3D Property Modeling of Void Ratio by Cokriging

Yao Lingqing* (姚凌青)
State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China; Key Laboratory of Orogenic Belts and Crustal Evolution, Ministry of Education, Peking University, Beijing 100871, China
Pan Mao (潘懋)
Key Laboratory of Orogenic Belts and Crustal Evolution, Ministry of Education, Peking University, Beijing 100871, China
Cheng Qiuming (成秋明)
State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan 430074, China; Department of Earth and Space Science and Engineering, York University, Toronto ON M3J1P3, Canada

ABSTRACT: Void ratio measures compactness of ground soil in geotechnical engineering. When samples are collected in certain area for mapping void ratios, other relevant types of properties such as water content may be also analyzed. To map the spatial distribution of void ratio in the area based on these types of point, observation data interpolation is often needed. Owing to the variance of sampling density along the horizontal and vertical directions, special consideration is required to handle anisotropy of estimator. 3D property modeling aims at predicting the overall distribution of property values from limited samples, and geostatistical method can be employed naturally here because they help to minimize the mean square error of estimation. To construct 3D property model of void ratio, cokriging was used considering its mutual correlation with water content, which is another important soil parameter. Moreover, K-D tree was adopted to organize the samples to accelerate neighbor query in 3D space during the above modeling process. At last, spatial configuration of void ratio distribution in an engineering body was modeled through 3D visualization, which provides important information for civil engineering purpose.

KEY WORDS: property modeling, cokriging, variogram, anisotropy, K-D tree.

INTRODUCTION
Since the concept of geostatistics was developed (Matheron, 1963), various kriging techniques have extensively been applied in various fields of geosciences such as reservoir estimation in mining and petroleum industries, and parameter estimation in environment inspection (Saby et al., 2006; Sepaskhah et al., 2005; Caers, 2001; Houlding, 1999). Kriging algorithm, which is the key method in geostatistics, yields the best linear unbiased estimator (BLUE) considering both spatial dependency and randomness between variables. 3D property modeling indicates process to build a reasonable mapping or predictor of
variables’ properties with spatial locations in 3D space. As the number of samples is usually limited, estimation or interpolation for locations without observations becomes essential in 3D property modeling. Although similar researches were abundant in 2D space using geostatistical methods or other interpolation techniques, 3D property modeling by kriging methods still remains a challenging job.

From 2D to 3D interpolation, the difficulty mainly emerges from implementation rather than geostatics theory. One possible way is to transform the problem from 2D to 3D space by dividing 3D space into a number of layers at various depths. This method neglects spatial correlation of variables along the vertical direction. Another way is to consider the whole interpolation process in 3D environment, which composes our major proposition. The key problems lie in two central aspects: (a) structural analysis about variations of different scales and anisotropic directions; and (b) efficient searching neighbor samples as predictors in 3D space. The first one can be figured out by analyzing the most remarkable variation directions and choosing the appropriate scale according to the samples’ spatial feature. As for the latter problem, the K-D tree was adopted to improve neighbor query efficiency.

In this article, the multivariate geostatistical method (Wackernagel, 1998) was used. While the primary variable to be predicted has correlations with secondary variables, cokriging is seemingly more feasible as the estimation variance can be reduced by incorporating cross-correlation among multiple variables. Additional structural analysis of spatial dependency in cross-correlation has to be done when using the cokriging method. Cokriging is used in this article for 3D property modeling of void ratio by incorporating void ratio and water content.

**COKRIGING METHOD IN 3D PROPERTY MODELING**

**Basic Principle of the Cokriging Method**

Cokriging is one of the commonly used multivariate geostatistical methods. In geostatistical theory, the geological variables named regionalized variables are regarded as random functions with stochastic patterns as well as spatial dependency. Geostatistical analysis uses variogram to characterize spatial correlations among regionalized variables. Kriging methods have been proved to acquire best linear unbiased estimator. However, common kriging methods are usually used in univariate conditions.

Cokriging is derived from the common kriging method, but it deals with multivariate instances. The variable to be predicted is called the primary variable, and the other types of variables correlated with it are termed secondary variables. Cokriging assumes some sort of spatial stationarity, and pre-specified models of spatial dependency in regionalized variables. These are usually variogram and cross-variogram, explaining the spatial dependency of the same type of variable from two samples at different locations. The cokriging estimator is a linear combination of observed data from the primary variable and the secondary variables, shown in equation (1)

$$Z_0^* = \sum_{k=0}^{K} \sum_{i=1}^{N_k} \lambda_{ik} z_{ik}$$

Without loss of generality, the variable with index 0 represents primary variable; where $Z_0^*$ indicates cokriging estimator at the location $x$; $K$ stands for the number of different secondary variables and $k$ is the index; $N_k$ indicates the total number of samples taken into account with index $k$; $z_{ik}$ indicates the $i$th value of variable with index $k$ and $\lambda_{ik}$ indicates its weight in estimation, and $x_{ik}$ indicates its position. The constraint ensuring unbiased estimation is shown as equation (2)

$$\sum_{i=1}^{N_k} \lambda_{ik} = \begin{cases} 0 & k \neq 0 \\ 1 & k = 0 \end{cases}$$

The cokriging system can be deduced by minimizing the estimation variance, denoted as equation (3)

$$\sum_{k=0}^{K} \sum_{i=1}^{N_k} \gamma_{ik}(x, x_{ik}) \lambda_{ik} + \mu_k = \gamma_{0k}(x, x_{0k})$$

for $k = 0, 1, \ldots, K; j = 1, 2, \ldots, N_k$

$$\sum_{i=1}^{N_k} \lambda_{ik} = 1$$

$$\sum_{i=1}^{N_k} \lambda_{ik} = 0$$

where $\gamma_{ik}(x, y)$ stands for the value of variogram or cross-variogram decided by the distance $(x-y)$, and $\mu_k$ represents Lagrange multipliers. Equation (4) lists the cokriging estimator variance, $\sigma_e^2$, for
med with variogram and cross-variogram

\[ \sigma^2_k = \sum_{i=1}^{N} \sum_{k=1}^{K} \lambda_{ik} \gamma_{ik}(x, x_k) + \mu_k \]  

(4)

After resolving the cokriging system, the weights of each observed sample in the specified range are assigned to yield the optimal estimator.

**Structural Analysis**

Geostatistics assumes some sort of spatial stationarity so that the variogram merely depends on the distance of two samples, independent of their locations. Under this assumption, variogram and cross-variogram are expressed as equations (5) and (6)

\[ \gamma(h) = \frac{1}{2} E[(Z(x) - Z(x + h))^2] \]  

(5)

\[ \gamma_{kl}(h) = \frac{1}{2} E[(Z_k(x + h) - Z_l(x))[(Z_k(x + h) - Z_l(x))] \]  

(6)

where \( E \) indicates the expectation of random variable; \( Z(x) \) is regionalized variable; \( h \) is displacement of two locations; \( k \) and \( l \) represent different regionalized variables; and \( \gamma(h) \) and \( \gamma_{kl}(h) \) indicate variogram and cross-variogram, respectively.

However, the assumption of stationarity is very restrictive. In several occasions, the structure of spatial dependency varies from location to location. In addition, spatial dependency may be anisotropic so that variograms or cross-variograms may have different spatial variation patterns in some directions. Sometimes, the scales of correlation observed on variograms also vary in different directions. This can be caused by different sampling sizes or sampling densities in different orientations. The ultimate model of variogram comes from curve-fitting of experimental variogram. Classical formulation to calculate experimental variogram is given as equation (7)

\[ \gamma'(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \]  

(7)

\( \gamma'(h) \) is the experimental variogram and \( N(h) \) indicates pairs of samples with distance \( h \).

Generally, several typical directions are chosen as the first step of structural analysis and the experimental variogram is achieved through statistical operation in these directions. Then, the variogram models are fitted to the experimental variograms. From the preliminary results, the major variation orientation can be speculated by searching the strongest spatial correlation, and similarly, the minor one with the weakest spatial correlation. From the variogram models fitted to the experimental results, the optimal correlation distance can be determined. At last, these variogram models are cumulated and transformed to a uniform nested model with elimination of anisotropy.

The structural analysis of cross-variogram is similar except that the estimation of cross-variogram values is based on sample pairs of different regionalized variables. However, when the samples of coregionalized variables are taken on the same positions, the process of modeling cross-variogram can be simplified by the following equation (8)

\[ \gamma_{kl}(h) = \frac{1}{2} \left[ \gamma_{kl}^*(h) - \gamma_{kl}(h) - \gamma_{kl}(h) \right] \]  

(8)

\( \gamma_{kl}(h) \) and \( \gamma_{kl}(h) \) indicate variograms of variables with index \( k \) and \( l \), respectively; \( \gamma_{kl}^*(h) \) indicates the variogram model determined by a new regionalized variable equivalent to the sum of variables with index \( k \) and \( l \), and \( \gamma_{kl}(h) \) is a cross-variogram.

**Three-Dimensional Neighbor Searching by K-D Tree**

The sample data are associated with spatial locations. These samples can be organized by special index tree for efficient spatial searching during the interpolation process. The K-D tree (Bentley, 1975) was adopted in this article because it behaves well in query scattered points. The principle of the K-D tree is illustrated in Fig. 1.

Figure 1 shows the K-D tree sketch map in 2D space. The points are inserted into left or right child recursively, by judging whether the coordinates of certain axis are less than the critical value. The critical value for dividing two branches can be set as the median in current tree node. When all the points finish inserting, a K-D tree index is accomplished. Neighbor query can be done by comparison with critical values in different depths of the tree, therefore, the query is considerably faster than brute-force global searching.

**APPLICATION BACKGROUND AND DATA ANALYSIS**

The case study is chosen for mapping the spatial distribution of void ratio in a geological body for
engineering purpose. The geological body lies in Yizhuang, a new urban economic development zone in Beijing; spreads extensively in horizontal, yet is thin along the vertical direction. In total, 29 drills were scattered in large scale horizontally with intervals ranging from 800 to 1 000 m. However, the intervals of sampling were considerably smaller in the vertical direction, varying within 1–3 m. Samples of void ratio and water content were taken from these drills, and measurements were taken on the two different variables in 201 different locations. The ultimate goal was to construct the 3D property model of void ratio from the samples.

From the samples, some basic statistical analysis should be made to assure the fulfillment of cokriging method’s preconditions. First, the Pearson correlation coefficient between void ratio and water content is calculated as 0.897. Besides, the QQ-plot of void ratio and water content almost coincides with a straight line (the dashed line in Fig. 2), except few points among the lowest and highest values. Therefore, it is obvious that the two parameters have positive correlation with
each other and most likely belong to the same kind of probability distribution. The statistics calculated from void ratios and water contents from these samples show high credence to the cokriging method for 3D property modeling. Cokriging relies on mutual correlations between different variables.

3D PROPERTY MODELING OF VOID RATIO

3D property modeling of void ratio was confined in certain geological body. The first step is to discretize the geometry into cubic units. This task was accomplished by voxelization algorithm (Kaufman, 1987). In this article, a typical voxelization algorithm (Karabassi et al., 1999) was adopted with minor change; however, its discussion is not given in this article. The geological body was divided into 128×128×128 grids of grid size 64 m×112 m×0.3 m.

The experimental variograms and cross-variograms were calculated in horizontal and vertical directions, respectively, and the basic lags were set to 800 m in horizontal with 300 m tolerance and 1 m in vertical with 0.5 m tolerance. Note that the scales of variations along the two orientations are quite diverse. Through curve-fitting algorithms, the final models were determined. There are six models together including four variograms and two cross-variograms. We only laid out two cross-variograms along the horizontal and vertical directions corresponding to Fig. 3 and Fig. 4, and the other four variograms are not shown to save the length of this article.

After the spatial models were determined, the cokriging system was set up and resolved using numerical linear programming algorithms. The coefficients of the cokriging system were totally decided by their neighbor samples. Thus, the K-D tree played an important role here and was used to retrieve these adjacent samples quickly.

RESULTS

Combined with the structural analysis results, cokriging systems were built through the observed values in neighbor. The void ratios of grid units were estimated, and then the 3D property model was constructed finally.

Figure 5 displays the void ratio’s property modeling results in 3D space. The figure on the left side in Fig. 5 shows the 3D visualization, and the figure on the right exhibits the cross-sections. Thus, it is quite clear and convenient for engineers to judge the physical features of geological body. The inner structures of properties revealed by this technique in 3D space are useful for engineering purposes.

CONCLUSIONS

The example introduced in this article exhibits that 3D property modeling provides a technique essential for visualization of spatial distribution of void ratio and water content. The same technique can be applied in several other fields including mining, groundwater modeling, and environmental modeling. This article proposes cokriging as the method of 3D property modeling. The structural analysis and the estimator are all considered comprehensively in real 3D space. From our experimentation done on void ratio, we can conclude that the difference between spatial structures along the horizontal direction and the vertical direction is fairly obvious, and therefore, its impact on estimation should not be ignored. The correlation between the void ratio and water content
was utilized by cokriging. It has been seen that it is important to find and refer to as much usable information as possible in 3D property modeling practice, to improve the reliability of the final result.

REFERENCES CITED


