

Automatic Variogram Modeling from Censored Variogram Volumes

Brandon J. Wilde and Clayton V. Deutsch

The variogram model is an important input statistic in geostatistical modeling. Care must be taken to ensure that the model honors the true spatial continuity. Determining the most appropriate variogram model can be difficult, particularly in three-dimensional cases. The typical procedure is to calculate a large number of experimental one-dimensional variograms. These one-dimensional variograms are studied and compared to determine which direction shows the greatest continuity (the major direction). The orthogonal direction to the major direction which shows the next greatest continuity is then identified. The variogram along the direction orthogonal to these two directions is computed. Now the procedure of determining an appropriate variogram model begins. The practitioner or some automatic software compares the variogram model to the experimental variograms. The variogram model is adjusted until a satisfactory fit is achieved. A variogram volume is the calculation of the variogram for all practical distances and directions. It is often used to identify directions of anisotropy prior to the calculation of experimental directional variograms. Though not commonly performed in practice, the step of calculating directional variograms could be dropped and the variogram model could be fit to the variogram volume. This note documents a GSLIB-style program which performs these two steps. The program `vormap_varfit` calculates a variogram volume and then fits a variogram model to the variogram volume automatically. The moment of inertia method is used to identify the principal directions of anisotropy. The program outputs a variogram model which best fits the variogram volume. This could be used as a starting point for further variogram analysis or it could be used directly in the geostatistical modeling.

Introduction

Estimation and simulation are popular methods for geostatistical modeling. Most estimation and simulation algorithms require that the variogram be defined for all possible directions and distances. It is conceivable that the separation vector between sample locations and the locations where we desire to perform estimation/simulation could involve a direction and distance for which we do not have a sample variogram value. Therefore, a parametric variogram model must be fit to the experimental variogram points in order to define the variogram for all directions and distances. Additionally, the variogram measure must have the mathematical property of positive definiteness. A positive definite model ensures that the kriging equations can be solved and that the kriging variance is positive (Deutsch, 2002).

Variograms are direction dependent, that is, the spatial continuity of a geologic variable can be different in one direction than another. Typically, the goal is to identify the direction in which spatial continuity is at a maximum, sometimes called the major direction. The spatial continuity in directions perpendicular to the major direction are evaluated and used to define a variogram ellipse. The assumption is that when the variogram is defined along the axes of this ellipse, the variogram can be determined for any vector by combining the axes variograms.

There are two common methods for determining the orientations of these ellipse axes. The first is to calculate experimental directional variograms for many different directions. Comparison of the directional variograms identifies the major directions as those directions with the greatest spatial continuity. The second method is the calculation of a variogram volume (3D) or variogram map (2D). This method considers all the experimentally available separation vectors and reports the variogram for each. Visualizing the variogram map or slices through the centroid of a variogram volume allow the major directions to be determined. A method for determining these directions automatically using the moment of inertia has been proposed (Hassanpour & Deutsch, 2008).

Once the major directions have been determined, a variogram model is fit to these directional variograms simultaneously. If the second method from above for determining the major directions was used, the experimental directional variograms must first be calculated. There are algorithms available for fitting the variogram model automatically (Larrondo et al., 2003; Cressie, 1985; Genton, 1998). The fitting

can also be performed manually. In either case, the fitting is done by determining the variogram model which best fits the experimental points calculated along the major directions.

This work proposes a method for fitting a variogram model to a variogram volume. Instead of fitting the variogram model to the select major directions, the variogram model is fit to all informed directions and distances simultaneously. The moment of inertia method is used to identify an initial guess of the directions and magnitudes of the anisotropy. The variogram fitting algorithm proposed by Larrondo and Deutsch (2003) is then used to fine tune the variogram model parameters with the objective of getting a best fit to the variogram volume. Not every value in the variogram volume is used for the fitting, select lags are censored from the fitting based on how well informed the lag is. A new GSLIB-style program, `varmap_varfit`, is documented which performs this fitting.

Variogram Volume

Isaaks and Srivastava (1989) were among the earliest to propose the use of a variogram volume for determining the major directions. The variogram volume is the logical limit of the idea of calculating the variogram in a number of distances and directions. A rectangular coordinate system is defined and the variogram value is calculated for all points separated by the vector $\mathbf{h} = (h_x, h_y, h_z)$. All pairs whose separation in the x direction is $h_x \pm \Delta x$, in the y direction is $h_y \pm \Delta y$ and in the z direction is $h_z \pm \Delta z$ are grouped together. The variogram values are then posted on a map where the center of the map is the lag distance of zero. The main use of the variogram map is to detect the major and minor directions of continuity. Typically the slices which pass through the center of the volume are posted on a map in each of the three orientations. These maps will point toward the major and minor directions when the major and minor directions are not parallel to any coordinate axes. The variogram volume will be very noisy and of little use in the presence of sparse data.

Variogram Volume Censoring

It is common for some lags in a variogram volume to be poorly informed or uninformed. This is due to the data configuration. The poorly informed lags can be detrimental to the fitting of the variogram as they typically introduce noise and indicate less continuity than is truly present. To avoid these issues, a censoring of the variogram volume is proposed. The censoring identifies the poorly informed lags. These lags are not used in the variogram fitting.

The quality of a variogram lag is assumed to be proportional to the number of pairs that inform that lag. As such, lags are censored based on the number of pairs informing the lag. Any lags with a number of pairs less than some cutoff are not considered. To determine the cutoff, the distribution of the number of pairs in the variogram volume is considered. A typical distribution for the number of pairs in a variogram volume is shown in Figure 1. Most of the lags in this variogram volume are uninformed, that is there are no pairs informing that lag. A cutoff could be arrived at by calculating the 20th percentile in the number of pairs distribution without considering uninformed lags. This is illustrated in Figure 2 for the number of pairs distribution shown in Figure 1 with the zeros removed. The 20th percentile of the distribution corresponds to a number of pairs of approximately 250. Any lags with less pairs than this would not be included in the variogram fitting. This censoring logic is implemented in the program `varmap_varfit`.

Variogram Volume Calculation and Moment of Inertia

The variogram volume calculations are performed first. This informs the variogram value at each lag. The user inputs the number and size of lags in each direction. The program scans over all the data, calculating the variogram between all pairs which fall within the specified lags. The center of the variogram volume corresponds to a distance of zero. The first thing that the `varmap_varfit` program does is calculates the chosen variogram measure for all informed lags.

Variogram volumes are commonly used to identify directions of anisotropy. This procedure has been automated by Hassanpour and Deutsch (2008) by the application of the moment of inertia to the variogram volume. Applying the moment of inertia to the variogram volume identifies the directions of anisotropy.

Variogram Fitting

The next operation that the `varmap_varfit` program performs is the variogram fitting. The azimuth, dip, and distance of each lag are known as well as the principal directions of anisotropy. The variogram fitting algorithm uses this information to determine a best-fit model. The fitting proceeds by starting with an initial guess for the variogram model. The user specifies the number of structures. This initial guess is perturbed by simulated annealing. A change to the model is suggested. If the fit is improved by the change, it may be kept. If the fit is not improved, the change may be rejected. The probability of an improvement being rejected or a non-improvement being accepted starts higher and decreases as the algorithm proceeds. Many changes are suggested and evaluated such that the variogram model converges towards a best fit. The fit is measured by comparing the variogram model value to the experimental value at each lag. The measure used is the mean squared error between the model and the experimental points. The variogram fitting aims to minimize the mean squared error. Once the error is minimized or a specified number of iterations have been performed, the algorithm terminates and outputs the final variogram model. To check the fit of the variogram model to the experimental points, the user can specify a number of directions they wish to check. The user also specifies direction tolerances. Any lags in the variogram volume which fall in the specified directions within the specified tolerances are plotted on the same plot as the final variogram model in that direction. This allows the user to observe the quality of the fit.

Parameter File

The parameter file for `varmap_varfit` is shown in Figure 3. It is a combination of the parameters for the GSLIB `varmap` program and the `varfit` program. Lines 5 through 17 are the same as for the `varmap` program described in Deutsch and Journel (1998). The remainder of the lines in the parameter file relate to the variogram fitting. Lines 20 through 23 are the same as for the `varfit` program described by Larrondo et al. (2003). Line 24 specifies the proportion of lags to censor. Lines 25 and 26 specify output files for the variogram model and postscript plot respectively. Lines 27 through 32 relate to the postscript plot. Line 27 specifies the maximum distance to plot on the x-axis of the variogram plots. Line 28 specifies the number of directions to plot. The directions are plotted three directions to a page in the postscript file. Line 29 specifies the azimuths for all the directions of interest. Line 30 specifies the dip for each direction. Line 31 specifies the azimuth and dip tolerances. Line 32 specifies the title for the variogram plots. The section of ADVANCED OPTIONS parameters is identical to Larrondo et al. (2003).

Conclusions

Determining an appropriate variogram model is an important step in most geostatistical modeling workflows. This is typically done by calculating many experimental directional variograms and studying and comparing these directional variograms to determine the principal directions of anisotropy. Then a model is fit to the experimental variograms which gives a best fit. Variogram volumes are often used to assist the practitioner in determining the principal directions. However, the middle step of calculating directional variograms is not necessary. A variogram model could be fit directly to the variogram volume.

References

- Cressie N, 1985. Fitting variogram models by weighted least squares. *Mathematical Geology*, Vol 17, No 5, 563-586.
- Deutsch CV & Journel AG, 1998. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press. New York. 369p.
- Deutsch CV, 2002. *Geostatistical Reservoir Modeling*. Oxford University Press. New York. 376p.
- Genton MG, 1998. Variogram fitting by generalized least squares using an explicit formula for the covariance structure. *Mathematical Geology*, Vol 30, No 4, 323-345
- Hassanpour R & Deutsch CV, 2008. Fitting local anisotropy with mass moments of inertia. *Center for Computational Geostatistics Annual Report 10*. University of Alberta. 9p.
- Isaaks EH & Srivastava RM, 1989. *An Introduction to Applied Geostatistics*. Oxford University Press. New York. 561p.
- Larrondo PF, Neufeld CT, & Deutsch CV, 2003. *VARFIT: A program for semi-automatic variogram modeling*. *Center for Computational Geostatistics Annual Report 5*. University of Alberta. 17p.

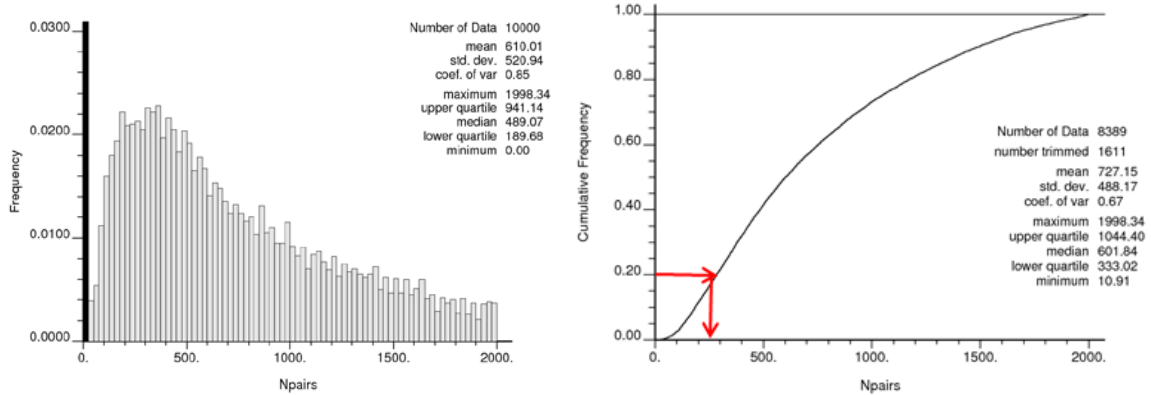


Figure 1 (left): A typical distribution of the number of pairs in a variogram volume. Figure 2 (right): Cumulative distribution of number of pairs with 1611 zeros removed. 20th percentile corresponds to approximately 250 pairs.

```

1          Parameters for VARMAP_VARFIT
2          *****
3
4  START OF VARMAP PARAMETERS:
5  hier3.dat          -file with data
6  1 3                - number of variables: column numbers
7  -1.0e21 1.0e21    - trimming limits
8  0
9  250 250 1         -l=regular grid, 0=scattered values
10 1.0 1.0 1.0      -if =1: nx, ny, nz
11 1 2 0            - xsiz, ysiz, zsiz
12 varmap.out        -if =0: columns for x,y, z coordinates
13 2 10 0           -file for variogram output
14 100.0 15.0 5.0  -nxlag, nylag, nzlag
15 0                -dxlag, dylag, dzlag
16 1                -minimum number of pairs
17 1 1 10 1        -standardize sill? (0=no, 1=yes)
18                -tail, head, variogram type (see below)
19 START OF VARFIT PARAMETERS:
20 1                -number of nested structures
21 1                -constant angle between structures
22 0                -inverse distance weighting (0=no, 1=yes)
23 0                -number of pairs weighting (0=no, 1=yes)
24 20              -percentage of points to drop
25 varfit.var       -file for variogram model
26 varfit.ps        -file for PostScript output
27 300              -maximum distance to plot
28 4                -number of directions to plot
29 0 75 90 165     -azimuth directions to plot
30 0 0 0 0         -dip directions to plot
31 10 10           -azimuth and dip tolerances
32 3D Semivariogram -project title
33
34 START OF ADVANCED OPTIONS:
35 0 0 0           -zonal Anis: Hmax, Hmin, Vert (0=no, 1=yes)
36 0 0 0           -cyclicality: Hmax, Hmin, Vert (0=no, 1=yes)
37 1 1.0          -fix sill
38 1 0.0          -fix nugget effect
39 0              -number of structure types to fix
40 0              -number of Hmax ranges to fix
41 0              -number of Hmin ranges to fix
42 0              -number of Vert ranges to fix
43 0              -number of azimuth angles to fix
44 0              -number of dip angles to fix
45 0              -number of plunge angles to fix
46 0              -number of variogram preferences
47 0 10.0         -fix Hmax/Vert anis. (0=no, 1=yes)
48 0 1.0         -fix Hmin/Hmax anis. (0=no, 1=yes)
49
50 Insert the required lines (from below) for the options chosen
51 above (must be placed below the option flag). Some parameters
52 can be fixed for each structure (range) while
53 others (hmax, hmin, hvert, anisotropy) are fixed for all structures
54 1 2            - structure number and structure type
55 1 1000        - structure number and Hmax range
56 1 200         - structure number and Hmin range
57 1 10         - structure number and Vert range
58 1 45         - structure number and azimuth angle
59 1 -30        - structure number and dip angle
60 1 30         - structure number and plunge angle
61 1 5          - variogram number, and scale factor for preference
62
63
64 type 1 = traditional semivariogram
65       2 = traditional cross semivariogram
66       3 = covariance
67       4 = correlogram
68       5 = general relative semivariogram
69       6 = pairwise relative semivariogram
70       7 = semivariogram of logarithms
71       8 = semimadogram
72       9 = indicator semivariogram - continuous
73      10 = indicator semivariogram - categorical

```

Figure 3: Parameter file for the varmap_varfit program.