CO-KRIGING - NEW DEVELOPMENTS

Donald E. Myers
Department of Mathematics
University of Arizona
Tucson, Arizona 85721

Abstract

Co-Kriging or Joint Estimation utilizes data from correlated variables to improve the estimation of all variables or to compensate for missing data on some variables. The general formulation of Co-Kriging in matrix form was given by the author. The matrix form emphasizes the analogy with Kriging of one variable utilizing only spatial dependence. General conditions are obtained for covariance matrix functions and variogram matrix functions. The extension to block co-kriging is delineated including the Co-Kriging variance. A simple algorithm is given for obtaining the "under-sampled" case from the general matrix formulation. Finally a method for reducing the size of the system of equations is given and a iterative method provided which allows solution of even singular systems in which entries are matrices.

INTRODUCTION

In (2), (3) the author has given in matrix form the general formulation of co-kriging and shown that it is a direct extension of kriging in that form. In this paper we extend those results in several ways. Section one is concerned with the general problem of the variance of vector linear combinations and conditions necessary for a matrix function to be a covariance or a variogram.

G. Verly et al. (eds.), Geostatistics for Natural Resources Characterization, Part 1, 295–305.

Although the extension to block estimation does not present any significant difficulties, it was not included in (2), (3), and for the sake of completeness is covered in Section two.

The "under sampled" case is the one that has received the most attention in mining applications but only for a few variables and a few sample locations. Section three describes a simple algorithm for obtaining the system of equations in matrix form for any "under-sampled" problem from the general form of Co-Kriging.

Finally Section four introduces methods for solving the large systems of equations that are generated in the use of Co-Kriging.

NOTATION

The notation used in (3) will be followed here. Recall that Tr denotes the trace, i.e. the sum of the diagonal entries and A^T denotes the transpose of A .

VARIANCE OF VECTOR LINEAR COMBINATIONS

One of the singular characteristics of Ordinary Kriging, i.e., the use of variograms and IRF-0's is that certain linear combinations can have finite variance even though the random function does not have finite variance. One of the properties of a variogram is the following

$$- \sum \sum \gamma(x_i - x_j)\lambda_i \lambda_j \geq 0 \qquad (1)$$

for all weights $\lambda_1, \ldots, \lambda_n$ such that $\lambda_1 + \ldots + \lambda_n = 0$. That is, $-\gamma$ is conditionally positive definite. In the case of a second order stationary random function the covariance must be positive definite. For Co-Kriging it is then appropriate to ask what kind of matrix functions can be covariance matrices or variogram matrices.

Let K(h) be an m x m matrix with entries $k_{ij}(h)$. In the case of second order stationary random functions K(h) should satisfy

$$\mathrm{Tr}\ E[\sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_i^T \bar{Z}(x_i)^T \bar{Z}(x_j)\ \Gamma_j]$$

$$= \mathrm{Tr}[\sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_i^T K(x_i - x_j)\ \Gamma_j] \geq 0 \qquad (2)$$

where we assume without loss of generality that

$$E[Z(x)] = [0, \dots, 0]$$

If the Z_i's are IRF-0's then K(h) should satisfy

$$\mathrm{Tr}E\left[\sum_{j=1}^{n}\sum_{i=1}^{n}\Gamma_i^T\bar{Z}(x_i)^T\bar{Z}(x_j)\Gamma_j\right]$$

$$= -\mathrm{Tr}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\Gamma_i^T K(x_i-x_j)\Gamma_j\right] \geq 0 \qquad (3)$$

for all $\Gamma_1, \dots, \Gamma_n$ with

$$\sum_{j=1}^{n}\Gamma_j = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix} \qquad (4)$$

Both cases may be considered at the same time. Let Γ^j_{st} be an m x m matrix whose only non-zero entry is λ^j_{st}. Then

$$\Gamma_j = \sum_{j=1}^{n}\sum_{t=1}^{n}\Gamma^j_{st} \qquad (5)$$

Substituting in (2) or (3) and recalling that (2) must be satisfied for all Γ's or (3) must be satisfied for all Γ_j's satisfying (4), it is necessary and sufficient that

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda^i_{ss}\,k_{ss}(x_i-x_j)\,\lambda^j_{ss} +$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda^i_{ss}k_{st}(x_i-x_j)\,\lambda^j_{tt} +$$

$$\sum\sum\lambda^i_{tt}\,k_{tt}(x_i-x_j)\,\lambda^j_{tt}$$

$$\geq 0 \qquad (6)$$

$$\leq 0 \qquad (7)$$

for all s, t . (6) corresponds to (2) and (7) to (3). This can be described somewhat simpler.

$$\lambda^1_{ss}, \lambda^2_{ss}, \dots, \lambda^n_{ss}$$

are diagonal entries in $\Gamma_1, \dots, \Gamma_n$ respectively. $\lambda^1_{tt}, \dots, \lambda^n_{tt}$ are also diagonal entries. Let

$$A_s = [\lambda_{ss}^1, \ldots, \lambda_{ss}^n] \quad (8)$$

$$B_t = [\lambda_{tt}^1, \ldots, \lambda_{tt}^n] \quad (9)$$

$$K_{uv} = \begin{bmatrix} k_{uv}(x_1-x_1) \ldots k_{uv}(x_1-x_n) \\ \vdots \qquad\qquad \vdots \\ k_{uv}(x_1-x_n) \ldots k_{uv}(x_n-x_n) \end{bmatrix} \quad (10)$$

then (16) can be written

$$A_s K_{ss} A_s^T + A_s K_{st} B_t^T + B_t K_{tt} B_t T > 0 \quad (11)$$

and (17) as

$$A_s K_{ss} A_s^T + A_s K_{st} B_t^T + B_t K_{tt} B_t^T \leq 0 \quad (12)$$

$$\sum_{i=1}^n \lambda_{ss}^i = 0 \ , \ \sum_{i=1}^n \lambda_{tt}^i = 0 \quad (13)$$

It is clearly necessary that

$$A_s K_{ss} A_s^T > 0$$

$$B_t K_{tt} B_t^T > 0 \quad (14)$$

which are consequences of (21). (24) is the usual positive definite condition for covariance functions.

BLOCK CO-KRIGING

In (2), (3), (4) the author indicated that the formulation of Co-Kriging given would extend to estimation of block values, the results were not given. For the sake of completeness they are icluded herein. Let V be a volume, area or length and let

$$\bar{Z}_V = [\frac{1}{V}\int_V Z_1(x)dx, \ldots, \frac{1}{V}\int_V Z_m(x)dx] \quad (15)$$

If $\bar{Z}_V^* = \sum_{k=1}^n Z(x_k)\Gamma_k$, the problem as before is to determine the Γ_k's so that Z_V^* is unbiased and the estimation error has minimum variance.

Stationary/Covariance case

If $Z_1, Z_2, \ldots, Z_m$ are second order stationary and

$$\bar{C}(h) = \begin{bmatrix} C_{11}(h) \ldots C_{1m}(h) \\ \vdots \qquad \vdots \\ C_{m1}(h) \ldots C_{mm}(h) \end{bmatrix}$$

then

$$E[\bar{Z}(x_i)]^T[\bar{Z}_V] = = \frac{1}{V}\int_V \bar{C}(x_i - x)dx = \bar{C}(x_i, V) \quad (16)$$

and

$$E[\bar{Z}_V]^T(\bar{Z}(x_j)] = \frac{1}{V}\int_v \bar{C}(x - x_j)dx = \bar{C}(v, x_j) \quad (17)$$

$$E[\bar{Z}_V)^T[\bar{Z}_V] = \frac{1}{V^2}\int_V\int_V \bar{C}(x-y)dxdy = \bar{C}(V,V) \quad (18)$$

The Kriging equations become

$$\sum_{j=1}^{n} \bar{C}(x_i x_j)\,\Gamma_j + \bar{\mu} = \bar{C}(x_i, V), \quad \sum_{j=1}^{n} \Gamma_j = I \quad (19)$$

with Kriging variance

$$\sigma^2_{CK} = \mathrm{Tr}\,\bar{C}(V,V) - \mathrm{Tr}\sum \bar{C}(x_j, V)\,\Gamma_j - \mathrm{Tr}\bar{\mu} \quad (20)$$

and as with punctual Co-Kriging the component corresponding to each Z_i may be selected out.

The Intrinsic Case

If the Z_i's are IRF-0's then we may write

$$\bar{Z}_V - Z^*_V = \sum_{k=1}^{n} \frac{1}{V}\int_V [\bar{Z}^* x(- \bar{Z}(x_k)]dx \quad (21)$$

and $E\,[\bar{Z}_V - \bar{Z}^*_V]^T[\bar{Z}_V - \bar{Z}^*_V]$ becomes

$$= \sum\sum \Gamma^T_i \frac{}{V^2} \int\int E[\bar{Z}(x) - \bar{Z}[x_i)]^T[\bar{Z}(y) - \bar{Z}(x_j)]dxdy\,\Gamma_j$$

$$= \sum\sum \Gamma_i^T \frac{1}{V^2}\int_V\int_V [\bar{\gamma}(x - x_j) + \bar{\gamma}(x_i - y) - \bar{\gamma}(x-y)$$

$$- \bar{\gamma}(x_i x_j)]dxdy\ \Gamma_j \tag{22}$$

From (22) it is easily seen that the Kriging equations are

$$\sum_{j=1}^{m} \bar{\gamma}(x_i - x_j)\ \Gamma_j + \bar{\mu} = \bar{\gamma}(x_i, V) \quad , \quad \sum \Gamma_j = I \tag{23}$$

where $\bar{\gamma}(x_i, V) = \frac{1}{V} \int_V \bar{\gamma}(x - x_i)dx$ (24)

If the samples have nonpunctual support then $\bar{\gamma}(x_i, x_j)$ is replaced by

$$\frac{1}{V_i V_j} \int_{V_i} \int_{V_j} \bar{\gamma}(x-y)dxdy \tag{25}$$

V_i, V_j being the supports at x_i, x_j .

The Co-Kriging Variance is

$$\sigma^2_{CK} = \mathrm{Tr} \sum_{1}^{m} \bar{\gamma}(x_i, V)\ \Gamma_i - \mathrm{Tr}\ \bar{\mu} - \mathrm{Tr}\ \frac{1}{V^2} \int_V \int_V \bar{\gamma}(x-y)dxdy \tag{26}$$

Since the off-diagonal entries in

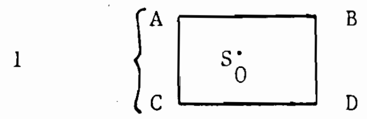
$$\frac{1}{V^2} \int_V \int_V \bar{\gamma}(x-y)dxdy \tag{27}$$

are not required otherwise the last term in σ^2_{CK} is simply

$$\sum_{i=1}^{n} \frac{1}{V^2} \int_V \int_V \gamma_{ii}(x-y)dxdy \tag{28}$$

THE "UNDER-SAMPLED" CASE

The following example is found in (1) Journel.



Data $Z_1(0)$,
$Z_2(0)$, $Z_2(A)$, $Z_2(B)$, $Z_2(C)$
$Z_2(D)$

V is the square with vertices A,B,C,D. It is desired to estimate Z_V where

$$Z_V = \frac{1}{V} \int_V Z_1(x)dx$$

using all the data. In (1), the problem was simplified by using the symmetry i.e.
$Z_2(S_R) = \frac{1}{4} [Z_2(A) + Z_2(B) + Z_2(C) + Z_2(D)]$. The system of equations given in (1) is a sub-system of the following

$$\begin{bmatrix} \bar{\gamma}(S_0,S_0) & \bar{\gamma}^*(S_0,S_R) & I \\ \bar{\gamma}_*(S_0,S_R) & \bar{\gamma}_*^*(S_R,S_R) & I_* \\ I & I^* & 0 \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \bar{\mu} \end{bmatrix} = \begin{bmatrix} \bar{\gamma}(S_0,V) \\ \gamma_*(\bar{S}_R V) \\ I \end{bmatrix} \quad (29)$$

where

$$\bar{\gamma}^*(S_0,S_R) = \begin{bmatrix} 0 & \gamma_{12}(S_0,S_R) \\ 0 & \gamma_{22}(S_0,S_R) \end{bmatrix} \quad (30)$$

$$\bar{\gamma}_*^*(S_R,S_R) = \begin{bmatrix} 0 & 0 \\ 0 & \gamma_{22}(S_R,S_R) \end{bmatrix} \quad (31)$$

$$\bar{\gamma}_*(S_0,S_R) = \begin{bmatrix} 0 & 0 \\ 0 & \gamma_{22}(S_R,S_R) \end{bmatrix} \quad (32)$$

$$I_* = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad I^* = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (33)$$

This system also provides for the estimation of Z_{2V} along with Z_{1V} . Moreover λ^2_{11} , λ^2_{12} will be arbitrary and hence may be taken to be zero, that is $Z_1(S_R)$ is not used in the estimation of Z_{1V} or Z_{2V} . In a geometric/graphical way then it is easy to see how the general Co-Kriging system is changed to the under-sampled version.

$$\begin{bmatrix} \bar{\gamma}(x_1-x_1) \cdot \cdot \cdot \bar{\gamma}(x_1-x_n) \\ \vdots \qquad\qquad \vdots \\ \bar{\gamma}(x_1-x_n) \cdot \cdot \cdot \gamma(x_n-x_n) \\ I \; \cdot \cdot \cdot \cdot \cdot \; I \end{bmatrix} \begin{bmatrix} I & \Gamma_1 \\ & \vdots \\ I & \Gamma_n \\ 0 & \bar{\mu} \end{bmatrix} = \begin{bmatrix} \bar{\gamma}(V,x_1) \\ \vdots \\ \bar{\gamma}(V,x_n) \\ I \end{bmatrix} \quad (34)$$

Each column/row in (35) corresponds to a sample location. Each column/row within a $\bar{\gamma}$ corresponds to a variable. For each variable that is not sampled at all locations, locate the column/row corresponding to the location, then in each $\bar{\gamma}$ (and I) change all of the column/row entries corresponding to the variable. In the case of the rows the deletion is applied on the right side of (35) as well. To see that constraining the variogram matrices is equivalent to constraining certain weights to be zero is easily seen.

For simplicity suppose Z_{i_0} is not sampled at x_{j_0}. Then Γ_{j_0} has additional constraints imposed namely

$\lambda^{j_0}_{i_0 1}, \ldots, \lambda^{j_0}_{i_0 m}$ are all zeros.

Let $\hat{\Gamma}_{j0}$ be the modified Γ_{j0}. Similarly let $\bar{\gamma}_*(x_{j0}-x_p)$ be the same as $\bar{\gamma}(x_{j0}-x_p)$ except that in the i_0 row all entries are zeros and $\bar{\gamma}^*(x_p-x_{j0})$ is the same as $\bar{\gamma}(x_p-x_{j0})$ except that all entries in the i_0 column are all zeros, that is, $\bar{\gamma}^*(x_p-x_{j0}) = \bar{\gamma}_*(x_{j0}-x_p)^T$. Let I_* be an identity matrix except that the i_0 row is all zeros and $(I^*)^T = I_*$.

If $\bar{Z}$ is a vector IRF-0 then the estimation variance in the under-sampled case is

$$
\begin{aligned}
&\mathrm{Tr}[\sum_{j\neq j0} \Gamma_j^T \bar{\gamma}(x_j-x) + \hat{\Gamma}_{j0}\, \bar{\gamma}(x_{j0}-x)] \\
&+ \mathrm{Tr}[\sum_{j\neq j0} \bar{\gamma}(x-x_j)\Gamma_j + \bar{\gamma}(x-x_{j0})\hat{\Gamma}_{j0}] \\
&+ \mathrm{Tr}[\sum_{i\neq j0}\sum_{j\neq j0} \Gamma_i^T \bar{\gamma}(x_n-x_j)\,\Gamma_j \\
&+ \sum_{i\neq j0} \Gamma_i^T \bar{\gamma}(x_i-x_{j0})\,\hat{\Gamma}_{j0} \\
&+ \sum_{j\neq j0} \hat{\Gamma}_{j0}^T \bar{\gamma}(x_{j0}-x_i)\,\Gamma_j \\
&+ \hat{\Gamma}_{j0}^T \bar{\gamma}(x_{j0}-x_{j0})\,\Gamma_{j0}]
\end{aligned} \qquad (35)
$$

The universality conditions can be written

$$\sum_{j\neq j0} I\Gamma_j + I\hat{\Gamma}_{j0} = I \qquad (36)$$

Consider the following identities

(i) $$I\hat{\Gamma}_{j0} = I_*\Gamma_{j0} \qquad (37)$$

(ii) $$\hat{\Gamma}_{j0}{}^{T}\bar{\gamma}(x_{j0}-x) = \Gamma_{j0}{}^{T}\bar{\gamma}_*(x_{j0}-x)$$

(iii) $$\bar{\gamma}(x-x_{j0})\hat{\Gamma}_{j0} = \bar{\gamma}*(x-x_{j0})\Gamma_{j0}$$

(iv) $$\hat{\Gamma}_{j0}{}^{T}\bar{\gamma}(x_{j0}-x_j)\hat{\Gamma}_j = \Gamma_{j0}{}^{T}\bar{\gamma}_*(x_{j0}-x_j)\Gamma_j$$

(v) $$\hat{\Gamma}_{j0}{}^{T}\bar{\gamma}(x_{j0}-x_{j0})\hat{\Gamma}_{j0} = \Gamma_{j0} = \Gamma_{j0}{}^{T}\bar{\gamma}^*_*(x_{j0}-x_{j0})\Gamma_{j0}$$

(vi) $$\Gamma_j{}^{T}\bar{\gamma}(x_j-x_{j0})\hat{\Gamma}_{j0} = \Gamma_j{}^{T}\bar{\gamma}*(x_j-x_{j0})\Gamma_{j0}$$

Then the estimation variance can be written in terms of the modified γ's instead of in terms of the modified Γ_j's. All arbitrary entries in the Γ_j's will be set equal to zero.

NUMERICAL METHODS

Because of the size, of the system of equations to be solved in Co-Kriging, makes it formidable for more than a few variables and a few sample locations. We describe first a method for reducing the system and then an iterative method that requires less core memory to solve the system.

Utilizing the matrix form given by (35) let

$$K = \begin{bmatrix} \bar{\gamma}(x_1-x_1) & \cdots & \bar{\gamma}(x_1-x_n) \\ \vdots & & \vdots \\ \bar{\gamma}(x_n-x_1) & \cdots & \bar{\gamma}(x_n-x_n) \end{bmatrix} \qquad (38)$$

$$E = [I\cdots I]^T \qquad (39)$$

$$\bar{\Gamma} = [\Gamma_1\cdots\Gamma_n]^T \qquad (40)$$

$$K_0 = [\bar{\gamma}(V,x_1)\cdots\bar{\gamma}(V,x_n)]^T \qquad (41)$$

Then (35) may be written as

$$\begin{bmatrix} K & E \\ E^T & 0 \end{bmatrix} \begin{bmatrix} \bar{\Gamma} \\ \bar{\mu} \end{bmatrix} = \begin{bmatrix} K_0 \\ I \end{bmatrix} \tag{42}$$

i.e. $K\bar{\Gamma} + E\bar{\mu} = K_0 \quad E^T\bar{\Gamma} = I$ (43)

If $KU = K_0$, $KV = E$ then (44) becomes

$$K(\bar{\Gamma} + V\bar{\mu}) = KU \quad \text{or} \quad \bar{\Gamma} = U - V\bar{\mu} \tag{44}$$

If moreover

$$(E^TV)W = EU - I \quad \text{then} \quad \bar{\Gamma} = U - VW \tag{45}$$

This not only reduces the size of the system but also avoids the possibility that (43) is ill-conditioned.

Kacmancz and Tanabe (6) have given an iterative method for solving systems of equations even when the system is singular: the author (5) has extended this to systems such as (35) or (43). For any two n x 1 matrices X,Y whose entries are m x m matrices let

$$\langle X,Y\rangle = Y^TX \quad \text{and} \quad (X,Y) = \mathrm{Tr}\,\langle X,Y\rangle \tag{46}$$

(X,Y) is an inner-product and $\| X \| = \sqrt{(X,X)}$ is a norm. For a system

$$AX = B \tag{47}$$

A a p x n matrix, X n x 1, B p x 1 whose entries are m x m matrices let $\bar{A}_i$ be the i^{th} column in A^T. Assume $\| \bar{A}_i \| > 0$ for all i. Let f_i be defined as follows

$$f_i(X) = X - \frac{1}{\alpha_i}[\bar{A}_i \langle X, \bar{A}_i\rangle - A_iB_i] \tag{48}$$

and $F(X) = f_1 \circ f_2 \circ \ldots \circ f_p(X)$. If X_0 is any initial element and

$$X_{i+1} = F(X_i) \tag{49}$$

then $X_0, X_1, \ldots, X_q, \ldots$ is a sequence converging to the solution of $AX = B$.

By writing

$$B = [B_1 \cdots B_p]$$

the original system $AX = B$ can be written

$$\langle X, \bar{A}_i \rangle = B_i \quad ; i=1, \ldots, p \tag{50}$$

If X is the solution of the system then $f_i(X)$ is the projection of X onto the hyperplane given by (51) even if A is invertible this method is useful since only one row of A need be in core at one time. Gaussian reduction would require all of A in core at one time. This projection method can be used for the under sampled case and also in conjunction with the reduction technique given earlier in this section.

A computer program has been written at the University of Arizona by J. Carr for Co-Kriging and which utilizes the projection algorithm for solving the system.

References

(1) Journel, A. G. 1977 "Geostatistique Minier" Centre de Geostatistique Fontainebleau, 737 p.

(2) Myers, D. E. 1981 "Joint Estimation of random functions: the matrix form" Research Report, Department of Mathematics, University of Arizona 32 p.

(3) Myers, D. E. 1982 "Matrix Formulation of Co-Kriging" Mathematical Geology, Vol. 14, No. 3, 249-257.

(4) Myers, D. E. 1983 "Estimation of Linear Combinations and Co-Kriging" to appear in Mathematical Geology Vol. 15, No. 5

(5) Myers, D. E. 1983 "On solving large scale linear systems", submitted for publication.

(6) Tanabe, K. 1971 "Projection method for solving a singular system of linear equations and its application", Numerische Math. 17, 203-214