Chapter 41
Mathematical Geology by Example: Teaching and Learning Perspectives

James R. Carr

Abstract Numerical examples and visualizations are presented herein as teaching aids for multivariate data analysis, spatial estimation using kriging and inverse distance methods, and the variogram as a standalone data analytical tool. Attention is focused on the practical application of these methods.

41.1 Introduction

An oxymoron. Mathematical geology has been characterized as such. Saying so, though, betrays ignorance, not of mathematics, but of geology. The science is inherently numerical. Minerals, for example, are quantifiable based on specific gravity, hardness, Miller index, and abundance. Rock classification in petrology and petrography is inherently dependent upon mineral frequency, determined in a manner identical to that which is used by the hematologist when classifying specimens of blood. Geologic structures are quantified by strike and dip, even abundance when characterizing the integrity of rock masses. Economic geologists and geochemists develop complex databases of samples, each associated with many elements, the analysis of which provides clues to ore genesis, water origin, environmental stresses, and rock classification, to name but a few applications. Geophysics and remote sensing provide enormous sets of numbers visualized as digital images. Far from being an oxymoron, mathematical geology is broadly defined as the application of theoretical and applied mathematics to the assessment of geologic data to aid in the interpretation of earth evolution.

The word, aid, is not chosen carelessly. No equation, no calculator, no computer, can substitute for the human ability to infer and interpret. Where equations, calculators, and computers can help with geologic interpretation is in the conversion of numbers to pictures, such as the case when converting numbers comprising a digital
image into what mimics a photograph on a computer screen. Scientific visualization is the process of converting numerical information of any kind into a picture, hopefully improving its interpretation. The responsibility of interpretation remains, always, with the human analyst.

There tends to be an element of distrust of numbers. The quote, attributed to Benjamin Disraeli, is well known, “There are three kinds of lies: lies, damned lies, and statistics.” Apparently, there is uncertainty regarding whether Disraeli actually made this quote. This saying was, however, widely used by the end of the 19th Century. Mark Twain, for example, writing in 1906: “Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: ‘There are three kinds of lies: lies, damned lies, and statistics.’”—*Autobiography of Mark Twain*.

Perhaps mistrust of numbers is not as accurate as saying that there exists a reverence of numbers due to a fundamental insecurity about mathematical understanding. The presentation of a statistical analysis can be quite intimidating to those whose confidence in understanding the analytical methods is weak. Of course, the weak confidence can be taken advantage of by those less scrupulous, stating interpretations of numbers for which there is no clear justification. Thus the skepticism surrounding statistics—lies worse than damned lies.

Despite this ignorance, statistical analysis of data is the most widely applied mathematical method in the geological sciences. Geologists draw maps, with geostatistics, geographic information systems (GIS), and remote sensing fundamentally contributing to the process. Mine *geologists* are increasingly charged with ore reserve estimation and ore control using geostatistics. Other examples of applied statistics included bivariate and multivariate methods important for understanding the correlation between two or more variables. Other numerical methods of importance to geologic understanding are finite difference modeling for understanding ground water flow, geostatistical simulation for modeling uncertainty of spatial data, time-series (Fourier) analysis for identifying cycles in data strings over time or space, linear algebra for modeling landform and geologic structure morphology, fractal geometry for understanding scaling in geologic processes, and the application of neural networks to the modeling of geologic processes.

Some of these applications have proven less interesting to students of mathematical geology than others. Three and a half decades of teaching applied mathematics to earth scientists and engineers at a hardrock mining school provide the backdrop for the following observations. One, *graduate* students of economic geology, moreover economic geology professionals eagerly seek instruction and advice in multivariate methods applied to rock geochemistry data, with an emphasis on interpretation for a better geologic understanding of ore deposits. These students and professionals typically want a pure course on multivariate data analysis. Two, teaching kriging theory in an *undergraduate* course is a waste of time when the heavily parametric practice of spatial estimation is considered; industry often views universities as workshops for training mine geologists and engineers on the use of a particular choice of mine planning software, such as SURPAC, and teaching how to use the software and what choices to make for parameter definition is more than one
course can provide. These students are less interested in kriging theory than they are in how to interpret a variogram, how to design a grid, what type of estimator to use, and more. The challenge in this case is how to answer these and other questions while not overstaying one’s welcome when explaining theory. Thirdly, the variogram is popular among students and professionals as an analytical tool when kriging is not the primary goal; students and practitioners of remote sensing, in particular, use the variogram for various types of digital image processing.

Multivariate data analysis, the practice of kriging, and the variogram as a stand-alone data analytical tool are presented in this chapter with an emphasis on their teaching. Both teacher and student perspectives are presented to balance the discussion between tips for learning and advice for teaching.

### 41.2 Multivariate Analysis of Geochemical Data

During the 1930s decade, psychologists began to apply principal components methods to help with the interpretation of their data (e.g., Hotelling 1933; Young 1937). Many psychologists collect data on patients characterizing their behavioral traits. Principal components methods allow psychologists to group patients of similar behaviors resulting in a better understanding of them. Three decades later, sedimentologists (e.g., Imbrie and Purdy 1962; Klovan 1966) used principal components analysis to group samples of sediment based on sedimentological characteristics. In this case, the sediment sample is analogous to the patient and the sample characteristics are analogous to behavioral traits. How sediment samples group can be an indication of sediment source, depositional environment, composition, or some other condition of importance to geologic interpretation. A collection of papers published in 1983 (Howarth 1983) reviewed the application of multivariate analysis to geochemical prospecting. Tomes written on geochemistry (e.g. Albarède 1995) often discuss multivariate analysis applied to the interpretation of geochemical data.

Many mathematical methods have been developed to help with the analysis of multivariate data. An important goal of each of these methods is a reduction in the number of variables to enable a more efficient understanding. If there are M original variables, in other words, a smaller number of variables, B, is sought that define a lower multivariate sub-space. Then, the original M dimensional data are projected onto the lower sub-space to yield a plot (graph) that is visually inspected to appreciate data similarities and differences. For students and teachers alike, the ultimate goal of multivariate analysis is the creation of these plots, the study of which motivates subjective conclusions about data associations (Greenacre 1984).

In order to develop the plots, some mutually orthogonal coordinate system is needed. Many of the mathematical methods used to analyze multivariate data involve the reduction of the original data information into some matrix that is eigendecomposed to obtain eigenvalues, each associated with a unique eigenvector. The eigenvectors are mutually orthogonal. Moreover, these eigenvectors define the
lower dimensional sub-space. They are the principal components of data intercorrelation information.

Can multivariate data analysis be taught without explaining, or at least reviewing, eigendecomposition of data? Answering yes leads to a teaching approach that treats the multivariate analytical algorithm as a black box. This approach relieves teachers of the chore of explaining a method that many students abhor. Undergraduate students, and even graduate students in some cases, are likely to skip class when eigenvalues and eigenvectors are to be discussed. Modern students are quick to dismiss that which they do not like, or find boring. For example, based on the experience of co-teaching mineralogy for the past five years, a discussion of crystallography often results in a rather empty classroom. A mental laziness is betrayed by students’ behavior in this regard. It frustrates teachers wanting students to achieve an understanding of analytical methods deeper than the data in–data out black box.

Of course, answering no to the foregoing question and teaching multivariate data analysis outside the black box is confounded by the same student attitudes. Their learning cannot be forced. It can, however, be enticed by numerical examples that are straight-forward, explained in class, and reinforced by extracurricular calculations. Students can be shown that an understanding deeper than black box mysticism is relatively easy to achieve. What follows is a demonstration of this concept and is intended as an aid to instruction. Student understanding can be assessed by substituting the starting data table with alternative data.

### 41.2.1 Numerical Insight to Multivariate Data Analysis

Geochemical data from seven rock samples, each characterized by five elements (variables), are presented in the following table:

| Sample | Gold (Au) | Silver (Ag) | Copper (Cu) | Lead (Pb) | Zinc (Zn) |
|--------|-----------|-------------|-------------|-----------|-----------|
| 1      | 15        | 0.4         | 21          | 21        | 15        |
| 2      | 25        | 0.3         | 14          | 15        | 3         |
| 3      | 19        | 0.5         | 12          | 19        | 4         |
| 4      | 37        | 0.5         | 24          | 17        | 7         |
| 5      | 33        | 0.3         | 14          | 13        | 5         |
| 6      | 12        | 0.4         | 21          | 29        | 5         |
| 7      | 12        | 0.4         | 13          | 19        | 5         |

*Note* values for Au, Ag, Cu, Pb, and Zn are in ppm

These data represent a five-dimensional variable space. The goal is to determine the eigenvectors for these data, the sub-space that will be used for plotting. A theorem presented by Eckart and Young (1936) holds that any real valued data matrix can be represented as the following product, \[ \text{[data]} = \text{[R-mode eigenvectors]} \cdot \text{[eigenvalues]} \cdot \text{[transposed Q-mode eigenvectors]} \]. In this case, Q-mode multivariate
analysis is focused on the relationships among samples. R-mode multivariate analysis is one that focuses on the relationships among the variables.

To obtain the Q-mode result, the data matrix is multiplied by its transpose in the following order, \([\text{transposed data matrix}] \cdot [\text{data matrix}]\), to yield a square matrix as follows: \([\text{data}]^T \cdot [\text{data}] = [\text{Q-mode}]\):

\[
\begin{pmatrix}
15 & 25 & 19 & 37 & 33 & 12 & 12 \\
0.4 & 0.3 & 0.5 & 0.5 & 0.3 & 0.4 & 0.4 \\
21 & 14 & 12 & 24 & 14 & 21 & 13 \\
21 & 15 & 19 & 17 & 13 & 29 & 19 \\
15 & 3 & 4 & 7 & 5 & 5 & 5
\end{pmatrix}
\]

\[
\begin{pmatrix}
15 & 0.4 & 21 & 21 & 15 \\
25 & 0.3 & 14 & 15 & 3 \\
19 & 0.5 & 12 & 19 & 4 \\
37 & 0.5 & 24 & 17 & 7 \\
33 & 0.3 & 14 & 13 & 5 \\
12 & 0.4 & 21 & 29 & 5 \\
12 & 0.4 & 13 & 19 & 5
\end{pmatrix}
\]

\[
\begin{pmatrix}
3957 & 61 & 2651 & 2685 & 920 \\
61 & 1.16 & 48.4 & 54 & 17.9 \\
2651 & 48.4 & 2163 & 2325 & 813 \\
2685 & 54 & 2325 & 2687 & 860 \\
920 & 17.9 & 813 & 860 & 374
\end{pmatrix}
\]

The result of this multiplication is a square matrix, \(M \times M\) in size, and \(M\) is the number of original variables, 5 in this case. This square matrix is the one from which eigenvalues and Q-mode eigenvectors are obtained [software is necessary for eigendecomposition, ironically rendering this step as a black box]:

\[
\text{Eigenvalues: } \begin{pmatrix} 91.8 & 0 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 \\ 0 & 0 & 10.1 & 0 & 0 \\ 0 & 0 & 0 & 4.8 & 0 \\ 0 & 0 & 0 & 0 & 0.2 \end{pmatrix}
\]

\[
\text{Eigenvectors: } \begin{pmatrix} 0.655 & -0.73 & -0.10 & -0.17 & -0.007 \\ 0.011 & 0.005 & -0.01 & -0.17 & 1.0000 \\ 0.500 & 0.228 & 0.336 & 0.765 & -0.008 \\ 0.536 & 0.618 & -0.47 & -0.33 & -0.017 \\ 0.183 & 0.181 & 0.809 & -0.53 & -0.007 \end{pmatrix}
\]

The eigenvectors are loaded column-wise. The eigenvalues are loaded into a matrix along the diagonal. All off-diagonal entries in the eigenvalue matrix are zero. These eigenvalues are actually the square roots of those computed directly from the R-mode matrix because the original data matrix is squared when multiplied by its transpose. By performing this multiplication to yield a square, symmetrical matrix,
the eigenvectors from which are guaranteed to be orthogonal to one another. For
example, if the first two eigenvectors are multiplied together, the result should be
precisely zero:

$$3.59 \times -4.04 + 0.063 \times 0.028 + 2.74 \times 1.26 + 2.94 \times 3.42 + 1 \times 1 = 0.005 \approx 0$$

and the result would be precise if not for round-off error.

Working toward the goal of plotting samples 1 through 7, a first step involves
multiplying the eigenvector matrix to the eigenvalue matrix:

$$\begin{pmatrix}
0.655 & -0.73 & -0.10 & -0.17 & -0.007 \\
0.011 & 0.005 & -0.01 & -0.17 & 1.0000 \\
0.500 & 0.228 & 0.336 & 0.765 & -0.008 \\
0.536 & 0.618 & -0.47 & -0.33 & -0.017 \\
0.183 & 0.181 & 0.809 & -0.53 & -0.007
\end{pmatrix}$$

$$\begin{pmatrix}
91.8 & 0 & 0 & 0 & 0 \\
0 & 25 & 0 & 0 & 0 \\
0 & 0 & 10.1 & 0 & 0 \\
0 & 0 & 0 & 4.8 & 0 \\
0 & 0 & 0 & 0 & 0.2
\end{pmatrix} = \begin{pmatrix}
60.13 & -18.25 & -1.01 & -0.82 & 0.00 \\
1.01 & 0.13 & -0.10 & -0.82 & 0.20 \\
45.90 & 5.70 & 3.39 & 3.67 & 0.00 \\
49.20 & 15.45 & -4.75 & -1.58 & 0.00 \\
16.80 & 4.53 & 8.17 & -2.54 & 0.00
\end{pmatrix}$$

The next step involves multiplying this resultant matrix by the original data
matrix. Because the fifth column of this matrix represents values of practically zero,
only the first four columns are used to obtain four factors:

$$\begin{pmatrix}
15 & 0.4 & 21 & 21 & 15 \\
25 & 0.3 & 14 & 15 & 3 \\
19 & 0.5 & 12 & 19 & 4 \\
37 & 0.5 & 24 & 17 & 7 \\
33 & 0.3 & 14 & 13 & 5 \\
12 & 0.4 & 21 & 29 & 5 \\
12 & 0.4 & 13 & 19 & 5
\end{pmatrix} \begin{pmatrix}
60.13 & -18.25 & -1.01 & -0.82 \\
1.01 & 0.13 & -0.10 & -0.82 \\
45.90 & 5.70 & 3.39 & 3.67 \\
49.20 & 15.45 & -4.75 & -1.58 \\
16.80 & 4.53 & 8.17 & -2.54
\end{pmatrix} =$$

|     | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|-----|----------|----------|----------|----------|
| 1   | 3151.00  | 238.40   | 78.80    | -6.84    |
| 2   | 2935.00  | -131.10  | -24.56   | -0.69    |
| 3   | 2696.00  | 33.38    | -36.13   | -12.13   |
| 4   | 4281.00  | -244.00  | 20.38    | 12.69    |
| 5   | 3351.00  | -298.90  | -6.80    | -9.17    |
| 6   | 3197.00  | 371.50   | -37.87   | 8.38     |
| 7   | 2337.00  | 171.40   | -17.49   | -5.18    |
The word, factor, heading each column is one of the new variables within the sub-space of the original data matrix. The numbers to the left are the sample numbers used in the original data able. The factors represent an orthogonal coordinate system to enable plotting these seven samples to determine their relationship to one another.

Relative significance of each factor with respect to the total data information is determined by summing the eigenvalues, then dividing each eigenvalue by this sum to obtain a proportion. The five eigenvalues sum to 131.9. Factor 1, for instance, represents $100 \times \left( \frac{91.8}{131.9} \right) = 70\%$ of the original data information content. The second factor associated with an eigenvalue equal to 25, incorporates 20% of the original data information content. If the seven samples are plotted using the first two factors, then the resultant plot represents 90% of the original data information content. This plot is shown in Fig. 41.1.

Notice that samples 2, 4, and 5 plot in the negative region with respect to Factor 2. These three samples are associated with the highest gold values. But, these samples are among the lowest for silver, lead, and zinc. Samples 1 and 6 are much higher in lead and zinc, but much lower for gold. Each factor is a function of all five of the data variables, Au, Ag, Cu, Pb, and Zn. For example, in the foregoing matrix multiplication involving the original data matrix, the Factor 1 “coordinate” for sample 1 is equal to:

$$15 \times 60.13 + 0.4 \times 1.01 + 21 \times 45.9 + 15 \times 49.2 + 15 \times 16.8 = 3151.0.$$ 

In reviewing this calculation, notice that it is:

$$\text{Au value} \times 60.13 + \text{Ag value} \times 1.01 + \text{Cu value} \times 45.9 + \text{Pb value} \times 49.2 + \text{Zn value} \times 16.8.$$ 

Literally, the coordinate of a sample in any of the factors is a function of all the original variables, not just any one, or two. Because of this, the way samples plot in

![Fig. 41.1](image-url)
Fig. 41.1 reflects their similarity over all variables. A grouping of samples, then, suggests a rock-chemistry similarity that likely has importance to the interpretation of ore genesis.

To obtain the R-mode result, the original data matrix is once again squared, but in a different order of multiplication: \([\text{data}][\text{data}]^T = \text{[R-Mode]}\). Following the same steps above for the Q-mode result, resulting factors for plotting of the five variables are:

|     | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 |
|-----|----------|----------|----------|----------|----------|
| Au  | 5542.00  | −463.80  | −7.27    | 2.41     | −0.28    |
| Ag  | 96.58    | 2.98     | −0.54    | 0.37     | −0.03    |
| Cu  | 4220.00  | 136.50   | 35.76    | −18.29   | −0.15    |
| Pb  | 4527.00  | 378.80   | −46.61   | 6.99     | −0.13    |
| Zn  | 1538.00  | 111.00   | 82.77    | 11.84    | −0.09    |

The relative importance of each factor with respect to original data information content is the same as for the Q-mode result because the eigenvalues are identical. Figure 41.2 presents a plot based on the first two factors.

Figure 41.2 suggests that gold (Au) is not closely associated with any one of the four other variables. Focusing only on factor 1, gold (Au) and silver (Ag) are on opposite sides of the horizontal axis. Often, variables plotting as such are inversely related; when one is higher in value, the other is lower in value. Further with respect to factor 1, zinc (Zn) is closer to silver and lead (Pb) and copper (Cu) are closer to gold. If, however, the focus is solely on factor 2, then gold and lead appear to be inversely related.

Software is necessary for larger data sets. Using this example and challenging students to follow it for data sets other than that which is used will not necessarily guarantee a deep understanding. But, when reviewing the output from multivariate software, students will have a general understanding of what happens to the input data and the jargon inherent to the method. Knowing why eigenvectors (factors) are
used for developing plots gives students greater confidence when interpreting these plots.

An actual multivariate data set consisting of about 1000 samples, each associated with 50 elemental variables, is analyzed using the multivariate method known as correspondence analysis (Benzecri 1973). This multivariate method is the one preferred by the author for actual data analysis, but its mathematical presentation is not as straightforward as that presented above for principal components analysis. It is the opinion of the author that correspondence analysis yields plots that separate the data better than other methods. The result is shown in Fig. 41.3 for variables only (to reduce the clutter of the plot).

How elements are related is interpreted from Fig. 41.3. Manganese (Mn) is polarizing with all other elements plotting away from it along factor 1. Rocks higher in manganese are inferred to be much lower in the other elements. Given that the likely manganese mineral in this deposit is MnO (wad), moreover knowing that this mineral is black and sooty, could be useful knowledge in the field of where ore is, or is not, present. Factor 2 separates barium (Ba) from the precious metals. The likely barium mineral is barite, an easily recognized mineral in the field if crystalline. This element, too, may be useful for the approximate delineation of the ore zone in the field based on visual inspection.

**Fig. 41.3** An R-mode plot of a multivariate geochemical set of data characterizing an ore deposit. The relative importance of each factor is indicated in the axis label. This plot was created by software that is presented in Carr (2002)
41.3 Geostatistics and Its Myriad Parameters

Decisions. A teacher of geostatistical estimation can spend weeks teaching geostatistical theory, broadly so by including polygonal and inverse distance strategies in addition to kriging. Weeks! Then faced with teaching the practice of estimation. The theory is complex, particularly in the case of kriging. And, yet, the outcome is highly vulnerable to the parameters selected for implementing theory. Figure 41.4, whereas not intended to be comprehensive, presents many of the decisions that must be made by a geostatistician when practicing the gridding of data.

A teacher can spend more or less time on geostatistical theory, lesser for undergraduates perhaps. Time, however, must still be devoted to explaining about and advising on the parameters that are necessary to estimation. Moreover, the

| A Noncomprehensive Table of Parameters |
|----------------------------------------|
| **Method of Gridding:**                |
| Polygonal                              |
| Inverse Distance Methods (ID)          |
| Kriging                                |
| No Data Transform                      |
| With Data Transform                    |
| Simple                                 |
| Log-normal; with or without bias adjustment |
| Ordinary                               |
| Indicator (true indicator using quantiles as cutoffs) |
| Support, Block or Punctual             |
| Variogram Parameters:                 |
| Type of model                          |
| Nugget, Sill and Range                 |
| **ID:**                                |
| Power term                             |
| **Data:**                              |
| Transform options                      |
| None                                   |
| Indicator                              |
| Natural logarithms                     |
| Hermitian                              |
| Capping                                |
| **Gridding:**                          |
| Grid dimensions                        |
| Number of rows                         |
| Number of columns                      |
| Number of levels (if three dimensional) |
| Spacing distance between rows, columns, and levels |
| Number, N, of nearest neighboring sample locations used for estimation |
| Search window radius for locating N nearest neighbors |
| Search window strategy: general, quadrant, octant |
| Anisotropy considerations: Major axis direction, plunge direction, dip, ratio of major axis/minor axis |

Fig. 41.4 Gridding a set of spatial data requires selecting the estimation algorithm for gridding, then defining parameters unique to the estimation algorithm. How to treat the data, raw or transformed, is another important decision. Likewise, the geometry of the grid must be designed
influence of one or more of these parameters is best appreciated by visualizing estimation outcomes.

Estimation outcomes are visualized as color contour maps in the following demonstration. A collection of 2,500 mercury values were collected to assess the severity of site contamination after a flood. The variogram for these data is shown in Fig. 41.5.

The shape of the variogram in Fig. 41.5 is modeled using a spherical variogram model. This variogram shape is the most commonly observed for spatial data, regardless of the spatial phenomenon under study. The model is explicitly defined by setting the parameters for the nugget (found by extrapolating the calculated variogram backwards to intersect the y-axis at h = 0), the sill, that value of the variogram that is more or less constant once the range (of spatial correlation) is reached. In this example, these parameters are: nugget = 20 (rounded), sill = 117 (rounded), and the range is 90 m. Other parameters used in the following data visualization demonstration are as follows: (1) no data transform; (2) block support; (3) ordinary kriging; (4) general, isotropic search strategy with a radius equal to one-half the variogram range; (5) up to N = 10 nearest neighboring samples used for estimation; (6) inverse distance (power term = 1) and inverse distance squared (power term = 2) weighting presented for comparison to the kriging outcome; (7) grid parameters: 50 rows, each with 50 columns, spacing between rows and columns is 10 m. Outcomes are presented in Figs. 41.6 and 41.7.

A lower nugget value is seen to yield lesser smoothing during estimation (Fig. 41.6). A larger nugget yields more smoothing. With inverse distance methods, the larger the power term is, the less smoothing that results during estimation. The aesthetic appeal of a map is a subjective assessment. The amount of smoothing controls the complexity of the map. If larger scale aspects of a spatial region are of more interest than smaller scale aspects, then more smoothing should be used during estimation to downplay the smaller scales. On the other hand, if the desire is to visualize spatial variability down to the smallest possible scale that is allowed by the data, no to minimal smoothing should be used during estimation.

![Variogram](image)

**Fig. 41.5** Variogram for 2,500 mercury values. The jagged line is the actual calculation outcome. The smooth, continuous line is a model fit to the calculation outcome. The model, in this case, is the spherical variogram model and its parameters, nugget, sill, and range, are listed above the variogram.
Indeed, there are similarities among these maps. Each map shows a zone of higher mercury values in the center, and two low zones at the left-center and top-center. These regions are associated with a higher density of spatial samples. Regions of the map that change appreciably when estimation parameters are changed are more sparsely sampled. The spatial distribution of mercury samples is shown in Fig. 41.8.

Differences among the contour map outcomes are noteworthy for spatial locations associated with sparser sampling. Moreover, these differences are more easily observed when increased smoothing is used during estimation.

**Fig. 41.6** Three visualizations from kriging. Top, left map is based on the variogram parameters, nugget, sill, and range, that are listed in Fig. 41.5. Top, right map is based on the same variogram parameters, except the nugget value is set equal to zero. The bottom, left map is based on the nugget value set equal to the sill value; in this case, the outcome is a simple average estimation. Integer labels are used for the contour lines to indicate relative value from smaller, 1, to larger, 10. Color also indicates relative value from lower, blue, to higher, red.
41.4 The Variogram as a Stand-Alone Data Analytical Tool

Kriging is not necessarily the ultimate goal of geostatistical analysis. The variogram as a stand-alone data analytical tool has a variety of uses that are independent of estimation. Examples are many and include noise isolation, texture classification of digital images, and self-affine fractal analysis and modeling. The concept of digital image texture is chosen for demonstration.
Four textures are illustrated in Fig. 41.9: water, playa, alluvium, and sedimentary rock outcrops.

Water and playa textures are similar, differing only in reflectivity. Variograms of these textures are likewise similar and indicate a predominant spatial randomness.

Fig. 41.9  The center mosaic shows four textures extracted from a Landsat TM image, clockwise from top, left: alluvium, playa, water, and sedimentary rock outcrops. Variograms for these textures are likewise arranged.
with little underlying signal. The variogram of alluvium texture indicates an underlying spatial structure that is heavily masked by noise (randomness). Unlike these other textures, sedimentary rock layer texture is predominantly signal. The variogram of this texture reveals the strong spatial structure and very low noise.

Digital image classification is a process that depends on automatic identification of classes, features on the ground, based on some form of signature, or characteristic for each class. The histogram of pixel values is one such signature that is often used when basing classification on pixel value. The variogram is a signature that is useful for classifying the texture of ground classes. The foregoing demonstration shows that variograms do differ for ground classes, but in ways that are not directly relatable to pixel values. Playa and water, for instance, are distinctly different in brightness, yet their variograms are similar in shape. The variogram has been used with considerable success for the classification of microwave images (e.g., Carr and Miranda 1998; Miranda et al. 1998). These images are inherently noisy due to microwave frequency additions and cancelations that impart what is known as speckle. The classification of texture using variogram signatures applied to less noisy images, such as those from the Landsat satellite, has not been extensively tested.

In the foregoing example, the images of alluvium, playa, water, and sedimentary rock are 100 × 100 pixel extracts from a band 3 (visible red) Landsat 7 image of

**Fig. 41.10** A band 3 (visible red) extract from a complete Landsat 7 scene, Path 39, Row 35, acquired on September 25, 2000
The variogram signatures shown in Fig. 41.9 were applied to this image for the classification of texture. The understanding of what constitutes texture in a digital image takes some time to develop. Texture is not brightness, per se, but rather the unique patterns exhibited by groups of pixels. The outcome of textural classification is shown in Fig. 41.11.

The predominant texture seen in Fig. 41.11 is that of alluvium. The texture of water is not unique and is confused with the texture of alluvium. The texture of the shoreline of the lake is identified as playa. This lake (Lake Mead, Clark County, Nevada, U.S.A.) is an artificial reservoir that has a fluctuating water level that leaves an almost pure white calcium carbonate staining on the shoreline. Like playa sediments, the reflectivity of this material often saturates the satellite sensor resulting in identical textures. Sedimentary outcrop texture was often confused with alluvium, and shadows (northwest facing slopes) were often confused with water. Given that this image is of a harsh, arid environment (precipitation is less than 7 cm per year), the predominant alluvium texture makes sense.

Fig. 41.11  Outcome of texture classification based on variograms applied to the Landsat image shown in Fig. 41.10. Colors represent: water (red), playa (green), alluvium (blue), and sedimentary outcrops (yellow)
References

Albarède F (1995) Introduction to geochemical modeling. Cambridge University Press
Benzècri JP (1973) L’analyse des données. Tome 1: La taxinomie. Tome 2: L’analyse des correspondances. Dunod, Paris
Carr JR (2002) Data visualization in the geosciences. Prentice-Hall
Carr JR, Miranda FP (1998) The semivariogram in comparison to the co-occurrence matrix for classification of image texture. IEEE Trans Geosci Rem Sens 36(6):1945–1952
Eckart C, Young G (1936) The approximation of one matrix by another of lower rank. Psychometrika 1:211–218
Greenacre MJ (1984) Theory and applications of correspondence analysis. Academic Press
Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24(417–441):498–520
Howarth RJ (ed) (1983) Statistics and data analysis in geochemical prospecting. Elsevier
Imbrie J, Purdy EG (1962) Classification of modern Bahamian carbonate sediments. In: AAPG classification of carbonate rocks, a symposium, Mem 1, pp 253–272
Klovan JE (1966) The use of factor analysis in determining depositional environments from grain-size distributions. J Sediment Pet 36:115–125
Miranda FP, Fonseca L, Carr JR (1998) Semivariogram textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. Int J Remote Sens 19(3): 549–556
Young G (1937) Matrix approximations and sub space fitting. Psychometrika 2:21–25

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.