CO-KRIGING AND PRINCIPAL COMPONENT ANALYSIS

BENTONITE DATA REVISITED

by Donald E. MYERS* and James CARR**

CONTENTS

---

*Department of Mathematics - University of Arizona - Tucson, Arizona 85721 - USA

**Department of Geological Engineering - University of Missouri-Rolla - Rolla, Missouri 65401 - USA

"ETUDES GEOSTATISTIQUES". Séminaire CFSG - 14-15 Juin 1984 - Fontainebleau, France, in Sc. de la Terre, Sér. Inf., n° 21

ABSTRACT

Borgman and Frahme used principal component analysis to reduce consideration of eleven characteristics of Bentonite to only five using data from deposits in Northeastern Wyoming, U.S.A. In this paper, the authors show that co-kriging is quite feasible for eleven variables and give numerical comparisons with those using the principal component formulation.

RESUME

UNE ETUDE SUR DES DONNEES DE BENTONITE

Borgman et Frahme ont utilisé l'analyse en composantes principales pour réduire de onze à cinq le nombre de caractéristiques à considérer pour des données provenant de gisements de bentonite du Nord-Est Wyoming (U.S.A.). Dans cet article, les auteurs montrent que le co-krigeage est parfaitement faisable avec onze variables et ils établissent une comparaison numérique avec l'analyse en composantes principales de Borgman et Frahme.

A - INTRODUCTION

Bentonite is a clay occurring in layered deposits in Northeastern Wyoming and also in the Black Hills District or South Dakota, U.S.A. It is used for a number of purposes such as binder in foundry sand, pelletizing of taconite and as a constituent of drilling mud for oil wells. Unlike metallic deposits, for Bentonite the economic value of a mining unit or of the deposit is not determined by an ore grade. There are a number of characteristics that are reflected in this value. One approach then is to form a value function which is a linear combination of a number of attributes. In order to evaluate a deposit it is necessary to utilize sample data and estimate "values" for mining units. Kriging is the tool frequently used for such an estimation of ore grades. When a value function is formed by a linear combination and the data is likewise transformed to such values it is known that Kriging the linear combination is not an optimal method. It has the further drawback that if the weights in the value function are changed then the entire estimation process must be repeated including modelling the variogram. Co-Kriging is a multivariate estimation method which avoids both of these drawbacks but which has not received as much attention or application until recently. At least two intermediate approaches have been used. G. Matheron (1979) obtained results for improving the Kriging of linear combinations. Borgman and Frahme (1976) used principal components analysis to reduce to a smaller number of uncorrelated variables, each could then be Kriged separately. The comparisons between these two approaches and the feasibility of co-Kriging is the subject of this paper.

B - THE DATA

Knechtel and Patterson (1956) reported measurements on 29 characteristics at 81 locations. Borgman and Frahme considered the data for only eleven variables and only for samples from the Clay Spur beds. The data set is tabulated on the back of map M-36. The locations are given by map number and for the 78 Clay Spur samples, there are only 45 distinct locations. Sample location coordinates in miles were read from the map using 45° N. Lat., 104°20' W. Long. as the origin. Where there were samples with the same map number, the average of the values was utilized for further analysis. The eleven variables included were the same as those in Borgman and Frahme, and are listed below.

NC Percent nonclay material.

GR Percent grit (by weight) retained on a 200 mesh screen.

GC Green compression strength in psi for 2% tempering water using sand bonded with 4% clay.

DC Dry compression strength in psi for 2% tempering water using sand bonded with 4% clay.

SW Swelling capacity of 2 grams of bentonite in milliliters.

YC Yield in barrels of 15 centipoise slurry that can be made from one ton of bentonite.

PH pH

SL Viscosity in centipoises for a slurry containing 6% clay by weight.

FL Thickness of filtrate after 30 minutes for a suspension containing 6% clay by weight.

IC Initial gel strength in grams for a slurry containing 6% clay by weight.

TC Gel strength in grams after 10 minutes for the slurry used in (10) above.

C - MULTIVARIATE METHODS

The Bentonite data constitutes a 45 x 11 array. This may be interpreted as 45 points in 11 dimensional space on 11 points in 45 dimensional space. In either case it is desirable to ask whether the cloud of points (or vectors) can be described in a lower dimensional space. Principal components analysis (PCA) is one method for achieving this. By using a Euclidean distance and after normalization, a set of eigenvalues and eigenvectors is extracted from the correlation matrix. The eigenvectors (principal components) are orthogonal (uncorrelated) and represent the directions of the cloud of points. The eigenvalues represent the moments of inertia in the respective directions. This approach was followed by Borgman and Frahme. Each original point is represented as a linear combination of the new components. For the Bentonite data using PCA, five factors explain 88% of the variance. Variograms and cross-variograms were then computed for the coefficients but no co-Kriging results were given by Borgman and Frahme. Let D_{ij} be the data for the i-th characteristic in the j-th sample. Set

$$\bar{D}_i = \frac{1}{45} \sum_{j=1}^{45} D_{ij} \qquad (1)$$

$$S_i^2 = \frac{1}{45-1} \sum (D_{ij} - \bar{D}_i)^2 \qquad (2)$$

$$\bar{D} = \begin{bmatrix} \bar{D}_1 \\ \vdots \\ \bar{D}_{11} \end{bmatrix}, \quad S^2 = \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & S_{11} \end{bmatrix} \qquad (3)$$

Then the normed data matrix is

$$Z = S^{-1} (D - \bar{D}1^T) \tag{4}$$

and the correlation matrix is

$$R = \frac{1}{45} Z Z^T \tag{5}$$

It is the eigenvectors and eigenvalues of R that are the principal components and indicate the percent of variance explained.

R-mode principal component factor analyses for the bentonite data yielded the first five eigenvalues as (4.84, 1.78, 1.60, 0.82, 0.64). Thus the first five eigenvectors "explain" 88% of the data variance. The corresponding five vectors $\underline{E}_1, \underline{E}_2, \ldots, \underline{E}_5$ and eleven scalars, σ_i are given in Table 1 below.

i	σ_i	$\underline{E}_1$	$\underline{E}_2$	$\underline{E}_3$	$\underline{E}_4$	$\underline{E}_5$
1	4.14	.179	.250	.493	-.105	-.285
2	1.59	.144	.291	.395	-.536	.584
3	0.72	-.141	-.555	-.120	.019	.510
4	10.80	.021	.501	-.020	.727	.400
5	7.91	-.400	-.106	.048	.065	.142
6	19.66	-.429	.012	.083	-.098	.014
7	0.84	-.216	.324	-.460	-.252	-.263
8	15.89	-.429	.076	.145	-.022	-.030
9	5.49	.188	-.411	.428	.244	-.224
10	23.55	-.390	.041	.313	.133	-.120
11	48.44	-.407	.042	.251	.111	-.047

Table 1. The first five eigenvectors for the bentonite data and the standard deviations, σ_i.

Another method for reducing the number of variables has been developed by Benzecri (1973). Correspondence analysis proceeds by normalizing the data set in a different manner. If all entries in the data array are added and then each entry divided by this sum, the resulting entries could be interpreted as probabilities. That is, the original array is analagous to a contingency table. Instead of a Euclidean metric a chi-square metric is used. For CA the matrix S is formed where

$$S_{jk} = \sum_{i=1}^{n} \frac{D_{ik}}{D_{i\cdot} D_{\cdot j}} \qquad (6)$$

$$\text{where } D_{i\cdot} = \sum_{j=1}^{m} D_{ij} \qquad (7)$$

$$D_{\cdot j} = \sum_{i=1}^{n} D_{ij} \qquad (8)$$

It is the eigenvalues and eigenvectors of this matrix that are used in the representation in lower dimensional space. One of the advantages of Correspondence Analysis (CA) is that the normalization is symmetric and the eigenvalues, eigenvectors can be extracted simultaneously whether rows or columns are viewed as points. When CA was applied to the Bentonite data 3 factors explained 88% of the variance. The coordinates of those factors in terms of the original variates is given in Table 2.

	J Factors (Variables Coordinates)				
NC	.268532	.082745	.010257	.076746	-.038046
GR	.322280	.144658	.151971	.093242	-.074404
GC	.169334	.365915	.915241	.010818	-.077004
DC	.246331	.120275	-.028997	.019158	.072338
SW	.073490	-.088080	.271276	-.062697	-.002881
YT	.070724	.159783	-.079847	-.070381	-.002881
PH	.509701	-1.142226	-.381855	-.043551	.009954
SL	-.350470	.006292	-.027142	-.204010	-.034154
FL	.336272	.119427	-.063144	.228225	-.119753
IC	-.488440	-.048314	-.148560	.016706	-.202629
TC	-.500948	-.105691	.012118	.090878	.091321

Table 2

	WEIGHT	AC(1)	RC(1)	AC(2)	RC(2)	AC(3)	RC(3)
NC	.046454	3.53	83.46	.52	7.92	.01	.12
GR	.006765	.74	64.07	.23	12.91	.42	14.25
GC	.024937	.75	2.85	5.46	13.31	56.11	83.24
DC	.188327	12.04	74.37	4.45	17.73	.43	1.03
SW	.103691	.59	5.95	1.31	8.54	20.50	81.03

YT	.263388	1.39	11.95	10.99	60.98	4.51	15.23
PH	.034020	9.31	15.17	72.54	76.20	13.32	8.52
SL	.050693	6.56	73.82	.00	.02	.10	.44
FL	.058784	7.00	57.18	1.37	7.21	.63	2.02
IC	.065900	16.56	78.40	.25	.77	3.91	7.25
TC	.157041	41.52	89.99	2.87	4.01	.06	.05

Table 3

Table 3 provides diagnostic information. Note that the AC column adds to 100(percent) for each factor. Likewise the row sums for RC would add to 100 if all factors were listed. Factor 1 is primarily determined by TC and this factor is the dominant component of all except GC, SW, YT, PH. Factor 2 is principally PH and is the dominant component of PH, YT. Factor 3 is principally GC and is the dominant component of GC, SW. The computer program used was adapted for use on the CDC 6400 from that given by David, Dagbert, Beauchemin (1977). The output also provides the factors for samples and plots the variables, samples on 2 factor coordinate systems. Valenchon (1982) has also described the application of CA to geochemical data. Zhou, Chang and Davis (1983) have described the appropriate applications for RQ mode PCA vs CA.

Note that neither PCA nor CA is an estimation method but rather are tools to reduce the number of variables.

D - CO-KRIGING

When the variables of interest are not only spatially correlated but also inter-correlated then joint estimation provides better estimation than separate estimation. This is known as co-Kriging and is a minimum variance linear estimation. It has not been widely used in mining applications except in the under sampled form and only with a limited choice of cross-variograms. The general formulation of co-Kriging described here is that given in Myers (1982), (1983), (1984). The numerical results were obtained using the computer program given in Carr, Myers and Glass (1984). The objective was to show that co-Kriging is feasible even for 11 variables and to compare the results with those obtained using co-Kriging on Borgman and Frahme's reduced variable model.

Let $Z_1(x), \ldots, Z_m(x)$ denote the m variables at location x or simply

$$\overline{Z}(x) = [Z_1(x), \ldots, Z_m(x)] \qquad (9)$$

The data is n row matrices $\overline{Z}(x_1), \ldots, \overline{Z}(x_n)$. For an unsampled location x_0, the co-Kriging estimator is given by

$$\overline{Z}(x_0) = \sum_{i=1}^{n} \overline{Z}(x_i)\Gamma_i \qquad (10)$$

where each Γ_i is an $m \times m$ matrix of weights. The co-Kriging equations may be written as

$$\sum_{i=1}^{n} \overline{\gamma}\,(x_i - x_j)\Gamma_j + \overline{\mu} = \overline{\gamma}\,(x_0 - x_i) \qquad (11)$$

$$\sum_{i=1}^{n} \Gamma_i = I$$

$\overline{\gamma}$ is the variogram matrix, $\overline{\mu}$ is the Lagrange multipler matrix and I an identity matrix. The Kriging variance is

$$\mathrm{Tr}[\Sigma\, \overline{\gamma}\,(x_0 - x_i)\, \Gamma_i + \overline{\mu}] \qquad (12)$$

where Tr denotes the trace. Myers (1983) describes the relationship of co-Kriging linear combinations to Kriging linear combinations. Davis and Greenes (1983) have given another example of the use of PCA to reduce the number of variables and effectively avoid co-Kriging. However such an approach is not really necessary as is shown by the results in this paper. Moreover, it is not possible to obtain a variance of the error of estimation of the original variables when PCA or CA is combined with Kriging or co-Kriging. For this reason it is preferable to use co-Kriging especially when estimating linear combinations.

E - VARIOGRAMS AND CROSS-VARIOGRAMS

In order to utilize Co-Kriging for the Bentonite data it is necessary to model variograms for 11 variables and cross-variograms for 55 pairs of variables. For variograms, standard valid models such as spherical, power, exponential or gaussian can be considered. For cross-variograms however the possible models are not so easily identified and plotting sample cross variograms is not an adequate guide for determining the models.

The method used is that described in Myers (1983), namely for each pair of variables form a new variable by the sum. Model the variogram for this new variable and form

$$\frac{1}{2}\left[\gamma_{12}^{+} + (h) - \gamma_1(h) - \gamma_2(h)\right] \qquad (13)$$

this is the cross-variogram. since γ_{12}^{+} , γ_1 , γ_2 are modelled separately it is necessary to check the Cauchy-Schwartz Inequality

$$\left|\gamma_{12}(h)\right| \leq \left[\gamma_1(h)\ \gamma_2(h)\right]^{1/2} \qquad (14)$$

The variograms can be validated by cross-validation in the usual way. Finally the cross-variograms and variograms can be cross-validated collectively by Co-Kriging.

Table 4 gives the variogram values for the eleven variables (BF denotes Borgman and Frahme and identifies the results given in that paper. MC refers to Myers and Carr).

Table 5 lists the variogram models actually used and Table 6 the cross-variogram models used. Table 7 tabulates the cross validation results for the eleven variograms as given by the three Kriging procedurs as well as for Co-Kriging.

Table 4

Comparison of Variogram Values

Variable		h (miles) 0	2	4	Variance
1	BF data	13.9	12.2	26.9	
	BF app	13.2	11.6	13.9	
	MC data	2.50	4.67	6.95	6.32
2	BF data	2.51	2.89	3.07	
	BF app	2.56	2.45	2.39	
	MC data	1.70	1.68	2.03	1.70
3	BF data	0.73	0.42	0.51	
	BF app	0.48	0.42	0.42	
	MC data	0.28	0.29	0.27	0.28
4	BF data	184.0	82.0	100.0	
	BF app	130.0	110.0	100.0	
	MC data	37.30	37.40	43.50	47.30
5	BF data	65.00	54.0	52.0	
	BF app	51.0	38.0	41.0	
	MC data	39.90	39.70	41.30	39.90
6	BF data	309.0	304.0	352.0	
	BF app	349.0	258.0	280.0	
	MC data	200.0	228.0	249.0	251.0
7	BF data	0.57	0.68	0.52	
	BF app	0.62	0.53	0.61	

	MC data	0.40	0.41	0.39	0.53
8	BF data	272.0	149.0	159.0	
	BF app	235.0	175.0	192.0	
	MC data	90.0	93.30	113.0	122.0
9	BF data	32.9	24.10	26.50	
	BF app	27.2	22.50	25.50	
	MC data	16.0	16.70	17.10	26.30
10	BF data	629.0	387.0	358.0	
	BF app	508.0	393.0	444.0	
	MC data	307.0	293.0	286.0	307.0
11	BF data	2203.0	2164.0	1709.0	
	BF app	2140.0	1620.0	1810.0	
	MC data	1800.0	2080.0	1720.0	1800.0

Table 5

Single Variable Relationships

Variable	Nugget	Sill	Range (miles)	Structure
1	2.50	6.32	4.0	Spherical
2	1.70	1.70	10.0	Random
3	0.28	0.28	10.0	Random
4	37.30	47.30	6.0	Spherical
5	39.90	39.90	10.0	Random
6	200.00	251.00	4.0	Spherical
7	0.40	0.53	12.0	Linear
8	90.00	122.00	6.0	Spherical
9	160.00	263.00	6.0	Gaussian
10	307.00	307.00	10.0	Random
11	1800.00	1800.00	10.0	Random

Table 6

Cross Variogram Parameters

Pair	Nugget	Sill	Range	Model
1-2	0.000	0.500	4.000	1
1-3	0.000	0.100	4.000	1
1-4	0.000	0.400	4.000	1
1-5	0.100	0.100	100.000	3
1-6	0.100	0.100	100.000	3
1-7	0.000	0.100	4.000	1
1-8	0.000	0.400	4.000	1
1-9	0.000	0.400	4.000	1
1-10	0.100	0.100	100.000	3
1-11	0.100	0.100	100.000	3
2-3	0.100	0.100	100.000	3
2-4	0.000	0.400	4.000	1
2-5	0.100	0.100	100.000	3
2-6	0.000	0.100	4.000	1
2-7	0.100	0.100	100.000	3
2-8	0.000	0.400	4.000	1
2-9	0.000	0.400	4.000	1
2-10	0.100	0.100	100.000	3
2-11	0.100	0.100	100.000	3

3-4	0.000	0.100	4.000	1
3-5	0.100	0.100	100.000	3
3-6	0.000	0.100	4.000	1
3-7	0.000	0.100	4.000	1
3-8	0.000	0.100	4.000	1
3-9	0.000	0.100	4.000	1
3-10	0.100	0.100	100.000	3
3-11	0.100	0.100	100.000	3
4-5	0.000	20.000	4.000	1
4-6	0.000	30.000	4.000	1
4-7	0.000	0.100	4.000	1
4-8	0.000	30.000	4.000	1
4-9	0.000	30.000	4.000	1
4-10	0.000	30.000	4.000	1
4-11	0.100	0.100	100.000	3
5-6	0.000	10.000	4.000	1
5-7	0.000	0.100	4.000	1
5-8	0.000	10.000	4.000	1
5-9	0.000	0.200	4.000	1
5-10	0.100	0.100	100.000	3
5-11	0.100	0.100	100.000	3
6-7	0.000	0.100	4.000	1
6-8	0.000	10.000	4.000	1
6-9	0.000	10.000	4.000	1
6-10	0.000	0.400	4.000	1
6-11	0.100	0.100	100.000	3
7-8	0.000	0.100	4.000	1
7-9	0.000	0.100	4.000	1
7-10	0.100	0.100	4.000	1
7-11	0.100	0.100	4.000	1
8-9	0.000	0.400	4.000	1
8-10	0.000	0.200	4.000	1
8-11	0.100	0.100	4.000	1
9-10	0.100	0.100	4.000	1
9-11	0.100	0.100	4.000	1
10-11	0.100	0.100	4.000	1

Model: 1 = Spherical

3 = Linear

Table 7

Mean Square Error Estimation Results

(42 sample locations)

Variable	Co-kriging	Myers/Carr	KRIGING 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	27.231	26.69
10	324.61	326.92	326.92	327.23
11	2037.27	2098.17	2098.17	2127.50

F - COMPUTATIONAL EXPENSES

In addition to complexity and the need to model many cross-variograms it is usually assumed that the computer time required to solve large Co-Kriging systems will be excessive. To estimate 11 variables for 45 locations required 30 minutes of CPU time on an IBM 4341. Separate Kriging of eleven variables required 5 minutes on the same system.

REFERENCES

Aboufirassi, M. and Marino , M.A., (1984), Co-Kriging of Aquifer Transmitivity from Field Measurements of Transmitivity and Specific Capacity, Math Geol. 16, No. 1, pp. 19-36.

Benzecri , J.P., (1973), l'Analyse des donnees , Tome II, l'Analyse des correspondances, Dunod, Paris, 619 p.

Borgman, L.E. and Frahme, R.B., (1976), Multivariate properties of Bentonite in Northeastern Wyoming, in Advanced Geostatistics In The Mining Industry, D. Reidel, pp. 381-390.

Carr, J., Myers, D.E. and Glass, C.H., (1984), Co-Kriging: A Computer Program, to appear in Computers And Geosciences.

Davis, B.M. and Greenes, K.A., (1983), Estimation using spatially distributed multivariate data: An example with coal quality, Math Geol. 15, #2, pp. 287-300.

Francois-Bongarcon, D., (1981), Les Corégionalisations, le Co-Krigeage, C-86, Centre de Géostatistique , Fontainebleau.

Hill, M.O., (1974), Correspondence Analysis: A neglected multivariate method, J. Royal Stat. Soc. Ser. C. (Applied Stat.) 23, No. 3, pp. 340.

Knechtel, M.M. and Patterson, S.H., (1956), Bentonite Deposits of the Northern Black Hills District, Montana, Wyoming and South Dakota, Mineral Investigations Field Studies Map MF-36, U.S. Geological Survey, Washington, D.C.

Matheron, G., (1979), Recherche de Simplification dans un problème de co-krigeage, N-628, Centre de Géostatistique, Fontainebleau.

Miesch, A.T., (1983), Correspondence Analysis in Geochemistry, Math Geol. 15, No. 3, pp. 501.

Myers, D.E., (1982), Co-Kriging: The Matrix Form, Math Geol. 14, No. 3, pp. 249-257.

Myers, D.E., (1983), Co-Kriging and Estimation of Linear Combinations, Math Geol. 15, No. 5, pp. 683-687.

Myers, D.E., (1984), Co-Kriging: New Developments, to appear in Proceedings of NATO ASI, GEOSTAT TAHOE 83, Reidel Publishing Co.

Orfeuil, J.P., (1973), L'Analyse des données, N-353, Centre de Géostatistique, Fontainebleau.

Teil, H., (1975), Correspondence factor analysis: an outline of its method, Math Geol. 7, No. 1, pp. 3-12.

Valenchon, F., (1982), The use of correspondence analysis in geochemistry, Math Geol. 14, No. 4, pp. 331-342.

Valenchon, F. (1983), Correspondence Analysis in Geochemistry: Reply, Math Geol. 15, No. 3, pp. 505.

Zhou, Di, Chang, T. and Davis, J.C., (1983), Dual Extraction of R-mode and Q-mode Factor Solutions, Math Geol. 15, No. 5, pp. 581.