Some Aspects of Multivariate Analysis

Donald E. Myers
Department of Mathematics
University of Arizona
Tucson, AZ 85721 U.S.A.

ABSTRACT. Classical multivariate methods for analyzing data tableaux do not explicitly incorporate spatial correlation. In contrast geostatistical techniques such as kriging, dual kriging and cokriging are primarily estimation methods. The multivariate formations of dual kriging and disjunctive kriging, dual cokriging and co-disjunctive kriging, provide methods which bridge the gap between the two approaches. Basic properties of both are presented.

INTRODUCTION

The importance of multivariate techniques has been recognized for a number of years. In particular, Cluster Analysis, Principal Components and Discriminant Analysis have had widespread use. However, these tools which had their origin in other disciplines other than the geosciences do not fully respect some characteristics of multivariate data sets arising out of problems in the geosciences, mining, environmental monitoring, hydrology and remote sensing. In particular the models underlying these do not specifically incorporate spatial correlation. The dichotomy between these techniques and estimation techniques such as kriging can be seen in the context of two perspectives of a data tableau. Consider

$$\begin{matrix} Z_1(x_1), \ldots, Z_m(x_1) \\ \vdots \qquad\qquad \vdots \\ Z_1(x_n), \ldots, Z_m(x_n) \end{matrix}$$

(F. Chung et al. (eds.), Quantitative Analysis of Mineral and Energy Resources, 669–687

In general multivariate techniques interpret the tableau as a (random?) sample of size n from a vector valued random variable whereas geostatistics a la Matheron would model the tableau as a non-random sample from one realization of a vector-valued random function where the spatial correlation is of principal interest. The above multivariate techniques do not require position coordinates and do not reflect the support of the sample. In general they are not estimation techniques but rather seek to identify underlying phenomena or discern patterns. The various forms of kriging have been developed and used for estimation almost exclusively. There have been a few instances where a multivariate technique such as principal components has been applied in conjunction with kriging, for example Borgman and Frahme (1976), Myers and Carr (1984), Myers (1984), Davis and Greenes (1983). In each of these the objective was to simplify or to avoid cokriging. Additional interesting applications are given in Bopp and Biggs (1981) or Hopke et al (1976) neither of whom used geostatistics but simply contoured the derived factors to identify patterns. Although pattern recognition is a tool much desired it is as yet not fully developed.

The results presented here incorporate some aspects of classical multivariate methods as well as incorporating spatial correlation and are intended to stimulate further thought.

I. POSSIBLE GENERALIZATION

A. Consider a data tableau with an infinite number of rows, i.e. one for each point in a domain of interest, i.e., each column is a function, not just a finite number of points and which is written simply as

$$Z_1, \ldots, Z_m$$

For a finite number of points, i.e. a finite number of lines, a uniform distribution is assumed, that is each line is equally weighted. For the infinite case a probability distribution must be assumed for each column. The analogy to choosing a vector which carries the largest amount of variance is to choose a random variable such that the sum of the squares of the projection is maximized. In

the finite version this is given as the inner product, for random variables it is the conditional expectation. Let the conditional expectation $E(Z_i/U)$ be written in operator form $D_U Z_i$ which easily extends to the vector $Z = [Z_1, \ldots, Z_m]$. Writing the inner product of two random variables V, W as $\langle V, W \rangle = E(VW)$, the problem is to maximize

$$\mathrm{Tr} \langle D_U Z, D_U Z \rangle \tag{1}$$

which is a positive quadratic form. However, unlike the finite dimensional case where the bivariate densities are implied by the data, in this case the operator D_U is characterized by bivariate densities that are unknown. One possible resolution of this problem is to assume that each pair Z_i, U is bivariate Gaussian and then it is sufficient to know the correlation for each pair. This is essentially the disjunctive kriging approach but recast in multivariate form, and will be discussed in Section IV.

B. Because the above formulation does not lead to a numerical solution without the use of strong distributional assumptions it is useful to re-consider the problem from another perspective. Principal components produces a representation of the data in the following form

$$Z = FA \tag{2}$$

where $F^T F$ is diagonal and the columns of F are ordered by percent of variance explained. One of the objectives of this transformation is to remove inter-variable correlation however this method only incorporates the inter-variable correlation at the same location. It would be more appropriate to remove spatial inter-variable correlation.

This suggests the following: if

$$\bar{Z}(x) = [Z_1(x), \ldots, Z_m(x)] \tag{3}$$

is a vector of second order stationary functions, find

$\bar{Y}(x) = [Y_1(x), \ldots, Y_p(x)]$ and C such that

$$\bar{Z}(x) = \bar{Y}(x)C \tag{4}$$

and the components of $\bar{Y}(x)$ are uncorrelated. Written in terms of the variograms and cross-variograms this becomes

$$\bar{\gamma}_Z(h) = C^T \bar{\gamma}_Y(h)C \tag{5}$$

and $\bar{\gamma}_Y(h)$ is to be a diagonal matrix. Here $\bar{\gamma}_Z(h)$, $\bar{\gamma}_Y(h)$ denote

$$\frac{1}{2} E([\bar{Z}(x + h) - \bar{Z}(x)]^T[\bar{Z}(x + h) - \bar{Z}(x)]) \tag{6}$$

and

$$\frac{1}{2} E([\bar{Y}(x + h) - \bar{Y}(x))]^T[\bar{Y}(x + h) - \bar{Y}(x)]) \tag{7}$$

$\bar{\gamma}_Z(h)$ corresponds to a correlation matrix. The representation

$$C^T\bar{\gamma}_Y(h)C \tag{8}$$

is analogous to the diagonalization of the correlation matrix as in Principal Components Analysis.

If all of the variograms of the components of $\bar{Z}(x)$ have sills then the total inertia of $\bar{Z}(x)$ is the sum of the sills, i.e.

$$\mathrm{Tr}\ \bar{\gamma}_Z(\infty) = \Sigma(\Sigma\ C_{jk}^2)\gamma_Y(\infty)_k \tag{9}$$

The columns of C can be normalized by requiring

$$\sum_{j=1}^{m} C_{jk}^2 = 1 \tag{10}$$

then the component of $\bar{Y}(x)$ with the largest percent of variance is the component of $\bar{Y}(x)$ whose variogram has the largest sill.

When the objective is pattern recognition as opposed to co-kriging, then the sample variogram matrix may substitute for the true variogram matrix. After the sample variogram matrix is computed for

various lags and diagonalized to find the eigenvalues, the sills for each variogram are estimated by the eigenvalues. When C is non-singular the equation

$$\bar{Z}(x) = \bar{Y}(x)C \tag{11}$$

can be solved as

$$\bar{Y}(x) = C^{-1}\,\bar{Z}(x) \tag{12}$$

and hence at least at data locations the values of $\bar{Y}(x)$ are determined. This would permit separate kriging of the components of $\bar{Y}(x)$ with subsequent reconstruction of kriged values of $\bar{Z}(x)$. This decomposition would also allow simulation of $\bar{Z}(x)$ by simulation of the components of $\bar{Y}(x)$.

The model formulation given by (11) is similar to that used by Wackernagel (1985) but the method of achieving it is different. In Wackernagel the procedure is to model the variogram (or covariances) of the components of $\bar{Z}(x)$ by nested models. It is assumed that the components of these nested models represent the components of $\bar{Y}(x)$ then the cross-variograms are modeled as linear combinations of the variogram models. In particular it is often assumed that the principal difference in the variogram models is reflected in the ranges. The approach proposed above allows a more general model for $\bar{\gamma}_Z(h)$. Not all such models can be diagonalized whereas the construction in Wackernagel ensures a diagonalizable model and hence is more restrictive in particular it assumes all variograms and cross variograms are proportional to the same model.

II. DUALITY - AN ALTERNATIVE

Before introducing a multivariate version it is useful to recall the formulation of the univariate kriging estimation in dual form. Not only is this form analogous to the use of splines for interpolation but it also provides a method for removing or reducing the spatial correlation.

Recall the equations determining the ordinary Kriging

estimator. Let $x_1, \ldots, x_n$ be data locations with data $Z(x_1), \ldots, Z(x_n)$. Then the kriging equations are written as

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{13}$$

where

$$\begin{aligned} \sum_{i=1}^{n} \lambda_i \gamma(x_i - x_j) + \mu &= \gamma(x_0 - x_j) \\ \Sigma \lambda_i &= 1 \end{aligned} \tag{14}$$

In matrix form we have

$$\begin{aligned} K\lambda + F\mu &= K_0 \\ F^T \lambda &= 1 \end{aligned} \tag{15}$$

which we can write as

$$\begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix} \begin{bmatrix} \bar{\lambda} \\ \mu \end{bmatrix} = \begin{bmatrix} K_0 \\ 1 \end{bmatrix} \tag{16}$$

where

$$K = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \gamma_{n1} & \cdots & \gamma_{nn} \end{bmatrix}, \ \gamma_{ij} = \gamma(x_i - x_j)$$

$$F = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \ \bar{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \ K_0 = \begin{bmatrix} \gamma_{01} \\ \vdots \\ \gamma_{0n} \end{bmatrix} \tag{17}$$

$$Z^*(x_0) = [Z(x_1), \ldots, Z(x_n), 0] \begin{vmatrix} \bar{\lambda} \\ \mu \end{vmatrix} \tag{18}$$

$$= [K_0^T \ 1] \begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{Z}^T \\ 0 \end{bmatrix} \tag{19}$$

(x_1),

Let
$$\begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix} \begin{bmatrix} \hat{Y} \\ b \end{bmatrix} = \begin{bmatrix} \hat{Z}^T \\ 0 \end{bmatrix} \tag{20}$$

(13)

If unique neighborhood kriging is used then $[\hat{Y}\ b]^T$ need only be computed once since all the information about the point to be estimated is contained in $[K_0^T\ 1]$. $\hat{Y}$ is a data vector "dual" to

(14) $\hat{Z}$. First introduced by Matheron, Dubrule (1983) showed that this formulation allowed for the use of a unique neighborhood in cross-validation and more recently Dubrule and Kostov (1986) have shown how it facilitates the incorporation of inequality constraints. Royer and Vieira (1984), Galli, Murillo and Thomann (1984) have suggested

(15) interpretations of $\hat{Y}$.

$\hat{Y}$, b have some properties which have perhaps not been noted. For example

(16)
$$\hat{Y}F = 0 \quad \text{and} \quad E(\hat{Y}) = [0,\ 0, \cdots 0,\ 0]$$

So that $\hat{Y}$ is centered in two different senses.

If $E(Z(x)) = m$ then $E(\hat{Z}) = mF^T$ and $E(b) = m$ hence b is an unbiased estimator of m. By writing

$$\begin{aligned} U &= K^{-1}K, \quad V = K^{-1}F \\ W &= K^{-1}\hat{Z}^T \end{aligned} \tag{21}$$

(17)

then we obtain

$$\begin{aligned} \text{Var}(b) &= \sigma^2 - (F^T W)^{-1} \\ &= \sigma^2 - (\hat{Z}\ V)^{-1} \end{aligned} \tag{22}$$

(18)

In a crude sense $\hat{Z}$ is replaced by new vector in which each entry is a linear combination but the components of $\hat{Y}$ are less correlated.

(19) If we compute

$$E[\hat{Z}^T Z] = (\sigma^2 + m^2)FF^T - K \tag{23}$$

then

$$E[\hat{Y}^T\hat{Y}] = V(F^TV)^{-1}V^T - K^{-1} \tag{24}$$

which is analogous to a correlation matrix.

III. CO-DUALITY

Although dual kriging is not quite a multivariate technique it suggests an extension of Cokriging that can properly be called multivariate.

Recall from Myers (1982, 1983, 1984) the general form of cokriging. As above $x_1, \ldots, x_n$ denote sample locations, $\bar{Z}(x_1), \ldots, \bar{Z}(x_n)$ are the data vectors where $\bar{Z}(x) = [Z_1(x), \ldots, Z_m(x)]$. Then the cokriging estimator is given by

$$\bar{Z}^*(x_0) = \sum_{i=1}^{n} \bar{Z}(x_i)\Gamma_i \tag{25}$$

where

$$\sum_{j=1}^{n} \bar{\gamma}(x_i - x_j)\Gamma_j + \bar{\mu} = \bar{\gamma}(x_0 - x_i)$$

$$\sum_{j=1}^{n} \Gamma_j = I \tag{26}$$

and

$$\bar{\gamma}(h) = \frac{1}{2}([\bar{Z}(x + h) - \bar{Z}(x)]^T[\bar{Z}(x + h) - \bar{Z}(x)]) \tag{27}$$

Unlike the kriging estimator $Z^*(x_0)$, $\bar{Z}^*(x_0)$ is not a scalar and hence is not equal to it's transpose but since

$$[[\bar{Z}^*(x_0)]^T]^T = \bar{Z}^*(x_0) \tag{28}$$

we shall work with $[\bar{Z}^*(x_0)]^T$ using notation analogous to that above for dual kriging

(24)

$$K = \begin{bmatrix} \bar{\gamma}(x_1-x_1) \cdots \bar{\gamma}(x_1-x_n) \\ \vdots \qquad\qquad \vdots \\ \bar{\gamma}(x_n-x_1) \cdots \bar{\gamma}(x_n-x_n) \end{bmatrix} \tag{29}$$

$$F = \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix}, \quad \text{each } I \text{ an } m \times m \text{ identity matrix}$$

e it
multi-

$$\bar{\Gamma} = \begin{bmatrix} \Gamma_1 \\ \vdots \\ \Gamma_n \end{bmatrix}, \quad K_0 = \begin{bmatrix} \bar{\gamma}(x_0-x_1) \\ \vdots \\ \bar{\gamma}(x_0-x_n) \end{bmatrix} \tag{30}$$

then the cokriging system becomes

'en by

(25)

$$\begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix} \begin{bmatrix} \bar{\Gamma} \\ \bar{\mu} \end{bmatrix} = \begin{bmatrix} K_0 \\ I \end{bmatrix} \tag{31}$$

We have then

(26)

$$[Z^*(x_0)]^T = [K_0^T \; I] \begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\bar{Z}}^T \\ 0 \end{bmatrix} \tag{32}$$

where

27)

$$\hat{\bar{Z}} = [\bar{Z}(x_1), \ldots, \bar{Z}(x_n)] \tag{33}$$

d

Let $\hat{\bar{Y}} = [\bar{Y}(x_1), \ldots, \bar{Y}(x_n)]$

with

28)

$$\begin{bmatrix} K & F \\ F^T & 0 \end{bmatrix} \cdot \begin{bmatrix} \hat{\bar{Y}}^T \\ B \end{bmatrix} = \begin{bmatrix} \hat{\bar{Z}}^T \\ 0 \end{bmatrix} \tag{34}$$

ove

Then $\hat{\bar{Y}}$ is the "dual" to $\hat{\bar{Z}}$ in that it is a vector obtained from $\hat{\bar{Z}}$ by incorporating the spatial and intervariace correlation

quantified by the variogram matrix. We see immediately that

$$\hat{\bar{Y}}^T = [I - K^{-1}F(F^TK^{-1}F)F^T]K^{-1}\hat{\bar{Z}}^T \tag{35}$$

and

$$B = (F^TK^{-1}F)^{-1}F^TK^{-1}\hat{\bar{Z}}^T \tag{36}$$

or alternatively

$$\hat{\bar{Z}} = K[I - K^{-1}F(F^TK^{-1})F^T]^{-1}\hat{\bar{Y}} \tag{37}$$

The original data tableau is then given as a linear combination of the entries in the new data tableau. Moreover, B is an estimator of $E[\bar{Z}(x)]^T$ when $\bar{Z}(x)$ is second order stationary and

$$\text{Trace } E[B - E(\hat{\bar{Z}}^T)][B - E(\bar{Z}^T)] = \text{Tr}[C(0) - K^{-1}F(F^TK^{-1}F)^{-1}] \tag{38}$$

is the sum of the variances of the estimation errors.

IV BIVARIATE GAUSSIAN MODEL

The first model proposed to generalize principal components would require minimizing the quadratic form given by equation (1) which is determined by the conditional expectation operator. In turn this would require knowledge of (at least) bivariate densities. Matheron (1976) extended the kriging estimator from a linear to a non-linear form by utilizing a bivariate Gaussian density function. Specifically the disjunctive kriging estimator is written in the form

$$Z^*_{DK}(x_0) = \sum_{j=1}^{n} f_j(Z(x_j))] \tag{39}$$

The f_j's are obtained by Hermite polynomial expansions which in turn are the eigenfunctions of the conditional expectation operator. By utilizing the general cokriging formulation given in Eq. (26), the DK estimator can be fully generalized.

A. Preliminaries

Let L^2_{ij} denote the space of functions f, such that $f(Z_i(x_j))$ is square integrable, i.e. $E[f(Z_i(x_j))]^2 < \infty$. Then form

the sum $L^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} L_{ij}^2$, which is not a direct sum since the constant functions are in each L_{ij}^2. If $\| \quad \|_{ij}$ is the norm in L_{ij}^2 then $[\sum_{i=1}^{m} \sum_{j=1}^{n} \| \quad \|_{ij}^2]^{1/2}$ is a norm on L^2. The non-linear estimator for $Z_j(x_0)$ is then the projection of $Z_j(x_0)$ onto the linear subspace of L^2 generated by the elements of the L_{ik}^2, that is,

$$Z_j^*(x_0) = \sum_{k=1}^{n} \sum_{i=1}^{m} f_{ij}^k(Z_i(x_k)) \tag{40}$$

where $f_{ij}^k \in L_{ik}^2$.

B. The Projections

Since $Z_j^*(x_0)$ is a projection, the errors $Z_j(x_0) - Z_j^*(x_0)$ are orthogonal to all the $f_i^k(Z_i(x_k))$ and hence we have

$$E[Z_j(x_0) f_{ij}^k(Z_i(x_k))] = E[Z_j^*(x_0) f_{ij}^k(Z_i(x_k))] \tag{41}$$

for all i, j, k. Since $Z_j^*(x_0)$ is a linear combination, each of these expectations only requires a bivariate density. However the functions f_{ij}^k are in general unknown and the problem is insoluable in this generality.

C. Transformations

As with single variable disjunctive kriging we assume that each random function may be transformed to one with a Gaussian distribution. More specifically we assume that for each Z_j there is a random function Y_j which is Gaussian and a function ϕ_j such that $Z_j(x) = \phi_j(Y_j(x))$ for each j, x. Moreover, we assume that for each i, j and each pair x, y

$$Y_i(x), \quad Y_j(y)$$

are bivariate Gaussian. If $i = j$ then this is the usual assumption for disjunctive kriging.

D. Hermitian Expansions

If $Y(x)$ is a Gaussian random function and f is a function such that $E[f(Y(x))]^2 < \infty$ then f may be represented by an expansion in terms of Hermite polynomials. In particular if f_{ij}^k is a function in L_{ik}^2 then

$$f_{ij}^k(Y(x)) = \sum_{p=0}^{\infty} f_{ij}^{kp} H_p(Y(x)) \tag{42}$$

where

$$f_{ij}^{kp} = E[f_{ij}^k(Y(x))H_p(Y(x))] \tag{43}$$

This provides a method for representing the unknown functions f_{ij}^k. Moreover the functions ϕ_j can be represented as

$$\phi_j(Y(x)) = \sum_{p=0}^{\infty} C_j^p H_p(Y(x)) \tag{44}$$

E. Bivariate Densities

Under the assumptions that each pair $Y_i(x)$, $Y_j(y)$ is bivariate Gaussian, the joint densities can be written in the form

$$g_{ij,xy}(u,v) = \sum_{p=0}^{\infty} (\rho_{ij}^{xy})^p H_p(u)H_p(v)g(u)g(v) \tag{45}$$

where $g(x)$, $y(y)$ are standard Gaussian densities, ρ_{ij}^{xy} is the correlation coefficient of $Y_i(x)$, $Y_j(y)$. As with univariate disjunctive kriging we could also consider more general bivariate densities such as the Hermitean models.

F. The Orthogonality Conditions

Combining Eq. (41), (42) we will have

$$E[Z_j(x_0)|Z_\ell(x_r)] = \sum_{i=1}^{m} \sum_{k=1}^{n} [f_{ij}^k(Z_i(x_k))|Z_\ell(x_r)] \tag{46}$$

for all j, ℓ, r

Using the transformations $Z_j(x) = \phi_j(Y_j(x))$ we obtain

is

$$E[\phi_j(Y_j(x_0))|Y_\ell(x_r)] = \sum_{i=1}^{m} \sum_{k=1}^{n} E[f_{ij}^k(\phi_i(Y_i(x_k)))|Y_\ell(x_r)] \qquad (47)$$

Since both the ϕ_j's and the f_{ij}^k's are unknown we replace $f_{ij}^k(\phi_j(Y_j(x_k)))$ by simply $f_{ij}^k(Y_j(x_k))$ where of course the f_{ij}^k's are not the original ones and then Equation (40) becomes

$$Z_j^*(x_0) = \sum_{k=1}^{n} \sum_{i=1}^{m} f_{ij}^k(Y_i(x_k)) \qquad (48)$$

or

$$Z_j^*(x_0) = \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{p=0}^{\infty} f_{ij}^{kp} H_p(Y_i(x_k)) \qquad (49)$$

If we write $\bar{Z}^*(x_0) = [Z_1^*(x_0), \ldots, Z_m^*(x_0)]$ and

$$\bar{H}_p(x_k) = \begin{bmatrix} H_p(Y_1(x_k)) \\ \vdots \\ H_p(Y_m(x_k)) \end{bmatrix}^T \qquad (50)$$

$$F_k^p = \begin{bmatrix} f_{11}^{kp} & \cdots & f_{1m}^{kp} \\ \vdots & & \\ f_{m1}^{kp} & \cdots & f_{mm}^{kp} \end{bmatrix} \qquad (51)$$

then we have

$$\bar{Z}^*(x_0) = \sum_{p=0}^{\infty} \sum_{k=1}^{n} H_p(x_k) F_k^p \qquad (52)$$

where

$$\bar{Z}(x_0) = [\phi_1(Y_1(x_0)), \ldots, \phi_m(Y_m(x_0))]$$

$$= \sum_{p=0}^{\infty} \bar{H}_p(x_0) \begin{bmatrix} C_1^p & & 0 \\ & \ddots & \\ 0 & & C_m^p \end{bmatrix} \qquad (53)$$

Now if we let F_k^p be written as

$$\begin{bmatrix} b_{11}^{kp} \cdots b_{1m}^{kp} \\ \vdots \\ b_{m1}^{kp} \cdots b_{mm}^{kp} \end{bmatrix} \begin{bmatrix} C_1^p & & 0 \\ & \ddots & \\ 0 & & C_m^p \end{bmatrix} = B_k^p \, C^p \qquad (54)$$

then

$$\bar{Z}(x_0) - \bar{Z}^*(x_0)$$

$$= \sum_{p=0}^{\infty} [\bar{H}_p(x_0) - \sum_{k=1}^{n} \bar{H}_p(x_k) B_k^p] \bar{C}^p \qquad (55)$$

and the estimation variance in L^2 is given by

$$\text{Trace} \; E[\bar{Z}(x_0) - \bar{Z}^*(x_0)]^T [\bar{Z}(x_0) - \bar{Z}^*(x_0)]$$

$$= \sum_{p=0}^{\infty} \text{Trace} \, (\bar{C}^p)^T \, W_p \, C^p \qquad (56)$$

$$W_p = E(\bar{H}_p(x_0))^T (\bar{H}_p(x_0))$$

$$- \sum_{k=1}^{n} (B_k^p)^T E[(\bar{H}_p(x_k))^T \bar{H}_p(x_0]$$

$$- \sum_{k=1}^{n} E[(\bar{H}_p(x_0))^T (\bar{H}_p(x_k))] B_k^p$$

$$+ \sum_{\ell=1}^{n} \sum_{k=1}^{n} (B_\ell^p)^T E[(\bar{H}_p(x_\ell))^T \bar{H}_p(x_k)] B_k^p \qquad (57)$$

(53) Since the C^p's are determined only by the transformations ϕ_i, to minimize the estimation variance it is sufficient to minimize, for each p separately,

$$\mathrm{Tr}(C^p)^T W_p \bar{C}^p \qquad (58)$$

Moreover, for each p, it is sufficient to consider

$$C^p = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

(54)

But minimizing

$$\mathrm{Tr}\ W_p$$

is exactly the problem of cokriging $\bar{H}_p(x_0)$ using the data $\bar{H}_p(x_1), \ldots, \bar{H}_p(x_n)$. Hence the matrices $B_1^p, \ldots, B_n^p$ are the solutions of the (simple) cokriging systems

(55)

$$\sum_{k=1}^{n} G_{jk}^p B_k^p = G_{0j}^p \qquad (59)$$

$$G_{jk}^p = E[(\bar{H}_p(x_j))^T \bar{H}_p(x_k)] \qquad (60)$$

(56) as formulated in Myers (1982)

For the bivariate Gaussian case the entries in G_{jk}^p are

$$(\rho_{jk}^{i\ell})^p$$

where $\rho_{jk}^{i\ell}$ = correlation of $Y_i(x_j)$, $Y_\ell(x_k)$

G. The Undersampled Case

If $\bar{Z}$ is under-sampled, i.e. if for i and some k, $Z_i(x_k)$ is missing then in equation (10) for all j, $f_{ij}^k \equiv 0$ and hence for all

p, $f_{ij}^{kp} = 0$. In the context of equation (27) this is exactly the under-sampled formulation of the cokriging problem as given in Myers (1984) and in Carr, Myers and Glass (1985) hence the algorithm given there still suffices.

V. DATA or MODEL DRIVEN

Principal Components is essentially a data driven technique, that is, no parameters need be separately fitted nor are there assumptions such as normality or stationarity. Some of the interpretations of these components and their eigenvalues follow only from the geometric formulation without the use of external models. Cokriging, dual kriging and cokriging in contrast require fitting variograms and cross-variograms, these fitting techniques are not on completely solid ground statistically and also require implicit modelling of the phenomena by random functions and the varograms must satisfy certain conditions. The extension from linear to non-linear techniques require very specific statistical models and assumptions these additional conditions are often untestable. It would appear, that multi-variate techniques that truly incorporate spatial correlation are either not possible or at least not as yet developed.

References

Armstrong, M., (1984) 'New Types of Disjunctive Kriging', Part I, N-872 Fontainebleau

Armstrong, M. and Matheron, G., (1985) 'Isofactorial models for granulo-densimetric data', N-944 Fontainebleau

Armstrong, M. and Matheron, G., (1985) 'New types of Disjunctive Kriging', Part II, N-943 Fontainebleau

Bopp, F. III, and Biggs, R. Metals in Estuarine Sediments. Factor Analysis and its Environmental Significance, Science v 214, p 441-443

Borgman L. and Frahme, R. B., (1976), 'Multivariate Properties of Bentonik in Northeastern Wyoming', in Advanced Geostatistics in the Mining Industry, Guarascio, M. et al (eds) D. Reidel Publishing, Dordrecht

Carr, J., Myers, D. E. and Glass, C., (1985) 'Co-Kriging: A Computer Program', Computers and Geosciences, 11, No 2, p 111.

Carr, J.
Co-C
p 67

Davis, B
distr
data'
Dubrule,
and k

Dubrule,
into
Geolo

Galli, A.
prope
for M
Reide

Gupta, A.
distri

Hopke, Ph.
D.J.,
J. Rad

Jackson, M
by dis
Comput
(O'Nei
of Min

Kim, Y. C.
geostat
DOE, Gr

Lantuejoul,
determi

Marechal, A
N-384 F

Marechal, A
gaussie

Marechal, A.
methods
Advance
Pub. Co.

Matheron, G.
Fontaine

Carr, J. and Myers, D. E., 'COSIM: (1985) A FORTRAN IV Program for Co-Conditional Simulation', Computers and Geosciences, 11, No. 6 p 675.

Davis, B. and Greenes, K., (1983), 'Estimation using spatially distributed multi-variate Data; An example with coal quality data', Math. Geology, 15, p. 287

Dubrule, O., (1983), 'Two methods with different objectives: splines and kriging', Math. Geology, 15, p 245

Dubrule, O. and Kostov, C., (1986) 'An interpolation method taking into account inequality constraints I. Methodology', Math. Geology, 18, No. 1, p 33

Galli, A., Murillo, E. and Thomann, J., (1984), 'Dual Kriging. Its properties and its uses in direct contouring', in Geostatistics for Natural Resource Characterization, G. Verley et al (eds.), D. Reidel Publishing, Dordrecht

Gupta, A. K., (1983) 'On the expansion of the bi-variate gamma distribution', Tech. Report, Bowling Green St. Univ.

Hopke, Ph., Ruppert, D.F., elute, P. R., Metzger, W.J. and Crowley, D.J., (176) 'Geochemical Profile of Chautauqua Lake Sediments,' J. Radioanalytical Chem. 29, p. 73.

Jackson, M. and Marechal, A., (1979) 'Recoverable reserves estimator by disjunctive kriging; a case study', in 16th Application of Computers and Operations Research in the Mineral Industry (O'Neil, T. J. ed.) Society of Mining Eng. of American Institute of Mining and Mettallurgy, N.Y. 240-249

Kim, Y. C., Myers, D. E. and Knudsen, H. P., (1977) 'Advanced geostatistics in ore reserve estimation and mine planning', U. S. DOE, Grand Junction 237 p

Lantuejoul, Ch. and Rivoirard, J., (1984) 'Une methode de determination d'anamorphose' N-916, Fontainebleau

Marechal, A., (1974) 'Generalities sur les fonctions de transfert', N-384 Fontainebleau

Marechal, A., (1975) 'Analyse numerique des anamorphoses gaussiennes', N-418 Fontainebleau

Marechal, A., (1976) 'The practice of transfer functions: Numerical methods and their applications'. M. Guarascio et al (eds) Advanced Geostatistics in the Mining Industry 253-206, D. Reidel Pub. Co., Dordrecht-Holland

Matheron, G., (1973) 'Le Krigeage Disjonctif', N-360/N-432 Fontainebleau

Matheron, G., (1974) 'Les fonctions de transfert des petits panneaux' N-395, Fontainebleau

Matheron, G., (1975) 'Modeles isofactoriels discrets et modele de Walsh', N-449, Fontainebleau

Matheron, G., (1975) 'Modeles isofactoriels pour l'effet zero' N-659, Fontainebleau

Matheron, G., (1976) 'A simple substitute for conditional expectation: The disjunctive kriging', in M. Guarascio et al (eds) Advanced Geostatistics in the Mining Industry, D. Reidel Pub. Co., Dordrecht-Holland

Matheron, G., (1976) 'Forcasting block grade distributions: The transfer functions' M. Guarascio et al (eds) Advanced Geostatistics in the Mining Industry, D. Reidel Pub. Co., Dordrecht-Holland 237-251

Matheron, G., (1977) 'Peut-on imposer des conditions d'universalite au krigeage disjonctif', N-539, Fontainebleau

Matheron, G., (1981) 'Quatre familles discrete', N.-703, Fontainebleau

Matheron, G., (1982) 'Pour une analyse krigeante des donnees reginalisees,' N-732, Fontainebleau

Myers, D. E., (1982), 'Co-Kriging: The Matrix Form', Math. Geology **14**, No 3 p. 249

Myers, D. E., (1983), 'Estimation of Linear Combinations and Co-Kriging', Math Geology, **15** No. 5 p 633

Myers, D. E., (1984), 'Co-Kriging: New Developments', in Geostatistics for Natural Resource Characterization, G. Verley et al (eds), D. Reidel Publishing Co, Dordrecht

Myers, D. E. and Carr, J., (1984), 'Co-kriging and Principal Components, Bentonite Data re-visited', Sciences de la Terre, **21**, p. 65

Myers, D. E., (1985) 'Co-kriging methods and Alternatives', in the Role of Data in Scientific Progress, P. Glaeser, (d), Elsevier Pub.

Rendu, Jean-Michel, (1980) 'Disjunctive Krigings: Comparison of Theory with actual results', Math. Geology, **12**, No. 4, p. 305

Rivoirard, J., (1985) 'Convergence des developpements en polynome d'Hermite', N-955, Fontainebleau

Royer, J. J. and Vieira, P.C., (1984) 'Dual Formalism of Kriging' in Geostatistics for Natural Resource Characterization, G. Verly et al (eds.), D. Reidel Publishing, Dordrecht

Wackernagel, H., (1985) "The inference of the linear model of the Coregionalization in the case of a geochemical data set,' Sciences del la Terre, No. 24, p. 81