Mathematical Geology, Vol. 19, No. 2, 1987

Production of Conditional Simulations via the LU Triangular Decomposition of the Covariance Matrix[1]

Michael W. Davis

This paper reviews the turning band method and fast Fourier transform method of producing a nonconditional simulation of a multinormal random function with a given covariance structure. A review of the two common methods of conditioning the simulation to honor the data shows that they are formally equivalent. Another method for directly pondering a conditional simulation based on the LU triangular decomposition of the covariance matrix is presented. Computational and implementation difficulties are discussed.

KEY WORDS: geostatistics, kriging, conditional simulation, fast Fourier transform.

INTRODUCTION

Conditional simulations are a useful research tool for studying new algorithms and theoretical developments in geostatistics. It is often useful, for performing sensitivity analysis, to have multiple conditional simulations from the same set of data.

This paper describes a simple technique, based on the lower-upper (LU) triangular decomposition of the covariance matrix, for producing multiple conditional simulations of a moderate-sized grid (up to 700 points). The method can be implemented easily by adapting a standard kriging program. A further advantage is that most operations performed in the algorithm can be vectorized to take advantage of the vector processing capability which exist on certain computers. Indeed, for machines such as the Cray 1 and the CDC Cyber 205, subroutines required to implement this procedure already exist as part of the standard FORTRAN library routines and are coded to take advantage of the machines architecture. The method has several advantages over other methods in common use such as the turning band method (Journel and Huijbregts, 1981)

[1]Manuscript received 15 March 1985; accepted 3 July 1986.
[2]Department of Applied Earth Sciences, Stanford University, Stanford, California, 94305.

0882-8121/87/0200-0091$05.00/1 © 1987 International Association for Mathematical Geology

and the fast Fourier transform method (FFT) (Borgman et al., 1984). The advantages are that it is integrated—i.e., simulation and conditioning are performed simultaneously and is simple to apply and can be applied to an arbitrary covariance structure. Also, data locations do not have to coincide with points on the simulated grid.

As presented, the method assumes that all grid points will be simulated at the same time and that all data will be used for conditioning. This restriction obviously can be relaxed and the method can be applied locally in much the same way as moving neighborhood kriging.

PRODUCTION OF A NONCONDITIONAL SIMULATION

Two methods are widely used at present that produce a nonconditional simulation of a random function, the turning band method of Matheron as described in Journel and Huijbregts (1981, p. 498) and the fast Fourier transform method as described in Borgman et al. (1984).

The turning band method works by simulating one-dimensional processes with a particular covariance structure along lines which are distributed equally in space. The appropriate one-dimensional covariance model is deduced from the covariance model of the two- or three-dimensional process. The simulated value at a given grid point is obtained by summing values taken at the closest point on each of the lines, by dropping a perpendicular from the grid point to the line. The method is fast, because only one-dimensional simulation is required. However, not all two-dimensional random functions have covariance models for which a simple one-dimensional covariance model exists, although in most cases a reasonable approximation can be found (see Journel and Huijbregts, 1981, p. 506 and Brooker, 1985). In three dimensions, the turning band method suffers from loss of accuracy due to a finite number of line simulations used. This is linked to the difficulty of generating lines which are distributed evenly in space. The usual approach is to take 15 lines joining midpoints of a regular icosahedron. This leads to what is termed striping in the simulation. Two-dimensional slices of the three-dimensional simulation will have stripes of large or small values caused by runs of large or small values on particular lines. This can be remedied only by generating many more lines. Unfortunately, this only can be done by generating lines drawn through points generated at random on the unit sphere. In this case, a large number of lines will have to be used to avoid the generation of anisotropic bias induced by nonuniform distribution of lines in space. The turning band method is also awkward to apply in cases where a strong anisotropy is present either in the variogram model or the actual geometry of the deposit. Anisotropy is usually rectified by applying a transformation to the coordinate system to remove the anisotropy. However, if the anisotropy is strong this can distort the simulated grid which is usually a cubic lattice. When the grid is back-transformed, un-

wanted anisotropies may be present. These points are all discussed in detail in Luster (1985). The principal advantages of the turning band method are its speed and the number of points in the three-dimensional simulation which can be generated by the method. To date, it has been the only method that can generate the large production oriented simulations, which may require as many as 2,000,000 grid points.

The fast Fourier transform method calculates the energy spectrum of the random function from the covariance model and uses the fact that the sum of independent multinormally distributed variables is also multinormal to produce a simulation with the appropriate covariance structure in the space domain. Being a numerical technique, the method can be applied for any given covariance model. Because FFT subroutines are becoming part of the standard FORTRAN library routines for many computers, particularly those with vector processing capabilities, the method is computationally efficient and easy to use. The present limitation of the method appears to be the size of the simulation which can be generated. This would appear to be an implementation as opposed to a theoretical difficulty.

CONDITIONING THE GRID

Two methods for conditioning the simulation to take the same values at sampled locations are described in Journel and Huijbregts (1981) and Borgman et al. (1984).

The first method works by adding a correlated error obtained from the simulation to a grid kriged from the data; the second method works by calculating the conditional distribution of grid point values given the data values assuming a multivariate normal distribution. The two methods are shown to be formally equivalent in the following paragraphs.

Steps for obtaining a conditional simulation from a nonconditional simulation using the first method are explained in Journel and Huijbreghts (1981, p. 494). Briefly, the idea is that a grid kriged from the data will be smoother than reality. To produce a grid which reproduces the spatial variability of the underlying random function, a new kriged grid is produced by using simulation values at data locations as values used in kriging. This new kriged grid is subtracted from the nonconditional grid to give a grid of correlated errors. This grid of errors is added to the original kriged grid to produce a conditioned simulation.

So

$$Z_{cs} = Z_{kd} + (Z_{us} - Z_{kus})$$

where Z_{cs} is the conditional simulation, Z_{kd} is the grid kriged from real data, Z_{us} is the unconditional simulation, Z_{kus} is the grid kriged from unconditional simulation data. This can be rewritten as

$$Z_{cs} = Z_{us} + (Z_{kd} - Z_{kus})$$

Because data locations are the same, $Z_{kd} - Z_{kus}$ can be obtained by kriging differences between data values and simulated values at data locations. The explicit representation of a kriged grid for the case of simple kriging with a zero mean gives the estimate for a single point as

$$Z^*(x, y) = \mathbf{k}'(x, y)\mathbf{K}^{-1}\mathbf{z}$$

where $\mathbf{K}$ is the matrix of covariances between the data points, $\mathbf{z}$ is the vector of data values, and $\mathbf{k}$ is the vector of covariances between the point and the data. [See Davis and Culhane, (1984, p. 602, for details.) For a grid of m points this can be rewritten as

$$\mathbf{Z} = \mathbf{C}'\mathbf{K}^{-1}\mathbf{z}$$

where $\mathbf{Z}$ is a vector of estimated grid values and $\mathbf{C}$ is an $m \times n$ matrix containing vectors $\mathbf{k}$ for each individual value estimated.

Expressing this in the same notation as Borgman et al. (1984)

$$\mathbf{Z}_{kd} - \mathbf{Z}_{kus} = \mathbf{C}_{21}\mathbf{C}_{11}^{-1}(\mathbf{z} - \mathbf{y})$$

where $\mathbf{z}$ is the actual data value, $\mathbf{y}$ is the simulated value at the data location, $\mathbf{C}_{11}$ is the covariance between the data (equivalent) $\mathbf{K}$, $\mathbf{C}_{21}$ is the matrix containing the vectors $\mathbf{k}$ equivalent $\mathbf{C}'$. So

$$Z_{cs} = Z_{us} + \mathbf{C}_{21}\mathbf{C}_{11}^{-1}(\mathbf{z} - \mathbf{y})$$

This is the same formula used by Borgman et al. (1984) to condition the simulation in the space domain. It is derived by considering the conditional covariance of the grid points given the data assuming a multivariate normal distribution. So, formally, the two methods of conditioning are demonstrated to be equivalent.

THEORETICAL BACKGROUND OF PROPOSED METHOD

Let $\mathbf{K}$ be the covariance matrix associated with data locations $(x_1, x_2, \ldots, x_n)$. Matrix $\mathbf{K}$ is symmetric and positive-definite and hence can be decomposed into the product of a lower triangular and an upper triangular matrix. Furthermore, because $\mathbf{K}$ is symmetric, the upper triangular matrix is the transpose of the lower triangular matrix

$$\mathbf{K} = \mathbf{LU} \qquad \text{where} \qquad \mathbf{L}' = \mathbf{U}$$

Now consider the random vector

$$\mathbf{y} = \mathbf{Lw}$$

where $\mathbf{w}$ is a vector of independent $N(0, 1)$ distributed random numbers. The

expected value of the $n \times n$ matrix $\mathbf{yy}'$ is given by

$$E(\mathbf{yy}') = E(\mathbf{Lww}'\,\mathbf{U})$$
$$= \mathbf{L}E(\mathbf{ww}')\mathbf{U}$$

Because $\mathbf{w}$ is a vector of independent $N(0, 1)$ random numbers

$$E(\mathbf{ww}') = \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix

$$E(\mathbf{yy}') = \mathbf{LIU}$$
$$= \mathbf{LU}$$
$$= \mathbf{K}$$

So vector $\mathbf{y}$ is a nonconditional autocorrelated simulation of the random function at data locations with covariance matrix $\mathbf{K}$.

CONDITIONAL SIMULATION OF A GRID

Consider the covariance matrix of data locations the grid to be generated, and partition the matrix as follows

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

where C_{11} is covariance between data points $(\mathbf{K})$, C_{21} is covariance between data points and grid points, and C_{22} is covariance between grid points. (If a data point and a grid point happen to coincide, the point should be considered as a data point and grouped accordingly. Also, no duplication of data should occur at the locations.)

Now decompose $\mathbf{C}$ as follows

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \mathbf{LU} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix}$$

Let $\mathbf{w}' = (\mathbf{w}_1', \mathbf{w}_2')$ be a vector of independent $N(0, 1)$ distributed random numbers.

Then, vector $\mathbf{y}' = (\mathbf{y}', \mathbf{y}_2')$ given by

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$$

is an unconditional simulation of the random function at the grid and data points with covariance matrix $\mathbf{C}$. Now

$$\mathbf{L}_{11}\mathbf{w}_1 = \mathbf{y}_1 \quad \text{and}$$

$$L_{21}\mathbf{w}_1 + L_{22}\mathbf{w}_2 = \mathbf{y}_2$$

To condition the simulation, replace $\mathbf{w}_1$ by $\mathbf{v}_1$ where $\mathbf{v}_1$ is the solution to the equation

$$L_{11}\mathbf{v}_1 = \mathbf{z}_1$$

where $\mathbf{z}_1$ is a vector of values at the data points. (Here data are assumed to have been transformed so that $\mathbf{Z}(\mathbf{x})$ is a $N(0, 1)$-distributed random function.) Note also that $\mathrm{E}(\mathbf{z}_1\mathbf{z}_1') = \mathbf{C}_{11}$.

Therefore

$$\mathbf{v}_1 = L_{11}^{-1}\,\mathbf{z}_1$$

So

$$\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}\begin{bmatrix} L_{11}^{-1}\,\mathbf{z}_1 \\ \mathbf{w}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_1 \\ L_{21}L_{11}^{-1}\mathbf{z}_1 + L_{22}\mathbf{w}_2 \end{bmatrix}$$

produces a conditional simulation of the random function. Multiple simulations may be produced simply be regenerating the term $L_{22}\mathbf{w}_2$ using different random numbers. Note that $\mathbf{w}_1$ is not needed.

COMPUTATIONAL DETAILS

Multiplying the decomposition

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix}\begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix}$$

$$= \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} + L_{22}U_{22} \end{bmatrix}$$

Therefore

$$C_{11} = L_{11}U_{11}$$

$$C_{21} = L_{21}U_{11}$$

$$C_{12} = L_{11}U_{12}$$

$$C_{22} = L_{21}U_{12} + L_{22}U_{22}$$

So L_{11} and U_{11} are calculated easily via the modified Cholesky decomposition. (The reason for using the modified Cholesky decomposition instead of the usual $\mathbf{R}'\mathbf{R}$ Cholesky decomposition of a positive definite matrix is that the former is more stable numerically and easier to implement. This is discussed in Davis and Culhane, (1984, p. 602.) Once U_{11} is known, L_{21} can directly calculated by

$$L_{21} = C_{21}U_{11}^{-1}$$

The FORTRAN subroutine UDUB given in Davis and Grivet (1984) easily can be adapted to perform this decomposition. Note that the algorithm factors a symmetric positive definite matrix $\mathbf{A}$ as follows

$$\mathbf{A} = \mathbf{U}'\mathbf{D}\mathbf{U}$$

where $\mathbf{D}$ is a diagonal matrix and $\mathbf{U}$ is an upper triangular matrix. So, using this progrom, the $\mathbf{L}$ matrix required to perform the simulation would be given by

$$\mathbf{L} = \mathbf{U}'\mathbf{D}^{1/2}$$

Because $\mathbf{D}$ is a diagonal matrix whose elements are reciprocals of diagonal elements of the matrix $\mathbf{U}$, $\mathbf{D}^{1/2}$ can be calculated easily. (Note that matrix $\mathbf{U}$ used in this paper is *not* the same as matrix $\mathbf{U}$ referred to by the program.)

CONCLUSION

The described technique provides a simple method of producing a conditional simulation of an intermediate-sized grid (at present the largest grid which has been simulated by the method contains 700 grid nodes plus data). The technique is useful for simulating a large number of small grids for studying sampling distributions. This is described in Solow (in preparation). The method as presented also can be applied locally to subareas, and this is often sufficient for most research and many production problems. Although the method as presented is restricted in the size of the grid it can generate, this is an implementational difficulty as opposed to a theoretical limitation. These problems can be circumvented by taking into account the particular structure of the covariance matrix, which will be sparse for a spherical covariance, or by adapting the algorithm to work locally as in moving neighborhood kriging, which is particularly appropriate in areas where a global stationary random function model is not adequate to describe the data. These topics will be discussed in a future paper.

REFERENCES

Brooker, Peter I., 1985. Two-Dimensional Simulation by Turning Bands: Math. Geol., v. 17, no. 1, p. 81-90.

Borgman, L., Taheri, M., and Hagan, R., 1984. Three-Dimensional, Frequency-domain Simulations of Geological Variables, *in* Verly, G., David, M., Journel, A. G., and Marechal, A. (Eds.), Geostatistics for Natural Resource Characterization, Part 2: Reidel, Dordrecht, Netherlands,—p.

Dahlquist, G. and Bjorck, A., 1974. Numerical Methods: Prentice-Hall, New Jersey,—p.

Davis, Michael W. and Culhane, P. G., 1984. Contouring Very Large Datasets Using Kriging, *in* Verly, G., David, M., Journel, A. G., and Marechal, A. (Eds.), Geostatistics for Natural Resource Characterization, Part 2: Reidel, Dordrecht, Netherlands, p.

Davis, Michael W. and Grivet, Cybil, 1984. Kriging in a Global Neighborhood: Math. Geol., v. 16, no. 3, p. 249–265.

Journel, A. G. and Huijbregts, Ch. J., 1981. Mining Geostatistics: Academic Press, London, 600 p.

Luster, G. R., 1985. Raw Materials for Portland Cement: Applications of Conditional Simulation of Coregionalization, Doctoral Dissertation, Department of Applied Earth Science: Stanford University, Stanford, California,—p.

Solow, A. R., in preparation, Bootstrapping Correlated Data: