CO-KRIGING: METHODS AND ALTERNATIVES

Donald E. Myers

Department of Mathematics, University of Arizona, Tuscon, Arizona 85721, U.S.A.

Co-Kriging is a linear estimation or contouring method for vector random functions which incorporates both spatial and intervariable correlation. A brief survey of Co-Kriging and types of applications will be given. Co-Kriging utilizes intervariable correlation for each pair of variables of interest and hence the number of variograms required increases rapidly, variogram and cross-variogram modelling will be discussed. Examples will be given for earthquake data from the Los Angeles basin and Bentonite deposits in Wyoming. Alternatives to Co-Kriging have been explored by several authors. Theoretical and empirical comparisons with Co-Kriging are given.

1. INTRODUCTION

The use of a linear estimator to interpolate or contour spatially distributed data is well-known for applications in mining, geology, soil science, hydrology, atmospheric sciences. Kriging, a regression method, is a minimum variance unbiased linear estimator (BLUE) and has received wide use for ore reserve estimation applications in mining. The theory and practice of Kriging is found in Journel and Huijbrechts (1978), Matheron (1971) as well as in numerous papers in MATHEMATICAL GEOLOGY, WATER RESOURCES RESEARCH, ECONOMIC GEOLOGY.

Kriging utilizes the spatial correlation, of the variable of interest with itself, to determine the weights in an optimal manner. There are at least two applications wherein intervariable as well as spatial correlation is relevant. When the data is sparse for the variable of interest but is readily available for a correlated variable it is desired to use the latter to improve the estimation of the former. A recent paper by Aboufirassi and Marino (1984) illustrates this "under sampled" problem. In other instances where there is not a single variable that is of primary interest, it may be relevant then to consider a linear combination of several variables. For each of these examples, joint estimation of several variables, using both spatial and intervariable correlation, is required. As was noted by Borgman and Frahme (1976), the number of variables may be large and this can create notational and computational problems. The objective of this paper is to review the theory of Co-Kriging (linear joint estimation), describe applications and consider alternative methods that have been used.

2. CO-KRIGING: A BRIEF REVIEW

Recall that in Kriging the sample data $Z(x_1)$, ... , $Z(x_n)$ is considered as taken from one realization of a random function $Z(x)$, x being a point in 1,2 or 3-space. The random function $Z(x)$ is assumed to have certain statistical properties, in particular a form of stationary. The spatial correlation of $Z(x)$ is represented by the variogram

$$(1) \quad \gamma(h) = \frac{1}{2} \, Var[Z(x + h) - Z(x)]$$

The linear estimator for $Z(x)$ at an unsampled location x_0 is given by

$$(2) \quad Z^*(x_0) = \sum_{i=1}^{n} \lambda_i \, Z(x_i)$$

where

$$(3) \quad \sum_{i=1}^{n} \lambda_i \gamma(x_i - x_j) + \mu = \gamma(x_0 - x_j);$$
$$j = 1,\ldots,n$$
$$\sum_{i=1}^{n} \lambda_i = 1$$

The linear system is obtained by minimizing the variance of the error of estimation. In order to determine the weights in a joint linear estimation either a scalar valued error or a scalar valued variance of the error must be identified since one can not minimize a vector quantity. For the "under sampled" problem and a single variable of primary interest, one may consider only the error of estimation of that variable. For the general problem the approach given in Myers (1982) is the relevant one and is described herein.

Let Z_1, ... , Z_m be the variables of interest for example the eleven characteristics for

Bentonite considered by Borgman and Frahme (1976), Myers and Carr (1984) or topsoil silt, subsoil silt, and subsoil sand as in McBratney and Webster (1983) or sulphur, ash, BTU as considered by Davis and Greenes (1983). Denote by

$$(4) \quad \overline{Z}(x) = [Z_1(x), \ldots, Z_m(x)]$$

The data for locations $x_1, \ldots, x_n$ is given by the n m-column matrices $\overline{Z}(x_1), \ldots, \overline{Z}(x_n)$. For an unsampled location x_0, the linear estimator for $\overline{Z}(x_0)$ is

$$(5) \quad \overline{Z}*(x_0) = \sum \overline{Z}(x_i)\Gamma_i$$

where $\Gamma_i = [\lambda_{jk}^i]$ is the weight matrix for location i, λ_{jk}^i is the weight assigned to $Z_j(x_i)$ in estimating $Z_k(x_0)$.

The Γ_i's are obtained as the solution to the system

$$(6) \quad \sum_{j=1}^{n} \overline{\gamma}(x_i - x_j)\Gamma_j + \overline{\mu} = \overline{\gamma}(x_0 - x_i)$$
$$\sum_{j=1}^{n} \Gamma_j = I$$

which is obtained by minimizing

$$(7) \quad \text{Trace } E [\overline{Z}(x_0) - \overline{Z}*(x_0)]^T [\overline{Z}(x_0) - \overline{Z}*(x_0)]$$
$$= \sum_{i=1}^{m} E[Z_i(x_0) - Z_i*(x_0)]^2$$

In (7) $\overline{\gamma}(h)$ denotes the matrix with entries

$$(8) \quad \gamma_{ij}(h) = \frac{1}{2} \text{Cov}(Z_i(x + h) - Z_i(x),$$
$$Z_i(x), Z_j(x+h) - Z_j(x))$$

and $\overline{\mu}$ is an m x m matrix of Lagrange multipliers.

In the application of Kriging a first step is to model the variogram from the data. For Co-Kriging a variogram must be modelled for each variable as well as a cross-variogram for each pair of variables.

As is shown in Myers and Carr (1984), modelling a large number of variograms and cross-variograms is quite feasible. Those results are described later. For n sample locations and m variables an (m + 1)n system of equations must be solved. Carr, Myers and Glass (1984) have given a computer program which shows that solution of the large systems is quite feasible.

3. MULTIVARIATE METHODS

When data are obtained for a large number of correlated variables it is frequently desirable to consider whether this information can be expressed by a smaller number of uncorrelated variables. Consider the data set

$\overline{Z}(x_1), \ldots, \overline{Z}(x_n)$. Each $\overline{Z}$ is an m-column matrix and might be interpreted as a point in m-space. For the collection of such points it is reasonable to ask about the center of gravity and the principal directions of this cloud of points as characteristics to describe the cloud. Principal Component Analysis (PCA) is the method for extracting the principal directions of the cloud. It characterizes angles by a scalar product which is equivalent to the Pearson Correlation Coefficient. After proper normalization of the data, the eigenvectors of the correlation matrix provide the desired uncorrelated new directions, the corresponding eigenvalues determine which directions are most important. Borgman and Frahme (1976) used PCA to reduce eleven characteristics for Bentonite to only five and an explained variance of 88%. Davis and Greenes (1983) used PCA for coal quality characteristics. PCA is not an estimation or contouring method but rather is primarily intended to reduce the number of variables under consideration.

PCA could be viewed as a method to obtain the representation by uncorrelated variables referred to in the previous section. The use of PCA does not ensure that all cross-variograms are zero. It should be noted that PCA does incorporate the position coordinates unless they are incorporated as additional variables. Davis and Greenes utilized the cross-variograms to show empirically that the coefficient functions were uncorrelated for that application.

Benzecri (1973) has given another form of principal component analysis. The data is viewed as an n x m array. If each entry is non-negative then the array may be considered as a contingency table, that is, if each entry is normalized by dividing by the sum of all the entries then each entry might be interpreted as a probability. It is this interpretation which motivates the use of a chi-square metric. Unlike PCA where rows and columns are treated in a non-symmetric manner, in Correspondence Analysis the factors are extracted simultaneously for rows and columns. Hill (1974) has provided a survey of these methods and their theoretical relationships. There is as yet not a large literature in English on the use of Correspondence Analysis in place of or in addition to PCA or Factor Analysis. None of the statistical packages such as BMDP, SSPS, IMSL incorporate CA. The numerical results for the use of CA given later in this paper were obtained with a version of the program by David, Dagbert and Beauchemim (1977) adopted for use on a CDC 6400.

EARTHQUAKE DATA

Following the event in Long Beach, California, building regulations provided for the installation of recording accelerometers in

buildings of more than one story especially in the Los Angeles region. While a great deal of attention has been given to possible methods for prediction of earthquakes, somewhat less has been devoted to predicting velocity and intensity of ground motion at different locations for events of a given magnitude. The design of buildings is dependent on such information. Carr, Myers and Glass (1984) have given an example of the use of Co-Kriging to simultaneously contour both of these variables. Co-Kriging is especially appropriate for multi-variate data, on a non-uniform grid, which is both spatially and inter-variable correlated. Table 1 tabulates the data for an event in the Los Angeles basin. Table 2 gives the results from Co-Kriging the data.

INPUT DATA

X-COORD	Y-COORD	VELOCITY	INTENSITY
132.360	91.170	10.200	7.000
133.210	102.280	15.600	7.000
71.850	182.890	1.000	5.000
76.490	173.440	3.800	5.000
141.490	94.500	8.200	7.000
167.240	71.710	2.300	6.000
119.210	92.611	5.100	7.000
108.810	163.430	11.700	6.000
169.670	58.920	3.900	5.000
189.820	130.080	2.000	5.000
132.550	63.370	6.100	5.000
220.260	93.390	1.500	5.000
0.000	135.640	1.700	5.000
97.860	141.200	6.200	6.000
143.470	152.310	7.600	6.000
72.370	44.470	3.500	6.000
248.490	57.810	2.300	5.000
44.410	98.950	3.200	6.000

TABLE 1

CO-KRIGING PROGRAM

SINGLE VARIABLE (VARIOGRAM) PARAMETERS

VARIABLE	NUGGET	SILL	RANGE
1	1.500	12.000	30.000
2	0.500	1.800	30.000

INTER-VARIABLE (CROSS-VARIOGRAM) PARAMETERS

NUGGET	SILL	RANGE
0.050	2.000	30.000

NORTH	WEST	VELOCITY	INTENSITY	VARIANCE
91.170	132.360	8.191	7.010	9.327
102.280	133.210	6.962	6.750	10.843
182.890	71.850	5.005	5.256	12.109
173.440	76.490	4.446	5.232	11.927
94.500	141.490	9.647	6.878	10.376
71.710	167.240	4.885	5.348	12.877
92.611	119.210	8.069	6.427	12.666

163.430	108.810	5.602	5.911	15.470
58.920	169.670	4.133	5.738	12.943
130.080	189.820	6.181	5.782	15.234
63.370	132.550	5.051	5.879	15.217
93.390	220.260	5.847	5.652	15.726
135.640	0.000	3.701	5.625	17.832
141.200	97.860	5.673	5.776	14.990
152.310	143.470	5.474	5.654	15.176
44.470	72.370	5.985	5.982	16.234
57.810	248.490	2.334	5.205	17.722
98.950	44.410	6.092	5.801	15.420

TABLE 2

BENTONITE DATA

In the Northern Blackhills District of Northeastern Wyoming and South Dakota, there are deposits of Bentonite. This clay is used for a number of purposes such as in mud for oil drilling. Unlike many metallic deposits, the economic value of the deposit is not characterized by an ore grade. Knetchel and Patterson (1956) reported measurements on as many as 29 characteristics at some locations. Borgman and Frahme (1976) considered only eleven characteristics all of which were observed at 43 locations in the Clay Spur beds. Although no specific linear combinations were given, Borgman and Frahme suggested that linear combination of these eleven is the appropriate proxy for ore grade. It is well-known that Kriging of linear composites is sub-optimal compared to forming linear combinations of Co-Kriged variables. Matheron (1979) has given a procedure for estimating the error resulting from Kriging linear composites. Although Borgman and Frahme did not report the results of Kriging the coefficients in the PCA representation, sample variograms were given and variograms modelled.

Myers and Carr (1984) have utilized Borgman and Frahme's results to provide a comparison with direct Co-Kriging. Variograms were modelled for the eleven variables as well as 55 cross-variograms. Although the results were not used in the Kriging, CA was also applied to the Bentonite data, three factors were sufficient to explain 88% of the variance as compared with five factors for 88% for PCA. It should be noted that Co-Kriging results in a computed variance of the error(s). The(se) variance(s) are determined solely by the variograms, cross-variograms, the sampling grid but not the data. The representations obtained by PCA or CA are data dependent and even if all factors are used the error variances are not determined in a comparable way. A complete discussion of the relationship of linear combinations of Co-Kriged values to Kriged linear combinations is given in Myers (1983). Finally it should be noted that the multi-variate methods such as PCA, CA incorporate intervariable correlation but not spatial intervariable correlation nor spatial correlation of individual variables. Table 4

gives a comparison of Co-Kriging with the
results using Kriging applied to the
Borgman/Frahme model. Since Kriging and Co-
Kriging are unbiased exact interpolators, cross-
validation can be used to quantify the
estimation reliability.

TABLE 3

Mean Square Error Estimation Results
(Kriging for 42 sample locations)

Variable	KRIGING Co-Kriging	Myers/ Carr	Borgman data	Borgman approx.
1	6.69	6.33	6.79	6.79
2	2.04	1.76	1.92	1.73
3	0.52	0.32	0.32	0.32
4	54.20	48.85	48.72	48.79
5	47.37	45.65	45.92	47.97
6	276.01	283.95	283.81	291.85
7	0.54	0.55	0.60	0.56
8	122.86	124.02	124.08	124.10
9	25.61	27.03	27.23	26.69
10	324.61	326.92	326.92	327.23
11	2037.27	2098.17	2098.17	2127.50

TABLE 4

REFERENCES

[1] Aboufirassi, M. and Mariño , M.A., "Kriging
     of water levels in the Souss Acquifer,
     Morocco", MATH GEOL. 15:4, (1983), 537.
[2] Benzecri, J.P., "L'Analyse
     des Données ",(Dunod, Paris, 1973).
[3] Borgman, L.E. and Frahme, R.B.,
     "Multivariate properties of Bentonite in
     Northeastern Wyoming", in Advanced
     Geostatistics in the Mining Industry,
     (Reidel Publishing, Amsterdam, 1976), 381.
[4] Carr, J., Myers, D.E. and Glass, C., "Co-
     Kriging: A computer program", to appear in
     COMPUTERS AND GEOSCIENCES, (1984).
[5] Davis, B. and Greenes, K, "Estimation using
     spatially distributed multi-variate data:
     An Example with Coal Quality", MATH GEOL.
     15, No. 2, (1983), 287.
[6] Davis, B. and Greenes, K, "Estimation using
     spatially distributed multivariate data:
     An Example with Coal Quality", MATH GEOL.
     15, No. 2, (1983), 287.
[7] François-Bongarcon, D., "La
     corégionalisation – le cokrigeage": C-86,
     Centre de Géostatistique, (Fontainebleau,
     France, 1981).
[8] Hill, M., "Correspondence Analysis: a
     neglected multivariate method", APPLIED
     STATISTICS 23, (1974), 340.
[9] Journel, A. and Huijbrechts, Ch, "Mining
     Geostatistics", (Academic Press, London,
     1978).
[10] Matheron, G., "Theory and Application of
     Regionalized Variables", Centre de
     Géostatistique , (Fontainebleau, France,
     1971).
[11] McBratney, A.B. and Webster, R., "Optimal
     interpolation and isorithmic mapping of
     soil properties, V. Co-regionalization and
     multiple sampling strategy", J. SOIL
     SCIENCE, 34, (1983), 137.
[12] Myers, D.E., "Co-Kriging: The Matrix
     Formulation", MATH GEOL., 14, No. 3,
     (1982), 249.
[13] Myers, D.E., "Estimation of Linear
     Combinations and Co-Kriging", MATH GEOL.
     15, No. 5, (1983), 633.
[14] Myers, D.E., "Co-Kriging: New
     Developments", to appear in Proceedings of
     NATO ASI, GEOSTAT TAHOE 83, (Reidel
     Publishing, Amsterdam, 1984).
[15] Myers, D.E. and Carr, J., "Co-Kriging and
     Principal Component Analysis: Bentonite
     Data Revisited", to appear in SCIENCES DE
     LA TERRE, (1984)