

Kriging in the Light of the Theory of Random Field Prediction

Athanasios Dermanis

Department of Geodesy and Surveying,
Aristotle University of Thessaloniki
University Box 503, 54124 Thessaloniki
Email: dermanis@topo.auth.gr,
Web page: <http://der.topo.auth.gr>

Summary

The method of kriging is critically examined from the viewpoint of the classical Wiener-Kolmogorov prediction theory for random fields, as well as from the viewpoint of the finite-dimensional statistical random effects model. It is shown that ordinary kriging is identical with the best homogeneous linear unbiased prediction and that its main characteristic is not the unbiased prediction but rather its homogeneous linear character (a strictly linear combination of the observations without an additional constant). The last argument is emphasized by deriving biased kriging on the basis of best homogeneous linear prediction which is biased.

1. Introduction

Kriging is a method developed in the beginning of the 50s by the mining engineer Krige (1951) in order to predict the ore content of a mining site using isolated measurements at particular points. Ore content is modeled as a random field, i.e. a stochastic function in three dimensions. The more general character of kriging as a method of prediction in relation to an underlying random field has been recognized by Matheron (1962) who studied the deep mathematical problems stemming from the infinite-dimensional character of the unknown random field. Thus the method found wider application in other fields such as hydrology. However similar prediction problems for random fields or stochastic processes (a standard term for functions of time) had already been studied independently by Kolmogorov (1941) and (1949), leading to what may be called a Wiener-Kolmogorov theory for prediction in random functions.

In geodesy a similar method has been introduced by Moritz (Heiskanen & Moritz, 1967) for prediction related to the gravity field of the earth. The method has

been deeply analyzed and generalized by Krarup (1969), who in addition showed its relation to the deterministic problem of interpolation for the harmonic function of the gravitational potential of the earth, which belongs to a Hilbert space with reproducing kernel. The relevant method has been named "collocation" and its deterministic aspects have been further studied by Dermanis (1976), Sansò (1978) and others.

Despite the presence of an underlying infinite-dimensional unknown function in any application, the problem can be reduced to a classical problem of stochastic prediction with finite dimensions in the framework of a so called "random effects model", since the number of data is finite and the problem of predicting the unknown random field can be reduced to that of predicting its value at any particular point of its domain of definition. Moreover the method can be generalized to the use of quantities other than point values for both data and predicted values, which must be however the values of continuous (bounded) linear functionals (mappings from a function to a real number) acting on the unknown field.

Despite its similarities with the Wiener-Kolmogorov prediction, kriging differs in an important aspect: it uses the variogram function instead of the covariance function of the relevant field. From a theoretical point of view this choice extends the applicability of kriging to random fields which the variogram is defined while the covariance function does not. This wider field of application is however insignificant from the viewpoint of practical applications. More important is the applicability of kriging to cases where the random field has unknown but constant mean function, while other methods require knowledge of the relevant constant value.

We shall limit ourselves here to the so called "ordinary kriging" with unknown constant mean function. "Universal kriging" where the unknown mean function is modeled as a linear combination of known (base)

function with unknown coefficients can also be treated in the framework of the classical finite-dimensional estimation-prediction statistical methodology where the so called “mixed effects model” is used. Nevertheless the essence of the comparisons and conclusions that will be drawn here does not need the generalization of universal kriging, which leads to somewhat more complicated algorithms which usually implement the variogram instead of the covariance function. More drastic is the generalization of the “intrinsic kriging”, which leads to solutions independent of the unknown mean value function by utilizing the so called “generalized covariance function”. Finally a recent generalization is the “generalized kriging” of Reguzzoni et al. (2005), which (as already done for geodetic collocation) allows the use of almost any real values related to the unknown random field, either as input data or as predicted quantities, provided that they can be expressed as continuous functionals of the underlying unknown field. On the basis of the comparison in the framework of the statistical random effects model still another generalization is possible, the one of “biased kriging” already proposed by Dermanis & Sansò (2007).

2. Prediction with the Random Effects Model

The random effects model is a linear model of the form $\mathbf{y} = \mathbf{G}\mathbf{s} + \mathbf{v}$, where \mathbf{s} is a vector of random variables with known mean values $E\{\mathbf{s}\} = \mathbf{m}$ and known covariance matrix $E\{(\mathbf{s} - \mathbf{m})(\mathbf{s} - \mathbf{m})^T\} = \mathbf{C}_{ss}$, \mathbf{v} is the vector of the observational errors with $E\{\mathbf{v}\} = \mathbf{0}$ and $E\{\mathbf{v}\mathbf{v}^T\} = \mathbf{C}_v$, \mathbf{G} is a known matrix and \mathbf{y} is the random vector of the observations for which a sample outcome is known as the result of specific measurements. The problem is the prediction (i.e. the estimation of the corresponding sample value) of any random variable s' with known mean value $E\{s'\} = m'$, which is correlated with the model random variables, through the known (cross)covariance matrix $E\{(\mathbf{s} - \mathbf{m})(s' - m')\} = \mathbf{c}_{ss'}$. The prediction is characterized as optimal (best) when it satisfies the criterion of minimizing the mean square prediction error $E\{\varepsilon^2\} = \min$, where $\tilde{s}' = \tilde{s}'(\mathbf{y})$ is the prediction and $\varepsilon = \tilde{s}' - s'$ prediction error. The prediction itself is a function $\tilde{s}' = \tilde{s}'(\mathbf{y})$ of the known observations \mathbf{y} , which is typically linear with two possible values the homogeneous (hom) linear one $s' = \mathbf{d}^T \mathbf{y}$ and the non-homogeneous (inhom) one $s' = \mathbf{d}^T \mathbf{y} + k$. Predictions are further distinguished into unbiased ones for which $E\{\tilde{s}'\} = E\{s'\}$ and biased ones for which this additional restriction does not hold. On the basis of the

above choices we may distinguish between four types of optimal linear predictions:

- inhomBLUP = inhomogeneous Best Linear Unbiased
- homBLUP = homogeneous Best Linear Unbiased Prediction
- inhomBLIP = inhomogeneous Best Linear Prediction
- homBLIP = homogeneous Best Linear Prediction

The optimal values of the coefficients \mathbf{d} or \mathbf{d} and k are determined by minimizing the function $E\{\varepsilon^2\} = \varphi(\mathbf{d})$ or $E\{\varepsilon^2\} = \varphi(\mathbf{d}, k)$, respectively, either directly or under the condition $\mathbf{d}^T \mathbf{G}\mathbf{m} = m'$ or $\mathbf{d}^T \mathbf{G}\mathbf{m} + k = m'$ for unbiased prediction. Based on the resulting values of \mathbf{d} and k we have the following prediction equations:

inhomBLUP = inhomBLIP:

$$\tilde{s}' = m' + \mathbf{c}_{ss'}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} (\mathbf{y} - \mathbf{G}\mathbf{m})$$

homBLUP:

$$\tilde{s}' = \alpha m' + \mathbf{c}_{ss'}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} (\mathbf{y} - \alpha \mathbf{G}\mathbf{m}),$$

$$\alpha = \frac{\mathbf{m}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{m}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} \mathbf{G}\mathbf{m}} \quad (2)$$

homBLIP:

$$\tilde{s}' = \alpha m' + \mathbf{c}_{ss'}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} (\mathbf{y} - \alpha \mathbf{G}\mathbf{m}),$$

$$\alpha = \frac{\mathbf{m}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} \mathbf{y}}{1 + \mathbf{m}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} \mathbf{G}\mathbf{m}} \quad (3)$$

We notice that the above predictors have the same form where only the parameter α differs, which can be taken as $\alpha = 1$ for the inhomBLUP = inhomBLIP prediction. In the case of zero mean values ($m_i = 0$, $m' = 0$) all the above predictions coincide taking the common form $\tilde{s}' = \mathbf{c}_{ss'}^T \mathbf{G}^T (\mathbf{G}\mathbf{C}_{ss}\mathbf{G}^T + \mathbf{C}_v)^{-1} \mathbf{y}$. In the case of non-homogeneous prediction the presence of the constant term k makes the bias disappear and the relevant prediction is unbiased even when this is not an a priori requirement, so that biased and unbiased predictions coincide. Homogeneous predictions have been proposed by Schaffrin (see e.g.. Grafrend & Schaffrin, 1993) as “robust” alternatives to the standard non-homogeneous ones since the former multiply the mean value with a factor α , which in a certain sense offers some type of partial protection against wrong assumptions about the mean values. Similar somewhat generalized expressions have been proposed by Dermanis (1987) in the framework of stan-

ard non-homogeneous prediction by including the unknown factor in the original model, either as a deterministic parameter or as a stochastic one with mean value 1 and known a priori variance.

3. Application to Random Field Prediction

In the case of an unknown random field $u(\mathbf{x})$, where e.g. $\mathbf{x} = [x \ y \ z]^T$ is the vector of Cartesian coordinates, the observations \mathbf{y} have the form $y_i = u(\mathbf{x}_i) + v_i$, $i = 1, 2, \dots, n$, i.e. the values $u(\mathbf{x}_i)$ of the field at n points \mathbf{x}_i are observed with additional random observational errors v_i , while it is required to provide a prediction $\tilde{u}_{\mathbf{x}} \equiv \widetilde{u(\mathbf{x})}$ of the value of the field $u(\mathbf{x})$ at any other point \mathbf{x} . We assume that the random field has constant mean value function $m(\mathbf{x}) \equiv E\{u(\mathbf{x})\} = \mu$ and known covariance function $C(\mathbf{x}, \mathbf{x}') \equiv E\{[u(\mathbf{x}) - \mu][u(\mathbf{x}') - \mu]\}$.

In order to compare the above results with kriging we have assumed a constant mean value $m(\mathbf{x}) = \mu$, an assumption which becomes necessary if we further assume that the random field is homogeneous, i.e. that $C(\mathbf{x}, \mathbf{x}') = C(\mathbf{x} + \mathbf{h}, \mathbf{x}' + \mathbf{h})$ for any displacement \mathbf{h} , in which case (choosing $\mathbf{h} = -\mathbf{x}$) the covariance function becomes a function of only the difference $\mathbf{x}' - \mathbf{x}$, thus having the form $C(\mathbf{x}' - \mathbf{x})$.

Setting $u_i = u(\mathbf{x}_i)$ and $C_{ik} = C(u(\mathbf{x}_i), u(\mathbf{x}_k))$, $c_i = C(u(\mathbf{x}), u(\mathbf{x}_i))$, we arrive at the random effects model $\mathbf{y} = \mathbf{u} + \mathbf{v}$, which corresponds to the more general model $\mathbf{y} = \mathbf{G}\mathbf{s} + \mathbf{v}$ with $\mathbf{G} = \mathbf{I}$, $\mathbf{s} = \mathbf{u}$, $s' = u(\mathbf{x}) \equiv u_{\mathbf{x}}$, $\mathbf{C}_{ss} = \mathbf{C}$, $\mathbf{c}_{s'} = \mathbf{c}$, $m' = \mu$, $\mathbf{m} = \mu\mathbf{s}$, where $\mathbf{s} = [1 \ 1 \ \dots \ 1]^T$. With these replacements the relevant predictions become

inhomBLUP = inhomBLIP:

$$\tilde{u}_{\mathbf{x}} = \mu + \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{y} - \mu\mathbf{s}) \quad (4)$$

homBLUP:

$$\tilde{u}_{\mathbf{x}} = \alpha \mu + \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{y} - \alpha \mu\mathbf{s}),$$

$$\alpha = \frac{1}{\mu} \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} \quad (5)$$

homBLIP:

$$\tilde{u}_{\mathbf{x}} = \alpha \mu + \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{y} - \alpha \mu\mathbf{s}),$$

$$\alpha = \frac{\mu \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{1 + \mu^2 \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} \quad (6)$$

4. Kriging as Best Linear Homogeneous Unbiased Prediction

As already pointed out by Dermanis (1984), kriging corresponds to best linear homogeneous prediction (homBLUP), while geodetic collocation corresponds to best linear inhomogeneous unbiased prediction (inhomBLUP). Therefore what distinguishes kriging is its homogeneous linear character rather than its unbiased one. In order to prove this statement we must translate the relevant result from the "language" of the covariance function to the one of the variogram, which is defined as

$$\gamma(\mathbf{x}, \mathbf{x}') = \frac{1}{2} E\{[u(\mathbf{x}) - u(\mathbf{x}')]^2\} \quad (7)$$

In kriging the so called "intrinsic" hypothesis holds, namely that the variogram is a function of only the displacement $\mathbf{h} = \mathbf{x}' - \mathbf{x}$, which corresponds to the hypothesis of a homogeneous field with covariance $C(\mathbf{x}, \mathbf{x}') = C(\mathbf{x}' - \mathbf{x})$, so that

$$\begin{aligned} \gamma(\mathbf{h}) &= \frac{1}{2} E\{[u(\mathbf{x} + \mathbf{h}) - u(\mathbf{x})]^2\} = \\ &= \frac{1}{2} E\{u(\mathbf{x} + \mathbf{h})^2\} - E\{u(\mathbf{x} + \mathbf{h})u(\mathbf{x})\} + \frac{1}{2} E\{u(\mathbf{x})^2\} = \\ &= C(\mathbf{0}) - C(\mathbf{h}) \end{aligned} \quad (8)$$

Under the intrinsic hypothesis the mean value function is necessarily constant. Introducing the variogram related matrices Γ , γ with elements $\Gamma_{ik} = \gamma(\mathbf{x}_i - \mathbf{x}_k)$, $\gamma_i = \gamma(\mathbf{x} - \mathbf{x}_i)$, respectively and setting $C_0 \equiv C(\mathbf{0})$, we obtain the transformation relations

$$\mathbf{C} = C_0 \mathbf{s}\mathbf{s}^T - \Gamma, \quad \mathbf{c} = C_0 \mathbf{s} - \gamma$$

as well as their inverse ones

$$\Gamma = C_0 \mathbf{s}\mathbf{s}^T - \mathbf{C}, \quad \gamma = C_0 \mathbf{s} - \mathbf{c}.$$

In order to show the correspondence of kriging with the best linear homogeneous unbiased prediction we will depart from the well known "kriging system" of the form

$$\begin{bmatrix} \Gamma - \mathbf{C}_v & \mathbf{s} \\ \mathbf{s}^T & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ k \end{bmatrix} = \begin{bmatrix} \gamma \\ 1 \end{bmatrix} \quad (9)$$

or explicitly

$$(\Gamma - \mathbf{C}_v)\lambda + k\mathbf{s} = \gamma, \quad \mathbf{s}^T \lambda = 1, \quad (10)$$

which solved for λ leads to the kriging prediction $\hat{u}(\mathbf{x}) = \lambda^T \mathbf{y}$. Replacing $\Gamma = C_0 \mathbf{s} \mathbf{s}^T - \mathbf{C}$, $\boldsymbol{\gamma} = C_0 \mathbf{s} - \mathbf{c}$, the kriging system becomes $\mathbf{s}^T \lambda = 1$ and $(C_0 \mathbf{s} \mathbf{s}^T - \mathbf{C} - \mathbf{C}_v) \lambda + k \mathbf{s} = C_0 \mathbf{s} \mathbf{s}^T \lambda - (\mathbf{C} + \mathbf{C}_v) \lambda + k \mathbf{s} = C_0 \mathbf{s} - (\mathbf{C} + \mathbf{C}_v) \lambda + k \mathbf{s} = C_0 \mathbf{s} - \mathbf{c}$, or simply

$$(\mathbf{C} + \mathbf{C}_v) \lambda - k \mathbf{s} = \mathbf{c}, \quad \mathbf{s}^T \lambda = 1. \quad (11)$$

The first relation gives $\lambda = (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{c} + k \mathbf{s})$, which replaced in the second one yields

$$\begin{aligned} \mathbf{s}^T \lambda &= \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{c} + k \mathbf{s}) = \\ &= \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{c} + k \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s} = 1, \end{aligned}$$

so that $k = \frac{1 - \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{c}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}}$. With this value of k the coefficients λ become

$$\begin{aligned} \lambda &= (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{c} + k \mathbf{s}) = \\ &= (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{c} + \frac{1 - \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{c}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s} \quad (12) \end{aligned}$$

and the corresponding prediction $\hat{u}(\mathbf{x}) = \lambda^T \mathbf{y}$ takes the form

$$\begin{aligned} \hat{u}(\mathbf{x}) &= \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y} + \frac{1 - \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{c}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} \mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y} = \\ &= \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y} + \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} - \\ &\quad - \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s} \\ &= \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} + \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \left[\mathbf{y} - \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}} \mathbf{s} \right] \quad (13) \end{aligned}$$

or simply

$$\begin{aligned} \hat{u}(\mathbf{x}) &= \beta + \mathbf{c}^T (\mathbf{C} + \mathbf{C}_v)^{-1} (\mathbf{y} - \beta \mathbf{s}), \\ \beta &= \frac{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{y}}{\mathbf{s}^T (\mathbf{C} + \mathbf{C}_v)^{-1} \mathbf{s}}. \quad (14) \end{aligned}$$

Comparing the above relation with equation (5) for the best linear homogeneous prediction, it is easy to certify that the two methods coincide if we recognize that $\beta = \alpha \mu$. Therefore ordinary kriging is identical to best linear homogeneous prediction applied to the case of random fields having a covariance function. This identification is not an absolute one for two reasons:

- Kriging is somewhat more general because it applies also to random fields having a variogram but no covariance function..
- Kriging does not require knowledge of the constant mean value μ of the relevant random field.

On first site the second of the above remarks appears not to be valid because the presence of μ in equation (5) is not a real one, a fact that is verified by the alternative relation (14) where the product $\alpha \mu$ has been replaced by the single parameter $\beta = \alpha \mu$. However knowledge of the value μ is in fact related not to the realization of the prediction but rather to the approximate determination of the variogram $\gamma(\mathbf{h})$ or the covariance function $C(\mathbf{h})$, in each case, on the basis of the corresponding relations

$$\gamma(\mathbf{h}) = \frac{1}{2} E\{[u(\mathbf{x}) - u(\mathbf{x} + \mathbf{h})]^2\} \quad (15)$$

$$C(\mathbf{h}) = E\{[u(\mathbf{x}) - \mu][u(\mathbf{x} + \mathbf{h}) - \mu]\} \quad (16)$$

Such approximations are based on the separation of the domain of definition D in covering independent subsets D_m ($\cup D_m = D$, $D_m \cap D_{m'} = \emptyset$ for $m' \neq m$) and the approximation of $\gamma(\mathbf{h})$ or $C(\mathbf{h})$ by step functions with constant values in every subset D_m ($\gamma(\mathbf{h}) = \gamma_m$, $\forall \mathbf{h} \in D_m$ and $C(\mathbf{h}) = C_m$, $\forall \mathbf{h} \in D_m$). Assuming uncorrelated random observation errors with the same variance $E\{v_i v_k\} = \sigma_v^2 \delta_{ik}$, the values γ_m or C_m are obtained from the values of the observations $y_i = u(\mathbf{x}_i) + v_i$ using the relations

$$2\hat{\gamma}_m + 2\sigma_v^2 = \frac{1}{N_m} \sum_{\mathbf{x}_k - \mathbf{x}_i \in D_m} (y_i - y_k)^2 \quad (17)$$

$$\hat{C}_m = \frac{1}{N_m} \sum_{\mathbf{x}_k - \mathbf{x}_i \in D_m} (y_i - \mu)(y_k - \mu) \quad (18)$$

Where N_m the number of point pairs with $\mathbf{x}_k - \mathbf{x}_i \in D_m$. The relevant estimates are unbiased, i.e. $E\{\hat{\gamma}_m\} = \gamma_m$ and $E\{\hat{C}_m\} = C_m$. For the variogram $\gamma(\mathbf{0}) = 0$ by definition, while the corresponding value $C_0 = C(\mathbf{0})$ is estimated using

$$\hat{C}_0 + \sigma_v^2 = \frac{1}{N_m} \sum_{\mathbf{x}_i} (y_i - \mu)^2 \quad (19)$$

with $E\{\hat{C}_0\} = C_0$.

5. Biased Kriging

In order to emphasize that the main characteristic of kriging is that of being a homogeneous linear prediction and not that of being an unbiased one, we will “translate” the (biased) best linear homogeneous prediction (6) into the “language” of the variogram using the transforming relations $\mathbf{C} = C_0 \mathbf{s} \mathbf{s}^T - \Gamma$, $\mathbf{c} = C_0 \mathbf{s} - \gamma$. Replacing the matrices \mathbf{C} and \mathbf{c} of equation (6) with the above values we arrive (after some extensive algebraic manipulation) to the relation for biased kriging

$$\hat{u}(\mathbf{x}) = \beta + \gamma^T (\Gamma - \mathbf{C}_v)^{-1} (\mathbf{y} - \beta \mathbf{s}),$$

$$\beta = \frac{H \mathbf{s}^T (\Gamma - \mathbf{C}_v)^{-1} \mathbf{y}}{H \mathbf{s}^T (\Gamma - \mathbf{C}_v)^{-1} \mathbf{s} - 1} \quad (20)$$

which contains instead of the mean value μ , the parameter $H \equiv C_0 + \mu^2$. Although H contains the mean value μ , it can be estimated directly from the observations using the relation

$$\hat{H} = \hat{C}_0 + \sigma_v^2 = \frac{1}{N_m} \sum_{x_i} y_i^2 \quad (21)$$

where $E\{\hat{H}\} = H$, i.e. we have an unbiased estimate.

References

- Dermanis, A., 1976. *Probabilistic and Deterministic Aspects of Linear Estimation in Geodesy*. Report No. 244, Dep. of Geodetic Science, The Ohio State University.
- Dermanis, A., 1984. Kriging and Collocation - A Comparison. *Manuscripta Geodetica*, vol. 9, no. 2, 159-167.
- Dermanis, A., 1987. *Geodetic Applications of Interpolation and Prediction*. International School of Geodesy "A. Marussi". 4th Course: "Applied and Basic Geodesy: Present and Future Trends". Ettore Majorana Centre for Scientific Culture, Erice-Sicily, 15-25 June 1987. *Eratosthenes*, 22, 229-262.
- Dermanis, A. and F. Sansò, 2007. *On the Feasibility of Biased Kriging*. Presented at the XXIV IUGG Congress, Perugia X-X July 2007 (to be published).
- Grafarend, E. and B. Schaffrin, 1993. *Ausgleichsrechnung in linearen Modellen*. Bibliographisches Institut Wissenschaftsverlag, Mannheim.
- Heiskanen W. and H. Moritz, 1967. *Physical Geodesy*. W. Freeman and Co., San Francisco.
- Karup, T., 1969. *A Contribution to the Mathematical Foundation of Physical Geodesy*. *Geodætisk Instituts*, Med. No. 44, Copenhagen.
- Kolmogorov, A.N., 1941. *Interpolation and extrapolation of stationary random sequences*. *Izvestiia Akademii Nauk SSR, Serii Matematicheskaia*, vol. 5, p. 3-14. Translation: Memo RM-3090-PR, Rand Corporation, Santa Monica, California, 1962.
- Krige, D. G., 1951, *A statistical approach to some basic mine valuation problems on the Witwatersrand*. *J. Chem. Metal. Min. Soc. South Africa*, v. 52, p. 119-139.
- Matheron, G., 1962, *Traité de géostatistique appliquée, vol. I: Mémoires du Bureau de Recherches Géologiques et Minières*, no. 14, Editions Technip, Paris, 333 pp.
- Reguzzoni, M., F. Sansò and G. Venuti, 2005. The theory of general kriging, with applications to the determination of a local geoid. *Geophys. J. Int.* 162, 303-314.
- Sansò, F., 1978. The Minimum Mean Square Estimation Error Principle in Physical Geodesy (Stochastic and Non-Stochastic Interpretation)". *Proceedings 7th Symposium on Mathematical Geodesy*, Assisi, 1978.
- Wackernagel, H., 2003. *Multivariate Geostatistics. An Introduction with Applications*. 3rd completely revised edition. Springer-Verlag, Berlin, Heidelberg.
- Wiener, N., 1949. *Extrapolation, interpolation and smoothing of stationary time series*: MIT Press, Cambridge, Massachusetts, 158 pp.

Originally published in:

Hydrogea. A volume in honor of Prof. Christos Jimopoulos, pp. 447-456. Publication of the School of Rural and Surveying Engineering, Aristotle University of Thessaloniki (in Greek).