of universal kriging with error free locations. The modified kriging is better than the naive kriging and moreover, it is more stable.

2. A measurement error model for random fields

2.1. The model and the assumptions

Consider the following linear geostatistical model

$$z(t) = \sum\nolimits_{j=0}^{k} \beta_j f_j(t) + \delta(t) + \varepsilon(t), \quad t \in D \subseteq \mathbb{R}^p, \tag{1}$$

where p is a fixed positive integer, $f_j(t)$ are known functions, $f_0(t) \equiv 1$, β_j are unknown parameters, $j = 0, 1, \dots, k$, $\delta(t)$ are random error terms with $\mathbb{E}\delta(t) = 0$, $\mathbb{E}\delta^2(t) < \infty$, $t \in D$. Random measurement error terms $\varepsilon(t)$, $t \in D$, are assumed to be independent of $\delta(t)$, $t \in D$. We shall always suppose that the functions $f_j(t)$, $j = 0, \dots, k$, are measurable and $\{\varepsilon(t),\ t \in D\}$, $\{\delta(t),\ t \in D\}$ are measurable random fields. The field $\{\varepsilon(t)\}$ is a white noise, $\mathbb{E}\varepsilon(t) = 0$, $\sigma^2 = \mathbb{E}\varepsilon^2(t) < \infty$, $t \in D$. We assume that $\varepsilon(t)$ is due to the measuring procedure and therefore we can assume that σ^2 is known.

We are interested in the field without an observation error, i.e., in

$$z_0(t) = \sum\nolimits_{j=0}^{k} \beta_j f_j(t) + \delta(t), \quad t \in D \subseteq \mathbb{R}^p, \tag{2}$$

but we cannot observe it directly. The observed random field is $z(t)$. Describe the scheme of observations:

$$z_i = z(\xi_i) = \sum\nolimits_{j=0}^{k} \beta_j f_j(\xi_i) + \delta(\xi_i) + \varepsilon(\xi_i), \quad i = 1, 2, \ldots, n,$$

where $\xi_1, \xi_2, \ldots, \xi_n \in D$ are random. Moreover, we do not know the precise value of ξ_i. We assume that the distribution of ξ_i is known, it is P_{ξ_i}. We use especially the expectation $\xi_i^0 = \mathbb{E}\xi_i$. We interpret it as follows. We intend to observe the field at location ξ_i^0. However, because of some measurement error the actual observation is made somewhere around ξ_i^0, namely at location ξ_i.

We assume that $\xi_1, \ldots, \xi_n$ are independent, and the three sets $\{\xi_i : i = 1, \ldots, n\}$, $\{\delta(t) : t \in D\}$, $\{\varepsilon(t) : t \in D\}$ of random variables are independent of one another.

The observations in the matrix form are the following:

$$Z(\xi) = F(\xi)\beta + \Delta(\xi) + \varepsilon(\xi),$$

where

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad F(\xi) = \begin{pmatrix} f_0(\xi_1) & \cdots & f_k(\xi_1) \\ \vdots & & \vdots \\ f_0(\xi_n) & \cdots & f_k(\xi_n) \end{pmatrix} = \begin{pmatrix} f_{\xi_1}^\top \\ \vdots \\ f_{\xi_n}^\top \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_{\widetilde{0}} \end{pmatrix}, \quad Z = Z(\xi) = \begin{pmatrix} z(\xi_1) \\ \vdots \\ z(\xi_n) \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix},$$

$$\Delta = \Delta(\xi) = \begin{pmatrix} \delta(\xi_1) \\ \vdots \\ \delta(\xi_n) \end{pmatrix}, \quad \varepsilon = \varepsilon(\xi) = \begin{pmatrix} \varepsilon(\xi_1) \\ \vdots \\ \varepsilon(\xi_n) \end{pmatrix}.$$

Let $s_0 \in D$ be fixed. Our aim is to predict $z_0(s_0)$ using $Z(\xi)$ and the distribution of ξ. The universal kriging gives the best linear unbiased predictor (see Krige (1951), Cressie (1991)). It approximates $z_0(s_0)$ with a linear function of $Z = Z(\xi)$:

$$z_0(s_0) \approx \sum\nolimits_{i=1}^{n} \lambda_i z_i = \lambda^\top Z,$$

where $\lambda = (\lambda_1, \dots, \lambda_n)^\top$ is a vector to be determined. The unbiasedness $\mathbb{E}z_0(s_0) = \mathbb{E}\sum_{i=1}^{n} \lambda_i z_i$ is equivalent to

$$f_j(s_0) = \sum\nolimits_{i=1}^{n} \lambda_i \mathbb{E} f_j(\xi_i), \quad j = 0, 1, \dots, k, \tag{3}$$

in other words $f(s_0) = \mathbb{E}(F(\xi))^\top \lambda$, where $f(s_0) = (f_0(s_0), f_1(s_0), \dots, f_k(s_0))^\top$. We remark that by the constraint (3) with $j = 0$, we have

$$\sum\nolimits_{i=1}^{n} \lambda_i = 1.$$

To obtain the best linear unbiased predictor we have to solve

$$\min_{\lambda} \mathbb{E}\left(z_0(s_0) - \lambda^\top Z\right)^2$$

under the constraint (3). Let

$$\gamma(h) = (1/2)\operatorname{var}\left(z_0(x+h) - z_0(x)\right),$$

where $x, x+h \in D$, be the semivariogram of z_0 (we assume that it does not depend on x, i.e., the field is intrinsically stationary). We assume that the semivariogram is known. We remark that the semivariogram of the observed random field $z(t) = z_0(t) + \varepsilon(t)$ is $\gamma(h) + \sigma^2$ if $h \neq 0$; and it is equal to $\gamma(0) = 0$ if $h = 0$ (as $\varepsilon(t)$ is a white noise). Let

$$\Gamma(\xi) = (\Gamma_{ij}(\xi))_{i,j=1}^{n}, \tag{4}$$

where

$$\Gamma_{ij}(\xi) = \gamma(s_0 - \xi_i) + \gamma(s_0 - \xi_j) - \gamma(\xi_i - \xi_j).$$

2.2. The mean squared error

Proposition 1. *Assume* (3). *Then the L_2-distance of $z_0(s_0)$ and $\sum_{i=1}^n \lambda_i z_i$ is*

$$\begin{aligned}(5)\qquad &\mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \lambda_i z_i\Big)^2 \\ &= \lambda^\top \mathbb{E}\Gamma(\xi)\lambda + \sigma^2 \sum\nolimits_{i=1}^n \lambda_i^2 + \beta_{\widetilde{0}}^\top \Big(\sum\nolimits_{i=1}^n \lambda_i^2 S_i\Big)\beta_{\widetilde{0}},\end{aligned}$$

where S_i is the covariance matrix of $(f_1(\xi_i), \dots, f_k(\xi_i))^\top$ while $\beta_{\widetilde{0}} = (\beta_1, \dots, \beta_k)^\top$.

Proof. The first part of the calculation is given in a more general setup. Let $\beta_0, \dots, \beta_k$ be random variables independent of $\{\xi_i\}$, $\{\delta(t)\}$, $\{\varepsilon(t)\}$ (a particular case is when $\beta_0, \dots, \beta_k$ are constants). From the constraint (3) we use only the case $j = 0$, that is $\sum_{i=1}^n \lambda_i = 1$.

To find the L_2-error, consider the following calculation.

$$\begin{aligned}&\mathbb{E}\Big\{\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \lambda_i z_i\Big)^2 \Big| \xi\Big\} \\ (6)\qquad &= \mathbb{E}\Big\{\Big(A_1 + A_2 + A_3\Big)^2 \Big| \xi\Big\} = \mathbb{E}\Big\{A_1^2 \Big| \xi\Big\} + \mathbb{E}\Big\{A_2^2 \Big| \xi\Big\} + \mathbb{E}\Big\{A_3^2 \Big| \xi\Big\} \\ &+ 2\mathbb{E}\Big\{A_1 A_2 \Big| \xi\Big\} + 2\mathbb{E}\Big\{A_1 A_3 \Big| \xi\Big\} + 2\mathbb{E}\Big\{A_2 A_3 \Big| \xi\Big\},\end{aligned}$$

where

$$A_1 = \sum\nolimits_{j=0}^k \beta_j \Big(f_j(s_0) - \sum\nolimits_{i=1}^n \lambda_i f_j(\xi_i)\Big),$$

$$A_2 = \delta(s_0) - \sum\nolimits_{i=1}^n \lambda_i \delta(\xi_i), \quad A_3 = -\sum\nolimits_{i=1}^n \lambda_i \varepsilon(\xi_i).$$

Now

$$\mathbb{E}\{A_1A_2|\xi\} = \sum\nolimits_{j=0}^{k}\Big[\Big(f_j(s_0) - \sum\nolimits_{i=1}^{n}\lambda_i f_j(\xi_i)\Big)\mathbb{E}\{\beta_j A_2|\xi\}\Big] = 0,$$

because, by independence,

$$\mathbb{E}\{\beta_j A_2|\xi = x\} = \mathbb{E}\Big\{\beta_j\Big(\delta(s_0) - \sum\nolimits_{i=1}^{n}\lambda_i\delta(x_i)\Big)\Big\} = 0$$

(here $x = (x_1, \dots, x_n)^\top$). Similarly

$$\mathbb{E}\{A_1A_3|\xi\} = 0, \quad \mathbb{E}\{A_2A_3|\xi\} = 0.$$

By independence

$$\mathbb{E}\{A_3^2|\xi = x\} = \mathbb{E}\Big[\sum\nolimits_{i=1}^{n}\lambda_i\varepsilon(x_i)\Big]^2 = \sigma^2\sum\nolimits_{i=1}^{n}\lambda_i^2,$$

$$\mathbb{E}\{A_3^2|\xi\} = \sigma^2\sum\nolimits_{i=1}^{n}\lambda_i^2.$$

As $\sum_{i=1}^{n}\lambda_i = 1$,

$$\mathbb{E}\{A_2^2|\xi = x\} = \mathbb{E}\Big\{\sum\nolimits_{i=1}^{n}\lambda_i(\delta(s_0) - \delta(x_i))\Big\}^2 = \lambda^\top\Gamma(x)\lambda,$$

$$\mathbb{E}\{A_2^2|\xi\} = \lambda^\top\Gamma(\xi)\lambda.$$

Using that $\sum_{i=1}^{n}\lambda_i = 1$ and $f_0 \equiv 1$,

$$\mathbb{E}\{A_1^2|\xi\} = \mathbb{E}\Big(\sum\nolimits_{i=1}^{n}\lambda_i\sum\nolimits_{j=1}^{k}\beta_j[f_j(s_0) - f_j(\xi_i)]\Big)^2.$$

Now substituting these into (6), we obtain

$$\mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^{n} \lambda_i z_i\Big)^2 = \lambda^\top \mathbb{E}\Gamma(\xi)\lambda + \sigma^2 \sum\nolimits_{i=1}^{n} \lambda_i^2$$

(7)

$$+\mathbb{E}\Big(\sum\nolimits_{i=1}^{n} \lambda_i \sum\nolimits_{j=1}^{k} \beta_j \left[f_j(s_0) - f_j(\xi_i)\right]\Big)^2.$$

From now on, we consider non-random $\beta_0, \ldots, \beta_k$ and use all equations in the constraint (3).

$$\mathbb{E}A_1^2 = \mathbb{D}^2 A_1 = \mathbb{D}^2\Big(\sum\nolimits_{i=1}^{n} \lambda_i \sum\nolimits_{j=1}^{k} \beta_j [f_j(s_0) - f_j(\xi_i)]\Big)$$

$$= \sum\nolimits_{i=1}^{n} \lambda_i^2 \mathbb{D}^2\Big(\sum\nolimits_{j=1}^{k} \beta_j [f_j(s_0) - f_j(\xi_i)]\Big) = \sum\nolimits_{i=1}^{n} \lambda_i^2 \mathbb{D}^2\Big(\sum\nolimits_{j=1}^{k} \beta_j f_j(\xi_i)\Big)$$

$$= \sum\nolimits_{i=1}^{n} \lambda_i^2 \beta_{\widetilde{0}}^\top S_i \beta_{\widetilde{0}} = \beta_{\widetilde{0}}^\top \Big(\sum\nolimits_{i=1}^{n} \lambda_i^2 S_i\Big)\beta_{\widetilde{0}},$$

where S_i is the covariance matrix of $(f_1(\xi_i), \ldots, f_k(\xi_i))^\top$. Now, substituting this into (7), we obtain (5). ∎

3. A modification of kriging

3.1. The new predictor

Our aim is to minimize the L_2-distance of $z_0(s_0)$ and $\sum_{i=1}^{n} \lambda_i z_i$ according to λ. However, the expression in (5) depends on the unknown value of β.

Assume that $\|\beta_{\widetilde{0}}\| \le R$, where R is a fixed positive constant, and $\|\beta_{\widetilde{0}}\|$ is the Euclidean norm of the vector $\beta_{\widetilde{0}}$. Therefore we have to minimize

$$\max_{\|\beta_{\widetilde{0}}\| \le R} \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^{n} \lambda_i z_i\Big)^2.$$

Therefore the objective function is

$$Q(\lambda) = \lambda^\top \mathbb{E}\Gamma(\xi)\lambda + \sigma^2 \sum\nolimits_{i=1}^n \lambda_i^2 + \Big\| \sum\nolimits_{i=1}^n \lambda_i^2 S_i \Big\| R^2, \tag{8}$$

where $\|A\|$ denotes the spectral norm of the matrix A, i.e., the maximal eigenvalue of A when A is symmetric and positive semidefinite. We have to minimize (8) under the constraint (3). The solution will be denoted by $\widehat{\lambda} = (\widehat{\lambda}_1, \dots, \widehat{\lambda}_n)^\top$.

We call this method modified kriging. The predictor

$$\widehat{z}(s_0) = \sum\nolimits_{i=1}^n \widehat{\lambda}_i z_i$$

will be called the modified kriging predictor. For this procedure we need the knowledge of the semivariogram γ, the functions f_i (γ and f_i are also used in the universal kriging), the distributions of $\varepsilon(t)$ and ξ_i, moreover the upper bound R. We consider the distributions of $\varepsilon(t)$ and ξ_i to be the characteristics of the measuring procedure therefore they are known previously. However, we have to find R.

In Gabrosek–Cressie (2002) for a similar measurement error setting the objective function was $\lambda^\top \mathbb{E}\Gamma(\xi)\lambda$ with the same constraint as (3).

Proposition 2. *By Lagrange multipliers one can get that the minimum of*

$$\lambda^\top \Gamma \lambda$$

under constraint $f = F^\top \lambda$ is attained at

$$\widehat{\lambda} = \Gamma^{-1} F (F^\top \Gamma^{-1} F)^{-1} f. \tag{9}$$

Example 1. Kriging a field with two regression parameters. In this case the model is

$$z_0(t) = \beta_0 + \beta_1 f_1(t) + \delta(t).$$

Now $S_i = \mathrm{var}(f_1(\xi_i))$, $i = 1, \ldots, n$, are one-dimensional, therefore the objective function is

$$Q(\lambda) = \lambda^\top \Big(\mathbb{E}\Gamma(\xi) + \sigma^2 I_n + R^2 \mathrm{diag}(S_1, \ldots, S_n) \Big) \lambda.$$

We have to minimize this function under the constraints

$$\sum\nolimits_{i=1}^{n} \lambda_i = 1, \qquad \sum\nolimits_{i=1}^{n} \lambda_i \mathbb{E} f_1(\xi_i) = f_1(s_0).$$

The solution is obtained by inserting $\Gamma = \mathbb{E}\Gamma(\xi) + \sigma^2 I_n + R^2 \,\mathrm{diag}(S_1, \ldots, S_n)$,

$$F = \begin{pmatrix} 1 & \mathbb{E} f_1(\xi_1) \\ \vdots & \vdots \\ 1 & \mathbb{E} f_1(\xi_n) \end{pmatrix}$$

and $f = (1, f_1(s_0))^\top$ into (9).

3.2. Asymptotic properties of modified kriging

Our aim is to prove that the error of modified kriging tends to zero if some regularity conditions are satisfied. We consider a fixed random field $z_0(t)$ of the form (2) which has to be predicted and another fixed random field $z(t)$ of the form (1) which is observed at some locations. The locations of observations are not fixed. During the m-th step we have observations at $n = n(m)$ points ($n(m) \to \infty$, as $m \to \infty$):

$$\xi_1 = \xi_1^{(m)}, \, \xi_2 = \xi_2^{(m)}, \ldots, \, \xi_n = \xi_{n(m)}^{(m)}.$$

(To avoid difficult notation we omit m, but we emphasize that the sites of observations depend on m.) The sites are random and their centres are $\xi_i^0 = \mathbb{E}\xi_i$, $i = 1, 2, \ldots, n$.

We have on our mind the so called infill asymptotics, that is the sites ξ_i^0, $i = 1, 2, \ldots, n$, become dense in the fixed domain D, as $m \to \infty$. Moreover, we think that the type of distribution of ξ_i is fixed but the variance tends to zero, as $m \to \infty$. We also consider that the functions $f_i(t)$ and the semivariogram γ satisfy some analytical conditions ($\gamma(h)$ is continuous at $h = 0$, say).

However, to obtain flexible conditions we shall express them in terms of expectations of certain functions of the random error terms. Then one can check them for particular semivariograms, mean value functions f_i (e.g. for polynomials), and error terms (e.g. for normally distributed ones). We shall give conditions for kriging at a fixed point s_0.

Assume that there is a sequence $r = r(m)$ of positive integers tending to infinity such that among the locations of observations ξ_i there are r points

$$\xi_{1*} = \xi_{1*}^{(m)},\ \xi_{2*} = \xi_{2*}^{(m)}, \dots,\ \xi_{r*} = \xi_{r(m)*}^{(m)}$$

with the following properties:

$$\alpha_j \le \frac{1}{r}, \quad j = 1, \dots, r, \tag{10}$$

$$\lim_{m\to\infty} \sum_{j=1}^{r} \alpha_j = 0, \tag{11}$$

where

$$\alpha_j = \max_{0\le i\le k} \left\| f_i(s_0) - \mathbb{E} f_i(\xi_{j*}) \right\|, \quad j = 1, \dots, r. \tag{12}$$

Conditions (10) and (11) mean that a subsequence of the locations ξ_i converges to s_0 in some sense.

Assume that among the locations ξ_i there are $k+1$ sites $t_0, t_1, \dots, t_k$, such that the matrix

$$\breve{\Phi} = \Big(\mathbb{E} f_i(t_j) \Big)_{i,j=0}^{k} \quad \text{is invertible,} \tag{13}$$

$$\|\breve{\Phi}^{-1}\| \le K, \tag{14}$$

where K is a finite constant not depending on m. This condition generally means that the locations $t_0, t_1, \dots, t_k$ do not converge to s_0. Let

$$\breve{\xi} = (\xi_{1*}, \dots, \xi_{r*}, t_0, t_1, \dots, t_k)^\top$$

be the vector of our specific locations. Let $\breve{S}_i$ be the covariance matrix S_i (defined in Proposition 1) calculated at the i-th coordinate of $\breve{\xi}$.

Assume that

$$\|\breve{S}_i\| \le L, \quad i = 1, \dots, r+k+1, \tag{15}$$

where L is a finite constant not depending on m.

Let $\Gamma(\breve{\xi})$ be the matrix Γ in (4) calculated at $\breve{\xi}$ and assume that

$$\left| \left(\mathbb{E}\Gamma(\breve{\xi}) \right)_{i,j} \right| \le M, \quad i,j = 1, \dots, r+k+1, \tag{16}$$

where M is a finite constant not depending on m. Moreover,

$$M_1 = \frac{1}{r^2} \sum\nolimits_{i=1}^{r} \sum\nolimits_{j=1}^{r} \left| \mathbb{E}\Gamma(\breve{\xi})_{i,j} \right| \to 0, \quad \text{as} \ \ m \to \infty. \tag{17}$$

Theorem 3. *Let R be an arbitrary upper bound of $\|\beta_{\widetilde{0}}\|$. Let $\widehat{\lambda}$ minimize* (8) *under the constraint* (3). *Assume that* (10)–(17) *are satisfied. Then for the modified kriging predictor $\widehat{z}(s_0)$*

$$\lim_{m \to \infty} \mathbb{E}\left(z_0(s_0) - \widehat{z}(s_0) \right)^2 = 0. \tag{18}$$

Proof. We apply some ideas of the proof of Theorem 2.1 in Yakowitz–Szidarovszky (1985). Consider the linear predictor

$$\breve{z} = \sum\nolimits_{j=1}^{r} \left[\frac{1}{r} - \alpha_j \right] z(\xi_{j*}) + \left[\sum\nolimits_{j=1}^{r} \alpha_j \right] \sum\nolimits_{i=0}^{k} \nu_i z(t_i), \tag{19}$$

where the coefficients $\nu_0, \dots, \nu_k$ are unknown. We need unbiased predictors, therefore (3) is to be satisfied:

$$\left[\sum\nolimits_{j=1}^{r} \alpha_j \right] \breve{\Phi}\nu = f - \Phi^* b, \tag{20}$$

where $f = f(s_0) = (f_0(s_0), f_1(s_0), \dots, f_k(s_0))^\top$, $b = (\frac{1}{r} - \alpha_1, \dots, \frac{1}{r} - \alpha_r)^\top$, $\nu = (\nu_0, \dots, \nu_k)^\top$ and

$$\Phi^* = \left(\mathbb{E} f_i(\xi_{j*}) \right)_{i=0,\, j=1}^{k \quad r}.$$

Assume now that $\sum_{j=1}^{r} \alpha_j \neq 0$. As $\breve{\Phi}$ is invertible, therefore (20) has a solution ν, so there is an unbiased linear predictor of the form (19). We shall use that choice of ν.

We have to prove that ν is bounded. Therefore consider

$$f - \Phi^* b = \Big[\sum\nolimits_{j=1}^{r} \alpha_j\Big] f + Ab,$$

where the (i,j)-th element of the matrix A is

$$a_{i,j} = f_i(s_0) - \mathbb{E} f_i(\xi_{j*}), \quad i = 0, \dots, k,\ j = 1, \dots, r.$$

Now we use the max-norm for vectors and the corresponding row-norm for matrices (both denoted by $\| \, . \, \|_\infty$).

$$\|Ab\|_\infty \leq \|A\|_\infty \|b\|_\infty = \max_{0 \leq i \leq k} \Big\{\sum\nolimits_{j=1}^{r} |a_{i,j}|\Big\} \|b\|_\infty$$

$$\leq \Big[\sum\nolimits_{j=1}^{r} \alpha_j\Big] \max_{0 \leq i \leq k} \left\{\left|\frac{1}{r} - \alpha_i\right|\right\} \leq \sum\nolimits_{j=1}^{r} \alpha_j,$$

because of (12) and (10). Therefore (20) and (14) imply

$$\|\nu\|_\infty \leq \|\breve{\Phi}^{-1}\|_\infty \Big(\|f\|_\infty + 1\Big) \leq K\Big(\|f\|_\infty + 1\Big).$$

If $\|\beta_{\widetilde{0}}\| \leq R$, then for the mean-square error of the modified kriging predictor $\widehat{z}(s_0)$ we have

$$\mathbb{E}\Big(z_0(s_0) - \widehat{z}(s_0)\Big)^2 \leq \widehat{\lambda}^\top \mathbb{E}\Gamma(\xi)\widehat{\lambda} + \sigma^2 \|\widehat{\lambda}\|^2 + \Big\| \sum\nolimits_{l=1}^{n} \widehat{\lambda}_l^2 S_l \Big\| R^2$$

$$\leq \breve{\lambda}^\top \mathbb{E}\Gamma(\breve{\xi})\breve{\lambda} + \sigma^2 \|\breve{\lambda}\|^2 + \Big\| \sum\nolimits_{l=1}^{r+k+1} \breve{\lambda}_l^2 \breve{S}_l \Big\| R^2. \tag{21}$$

Here

$$\breve{\lambda} = \left(\frac{1}{r} - \alpha_1, \ldots, \frac{1}{r} - \alpha_r, \nu_0 \sum_{j=1}^{r} \alpha_j, \ldots, \nu_k \sum_{j=1}^{r} \alpha_j\right)^\top$$

is the coefficient vector of the unbiased predictor $\breve{z}$.

Now it is enough to prove that (21) converges to 0, as $r = r(m) \to \infty$. We have

$$\|\breve{\lambda}\|^2 \leq \frac{1}{r} + (k+1)\|\nu\|_\infty^2 \left[\sum_{j=1}^{r} \alpha_j\right]^2.$$

Therefore $\sigma^2\|\breve{\lambda}\|^2 \to 0$. Because of (15), $\left\| \sum_{l=1}^{r+k+1} \breve{\lambda}_l^2 \breve{S}_l \right\| R^2 \to 0$.

Breaking $\mathbb{E}\Gamma(\breve{\xi})$ into appropriate blocks, using conditions (16), (17) and the form of $\breve{\lambda}$, we obtain that $\breve{\lambda}^\top \mathbb{E}\Gamma(\breve{\xi})\breve{\lambda} \to 0$. Therefore each summand of (21) converges to 0.

During the proof we assumed that $\sum_{j=1}^{r} \alpha_j \neq 0$. Now suppose that $\sum_{j=1}^{r} \alpha_j = 0$. Then $\breve{z} = \sum_{j=1}^{r} \frac{1}{r} z(\xi_{j*})$ is an unbiased predictor. For this predictor the value of the objective function Q converges to 0, because $r = r(m) \to \infty$, as $m \to \infty$. ∎

Corollary 4. *From the above calculation we can get an explicit formula for the L_2-error. Breaking $\mathbb{E}\Gamma(\breve{\xi})$ into four blocks (of sizes $r \times r$, $r \times (k+1)$, $(k+1) \times r$ and $(k+1) \times (k+1)$) gives*

$$\mathbb{E}\Big(z_0(s_0) - \widehat{z}(s_0)\Big)^2 \leq M_1 + M(k+1)^2\|\nu\|_\infty^2 \left[\sum_{j=1}^{r} \alpha_j\right]^2$$

$$+2M(k+1)\|\nu\|_\infty \left[\sum_{j=1}^{r} \alpha_j\right] + \sigma^2\|\breve{\lambda}\|^2 + \|\breve{\lambda}\|^2 LR^2.$$

3.3. Fields with Gaussian semivariogram

Example 2. We describe the modified kriging when $z_0(t)$ is a random field with the Gaussian semivariogram $\gamma(h)$, the polinomial trend function (i.e., $f_i(t)$ is a power function for each i), and the location ξ_i is normally distributed for each i. We calculate the ingredients for the objective function Q in (8) and for the constraint (3). Moreover, we check if the conditions of L_2-consistency in Theorem 3.4 can be fulfilled.

For simplicity, we give the details for a random field over the plain, that is $t \in D \subseteq \mathbb{R}^2$.

Let ξ_i be normally distributed around the centre ξ_i^0, namely

$$\xi_i = \begin{pmatrix} \xi_{i,1} \\ \xi_{i,2} \end{pmatrix} \sim \mathcal{N}_2 \left(\begin{pmatrix} \xi_{i,1}^0 \\ \xi_{i,2}^0 \end{pmatrix}, d^2 I \right),$$

where I is the identity matrix of size 2×2. Therefore the $2n$ random variables $\xi_{i,1}, \xi_{i,2}$, $i = 1, \ldots, n$, are independent normal.

A point of the plain we shall denote by $t = (x, y)^\top$. We predict the field at the location $s_0 = (s_{0,1}, s_{0,2})^\top$. We consider the polinomial trend function

$$\beta_0 + \beta_1 x + \beta_2 y + \beta_3 x^2 + \beta_4 y^2.$$

The moments of the normal distribution are easily calculated. However, for convenience, we give the following list which is necessary to describe the constraints $f_j(s_0) = \sum_{i=1}^n \lambda_i \mathbb{E} f_j(\xi_i)$, $j = 0, 1, \ldots, k$, in (3).

$$f_0(t) = 1, \quad \mathbb{E} f_0(\xi_i) = 1, \qquad f_0(s_0) = 1;$$

$$f_1(t) = x, \quad \mathbb{E} f_1(\xi_i) = \xi_{i,1}^0, \qquad f_1(s_0) = s_{0,1};$$

$$f_2(t) = y, \quad \mathbb{E} f_2(\xi_i) = \xi_{i,2}^0, \qquad f_2(s_0) = s_{0,2};$$

$$f_3(t) = x^2, \quad \mathbb{E} f_3(\xi_i) = d^2 + (\xi_{i,1}^0)^2, \quad f_3(s_0) = s_{0,1}^2;$$

$$f_4(t) = y^2, \quad \mathbb{E} f_4(\xi_i) = d^2 + (\xi_{i,2}^0)^2, \quad f_4(s_0) = s_{0,2}^2.$$

From here we also see that the quantities $\alpha_j = \max_{0 \le i \le k} \| f_i(s_0) - \mathbb{E} f_i(\xi_{j*}) \|$, $j = 1, \ldots, r$, can satisfy conditions (10) and (11) if and only if there is a subsequence of the locations ξ_i tending to s_0 and the dispersion d converges to 0, as $m \to \infty$.

It is seen that one can choose the locations ξ_i, $i = 1, \ldots, n$, such that among them there are $k+1$ sites $t_0, t_1, \ldots, t_k$, such that the matrix $\breve{\Phi} = \Big(\mathbb{E} f_i(t_j) \Big)_{i,j=0}^{k}$ is invertible and its inverse is bounded, i.e., condition (14) is satisfied.

Now we calculate S_i, i.e., the covariance matrix of $(f_1(\xi_i), \dots, f_k(\xi_i))^\top$. Using the moments of the normal distribution, we get for $i = 1, \dots, n$, that

$$S_i = \begin{bmatrix} d^2 & 0 & 2d^2\xi^0_{i,1} & 0 \\ 0 & d^2 & 0 & 2d^2\xi^0_{i,2} \\ 2d^2\xi^0_{i,1} & 0 & 2d^4 + 4d^2(\xi^0_{i,1})^2 & 0 \\ 0 & 2d^2\xi^0_{i,2} & 0 & 2d^4 + 4d^2(\xi^0_{i,2})^2 \end{bmatrix}.$$

Now, it is easy to see that if the points $\xi_{1*}, \dots, \xi_{r*}, t_0, t_1, \dots, t_k$ are chosen from a bounded region, then S matrices calculated at these points are bounded, i.e., condition (15) is satisfied.

Now, we turn to the matrix Γ. Let γ be the Gaussian semivariogram

$$\gamma(h) = \omega\Big(1 - \exp\big(-\|h\|^2/a^2\big)\Big)$$

for $h \in \mathbb{R}^2$, where $\|h\|^2 = h_1^2 + h_2^2$. The (i,j)-th element of $\Gamma(\xi)$ is

$$\Gamma_{ij}(\xi) = \gamma(s_0 - \xi_i) + \gamma(s_0 - \xi_j) - \gamma(\xi_i - \xi_j) = \tag{22}$$

$$\omega\Big\{ \exp\big[-\|\xi_i - \xi_j\|^2/a^2\big] - \exp\big[-\|s_0 - \xi_i\|^2/a^2\big] - \exp\big[-\|s_0 - \xi_j\|^2/a^2\big] + 1 \Big\},$$

$i, j = 1, \dots, n$. Here

$$s_0 - \xi_i \sim \mathcal{N}_2\left(s_0 - \xi_i^0,\, d^2 I\right), \quad \xi_i - \xi_j \sim \mathcal{N}_2\left(\xi_i^0 - \xi_j^0,\, 2d^2 I\right),$$

if $i \neq j$; while $\xi_i - \xi_j = 0$, if $i = j$. The squared norm of these variables can be described by χ^2 distribution:

$$\text{if} \quad X \sim \mathcal{N}_2\left(m, I\right), \quad \text{then} \quad \|X\|^2 \sim \chi^2_2(\|m\|^2),$$

where $\chi^2_2(\|m\|^2)$ denotes the chi-square distribution with degree of freedom 2 and with noncentrality parameter $\|m\|^2$. We know the moment generating function of the chi-square distribution $\chi^2_l(p)$:

$$\mathbb{E}\exp(v\chi^2_l(p) = (1-2v)^{-l/2}\exp\left(\frac{pv}{1-2v}\right). \tag{23}$$

Actually, we have to calculate the expectation of $\Gamma_{i,j}(\xi)$. Applying (23), we can obtain the expectation of each summand in (22). We have

$$\mathbb{E}\exp\Big[-\|s_0-\xi_j\|^2/a^2\Big] = \mathbb{E}\exp\left[-\frac{d^2}{a^2}\ \frac{\|s_0-\xi_j\|^2}{d^2}\right]$$

$$= \frac{a^2}{2d^2+a^2}\exp\left[-\frac{(s_{0,1}-\xi^0_{j,1})^2}{2d^2+a^2}\right]\exp\left[-\frac{(s_{0,2}-\xi^0_{j,2})^2}{2d^2+a^2}\right];$$

$$\mathbb{E}\exp\Big[-\|\xi_i-\xi_j\|^2/a^2\Big] = \mathbb{E}\exp\left[-\frac{2d^2}{a^2}\ \frac{\|\xi_i-\xi_j\|^2}{2d^2}\right]$$

$$= \frac{a^2}{4d^2+a^2}\exp\left[-\frac{(\xi^0_{i,1}-\xi^0_{j,1})^2}{4d^2+a^2}\right]\exp\left[-\frac{(\xi^0_{i,2}-\xi^0_{j,2})^2}{4d^2+a^2}\right],$$

if $i \neq j$, and it is equal to 1, if $i = j$. Therefore we can calculate all ingredients of the objective function Q in (8).

Finally, we can see that condition (16) for $\mathbb{E}\Gamma(\breve{\xi})$ is automatically satisfied. Moreover, condition (17) is also satisfied if there is a subsequence of the locations ξ_i tending to s_0 and the dispersion d converges to 0, as $m \to \infty$.

Therefore in this example we obtained realistic conditions which imply (18).

4. NAIVE KRIGING

Naive kriging means that we apply the universal kriging procedure as if ξ_i^0, $i = 1, 2, \ldots, n$, were the exact locations of the observations. Therefore we ignore information about the distribution of ξ_i.

As universal kriging produces the best linear unbaised predictor, we approximate $z(s_0)$ with a linear funcion:

$$z_0(s_0) \approx \sum\nolimits_{i=1}^{n} \lambda_i z_i.$$

With exact sites the unbiasedness is

$$f_j(s_0) = \sum\nolimits_{i=1}^{n} \lambda_i f_j(\xi_i^0), \quad j = 0, 1, \ldots, k. \tag{24}$$

To obtain the predictor we have to solve the minimum problem

$$\min_{\lambda} \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^{n} \lambda_i z_i\Big)^2$$

under constraints (24). With exact sites the value of this objective function is

$$Q_{\mathrm{N}}(\lambda) = \lambda^\top \left(\Gamma(\xi^0) + \sigma^2 I\right) \lambda, \tag{25}$$

where $\xi^0 = (\xi_1^0, \ldots, \xi_n^0)^\top$. We have to minimize (25) under constraints (24).

Denote the solution of this minimization problem by $\widetilde{\lambda}$. Then the naive kriging gives the prediction $z_0(s_0) \approx \sum_{i=1}^{n} \widetilde{\lambda}_i z_i$.

One can easily calculate the L_2-error of this predictor:

$$H(\widetilde{\lambda}) = \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^{n} \widetilde{\lambda}_i z_i\Big)^2$$

$$= \widetilde{\lambda}^\top \mathbb{E}\Gamma(\xi)\widetilde{\lambda} + \sigma^2 \sum\nolimits_{i=1}^{n} \widetilde{\lambda}_i^2 + \beta_0^\top \Big(\sum\nolimits_{i=1}^{n} \widetilde{\lambda}_i^2 S_i\Big)\beta_0 + \widetilde{M}(\beta),$$

where $\widetilde{M}(\beta) = \Big(\sum_{i=1}^{n} \widetilde{\lambda}_i \sum_{j=1}^{k} \beta_j \left[f_j(\xi_i^0) - \mathbb{E} f_j(\xi_i)\right]\Big)^2$.

Example 3. Let $f_j(t)$, $j = 0, 1, \dots, k$, be constant or linear functions. Then $\mathbb{E}f_j(\xi_i) = f_j(\xi_i^0)$ for every i and j. Therefore constraints for the modified kriging and those for the naive kriging (i.e., (3) and (24)) coincide. Moreover, the form of the L_2-errors are also the same, because in this special case

$$H(\widetilde{\lambda}) = \widetilde{\lambda}^\top \mathbb{E}\Gamma(\xi)\widetilde{\lambda} + \sigma^2 \sum\nolimits_{i=1}^n \widetilde{\lambda}_i^2 + \beta_{\widetilde{0}}^\top \Big(\sum\nolimits_{i=1}^n \widetilde{\lambda}_i^2 S_i \Big) \beta_{\widetilde{0}}. \tag{26}$$

(We recall that, by (5), the L_2-error of the modified kriging is $H(\widehat{\lambda})$, where function H is given in (26), but $\widehat{\lambda}$ is the solution of the minimum problem (8) under the constraint (3).)

Therefore for the L_2-error of the modified kriging and that of the naive kriging we obtain:

$$\sup_{\|\beta_{\widetilde{0}}\| \le R} \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \widehat{\lambda}_i z_i \Big)^2 \le \sup_{\|\beta_{\widetilde{0}}\| \le R} \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \widetilde{\lambda}_i z_i \Big)^2. \tag{27}$$

Example 4. Let $k = 0$, and $f_0(t) \equiv 1$. Now we are in the situation of ordinary kriging

$$z_0(t) = \beta_0 + \delta(t).$$

Inequality (27) for the L_2-error of the modified kriging and that of the naive kriging now is as follows:

$$\mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \widehat{\lambda}_i z_i \Big)^2 \le \mathbb{E}\Big(z_0(s_0) - \sum\nolimits_{i=1}^n \widetilde{\lambda}_i z_i \Big)^2. \tag{28}$$

Here the left hand side is the value of the objective function $Q(\lambda) = \lambda^\top \left(\mathbb{E}\Gamma(\xi) + \sigma^2 I \right) \lambda$ at $\widehat{\lambda}$, while the right hand side is its value at $\widetilde{\lambda}$. Moreover, $\widehat{\lambda}$ is the minimum of $Q(\lambda)$ under constraint $e^\top \lambda = 1$. On the other hand, $\widetilde{\lambda}$ is the minimum of $Q_{\mathrm{N}}(\lambda) = \lambda^\top \left(\Gamma(\xi^0) + \sigma^2 I \right) \lambda$ under constraint $e^\top \lambda = 1$. Therefore, by (9),

$$\widehat{\lambda} = \frac{\left(\mathbb{E}\Gamma(\xi) + \sigma^2 I\right)^{-1} e}{e^\top \left(\mathbb{E}\Gamma(\xi) + \sigma^2 I\right)^{-1} e}, \quad \widetilde{\lambda} = \frac{\left(\Gamma(\xi^0) + \sigma^2 I\right)^{-1} e}{e^\top \left(\Gamma(\xi^0) + \sigma^2 I\right)^{-1} e}.$$

The function $Q(\lambda)$ is strictly convex, so it has a unique minimum on the convex set $e^\top \lambda = 1$. Therefore in (28) there is a strict inequality unless $\widetilde{\lambda} = \widehat{\lambda}$.

5. Simulation results

Consider the domain $D = [-th, +th] \times [-th, +th] \subset \mathbb{R}^2$ and a Gaussian random field $z_0(x, y)$ on D with the linear mean value function:

$$z_0(x, y) = \beta_0 + \beta_1 x + \beta_2 y + \delta(x, y),$$

where $(x, y) \in D$. The parameters $\beta_0, \beta_1, \beta_2$ are unknown. (For simulation we chose $\beta_0 = 1$, $\beta_1 = 1$, $\beta_2 = -1$.) The field $\delta(x, y)$ is a zero mean Gaussian random field with the Gaussian semivariogram described in Example 2. We assume that the semivariogram is completely known. We used the following values of the parameters: $\omega = \pi/2$, $a^2 = 2$. The error term $\varepsilon(x, y)$ was a Gaussian white noise with variance $\sigma^2 = 0.5^2$. By simulation we generated the field z_0 on the h-lattice points of D and generated the field z at certain random points described below.

We applied three kriging methods: universal kriging, naive kriging and modified kriging (described in Section 3). The universal kriging was applied for data without error. For the universal kriging we used observations of $z_0(x, y)$ at locations $(x, y) = ((-t + 2k)h, (-t + 2l)h)$, $k, l = 0, 1, \ldots, t$. With the help of these observations we predicted the field $z_0(x, y)$ at the remaining h-lattice points $(x, y) = (ih, jh) \in D$. Naive kriging and modified kriging were applied for data with error. For naive and modified kriging we used the observations of the field $z(x, y)$ at random points having independent normal distributions around the previous locations with variance matrix $d^2 I$. With the help of these observations we predicted the field $z_0(x, y)$ at the same points as it was done by the universal kriging. In the modified kriging we chose $R = 2$.

We chose $t = 6$, therefore we had $n = 49$ observations for kriging and we made predictions at 120 points. We remark that at these 120 points the true values of the field z_0 were used to measure the precision of the kriging procedures.

The dispersion d of the location error and the step h were fixed on different values to check the behaviour of different kriging methods. For each fixed set of parameters we made 500 replications.

$h = 0.5,\ d = 0.5$			
	Universal kriging	Naive kriging	Modified kriging
aMSE	$3.6 \cdot 10^{-4}$	3.058	2.217
aVSE	$9.2 \cdot 10^{-7}$	16.092	7.163
$h = 1,\ d = 1$			
	Universal kriging	Naive kriging	Modified kriging
aMSE	$3.4 \cdot 10^{-3}$	3.364	1.756
aVSE	$3.5 \cdot 10^{-5}$	21.343	5.180

Now we describe the abbreviations in the tables below. After making kriging at any fixed point s_0, we calculated the squared error of kriging: $SE = (z_0(s_0) - \widehat{z}_0(s_0))^2$. Here $\widehat{z}_0(s_0)$ is the predicted value of the field z_0. Then we calculated the mean and the variance of the squared errors at the 120 locations. These values are the mean squared error (MSE) and variance of the squared error (VSE). Then we calculated the average of these quantities over 500 replications. $aMSE$ is the average of the mean squared errors, $aVSE$ is the average of the variances of the squared errors. $aVSE$ measures if the prediction fits uniformly well to the surface of our field.

The results show that by naive kriging and modified kriging we can not expect the precision of universal kriging. (But universal kriging was used for data without any observation error!) Modified kriging is better than naive kriging and it is more stable. The simulations were performed with MATLAB.

Acknowledgements

This research was partially supported by the Hungarian Foundation of Scientific Researches under Grant No. OTKA T032361/2000 and Grant No. OTKA T032658/2000. The research was partially realized while A.G. Kukush was visiting University of Debrecen.

References

[1] O. Berke, *On spatiotemporal prediction for on-line monitoring data*, Comm. Statist. Theory Methods **27** (9) (1998), 2343–2369.

[2] N.A.C. Cressie, *Statistics for Spatial Data*, Wiley, New York 1991.

[3] I. Fazekas, S. Baran, A.G. Kukush, and J. Lauridsen, *Asymptotic properties in space and time of an estimator in nonlinear functional errors-in-variables models*, Random Oper. Stoch. Equ. **7** (4) (1999), 389–412.

[4] I. Fazekas and A.G. Kukush, *Errors-in-variables and kriging*, p. 261–273 in: Proc. 4th International Conference on Applied Informatics, Eger 1999.

[5] J. Gabrosek and N. Cressie, *The effect on attribute prediction of location uncertainity in spatial data*, Geographical Analysis **34** (3) (2002), 261–285.

[6] D.G. Krige, *A statistical approach to some basic mine valuations problems on the Witwatersrand*, Journal of the Chemical, Metallurgical and Mining Society of South Africa **52** (1951), 119–139.

[7] S.J. Yakowitz and F. Szidarovszky, *A comparison of kriging with nonparametric regression methods*, J. Multivariate Anal. **16** (1985), 21–53.

Received 10 February 2004
Revised 2 August 2005