Priestley (1981) gives a much more comprehensive account of these random processes. Chapter 3 of that book is especially relevant, and we must leave the reader to pursue the matter there.

5.2.2 Bounded models

In our experience bounded variation is more common than unbounded variation, and the variograms have more varied shapes. In most of these models the variance has a maximum, which is the *a priori* variance of the process, known in geostatistics as the *sill* variance. The variogram may reach its sill at a finite lag distance, the *range*. Alternatively, the variogram may approach its sill asymptotically. In some models the semivariance reaches a maximum, only to decrease again and perhaps fluctuate about its *a priori* variance. These variograms represent second-order stationary processes and so have equivalent covariance functions. They are illustrated in Figures 5.4 and 5.5.

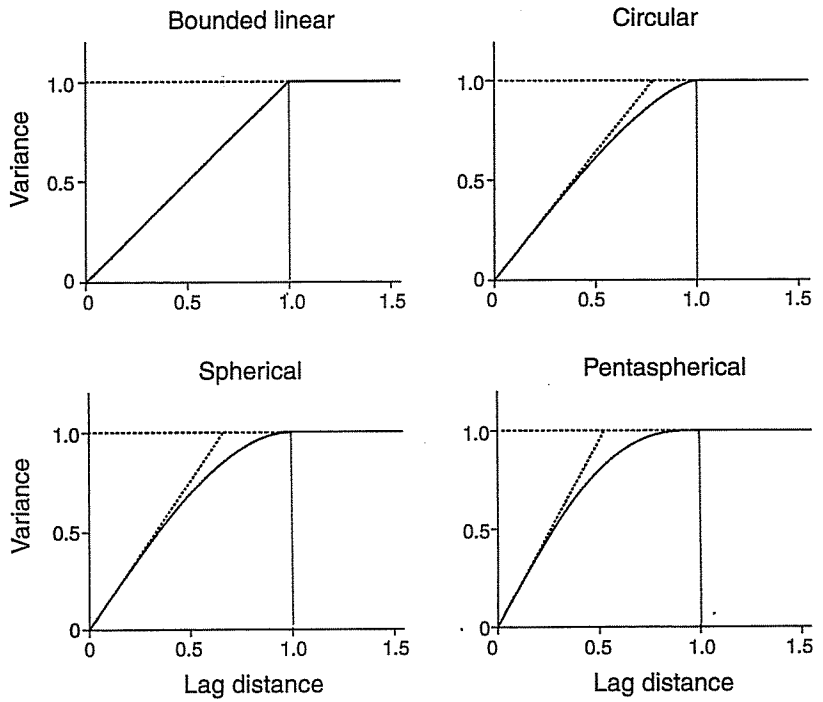


Figure 5.4 Bounded models with fixed ranges: (a) bounded linear; (b) circular; (c) spherical; (d) pentaspherical.

Bounded linear model

The simplest function for describing bounded variation consists of two straight lines, as in Figure 5.4(a). The first increases and the other has a constant variance:

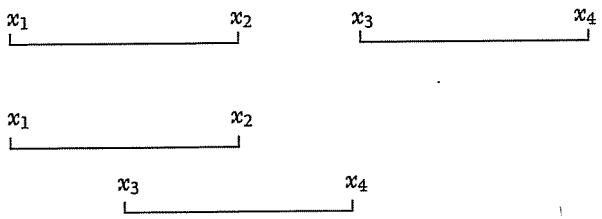
$$\gamma(h) = \begin{cases} c\left(\dfrac{h}{a}\right) & \text{for } h \le a \\ c & \text{for } h > a, \end{cases} \tag{5.13}$$

where c is the sill variance and a is the range. Evidently its slope at the origin is c/a. It is CNSD in one dimension ($\mathbb{R}^1$) only; it may not be used to describe variation in two and three dimensions.

We can derive the variogram for the bounded linear model heuristically as follows. We start with a stationary 'white noise' process, $Y(x)$, in one dimension, i.e. a random process with random variables at all positions along a line but in which there is no spatial dependence or autocorrelation. It has a mean μ and variance σ_Y^2. Suppose that we pass the process through a simple linear filter of finite length a to obtain

$$Z(x) - \mu = \frac{1}{a}\int_x^{x+a} Y(v)\mathrm{d}v. \tag{5.14}$$

Thus, we average $Y(x)$ within the interval a to obtain the corresponding $Z(x)$. Consider now the variable $Z(x)$ derived from two segments of the process $Y(x)$, one from x_1 to x_2 and the other from x_3 to x_4. They may overlap or not, as below.



Evidently, if the two segments do not overlap, as in the upper example, then we should expect their means in $Z(x)$ to be independent. But if they do overlap, as in the lower example, then they will share some of the original white noise series; their means will not be independent, and we should expect some autocorrelation. In general, the closer is x_1 to x_3 (and x_2 to x_4) and the longer is a, the stronger should be the correlation. In fact when x_1 coincides with x_3 (and x_2 with x_4) we should have perfect correlation. The only question is what form the correlation takes as x_3 approaches x_1.

To answer this we consider the discrete analogue of equation (5.14):

$$Z(x+d)-\mu=\lambda_0 Y(x+d)+\lambda_1 Y(x+d+1)+\lambda_2 Y(x+d+2) + \cdots + \lambda_{a-1} Y(x+d+a-1), \tag{5.15}$$

where the $\lambda_0, \lambda_1, \ldots, \lambda_{a-1}$ are weights, here all equal to $1/a$, and $d = 1/2a$ is half the distance between two successive points in the sequence. All more distant members, say $Y(x+d+a-1+b)$, of the series carry zero weight. Suppose that $Y(x)$ is a white noise process; then $Z(x)$ is a moving average process of order $a-1$. Further, if the variance of $Y(x)$ is σ_Y^2 then that of $Z(x)$ is

$$\begin{aligned}\sigma_Z^2 &= \lambda_0^2\sigma_Y^2+\lambda_1^2\sigma_Y^2+\lambda_2^2\sigma_Y^2+\cdots+\lambda_{a-1}^2\sigma_Y^2\\ &= \sigma_Y^2\sum_{i=0}^{a-1}\lambda_i\lambda_i\\ &= \sigma_Y^2/a,\end{aligned} \tag{5.16}$$

which is familiar as the variance of a mean. It is also the covariance at lag 0, $C(0)$. We now want the covariances for the larger lags. These are obtained simply by extension from the above equation:

$$C(h)=\sigma_Y^2\sum_{i=0}^{a-1-h}\lambda_i\lambda_{i+h}=\sigma_Y^2\frac{a-h}{a^2}. \tag{5.17}$$

The covariances are in order, for $h = 0, 1, 2, \ldots, a-1, a$,

$$\frac{a-0}{a^2}\sigma_Y^2,\ \frac{a-1}{a^2}\sigma_Y^2,\ \frac{a-2}{a^2}\sigma_Y^2,\ \ldots,\ \frac{a-a+1}{a^2}\sigma_Y^2,\ \frac{a-a}{a^2}\sigma_Y^2.$$

Dividing through by the $C(0)$ we obtain the autocorrelations, $\rho(h)$, as

$$1,\ (a-1)/a,\ (a-2)/a, \ldots,\ (a-a+1)/a,\ 0.$$

In words, the covariance and autocorrelation functions decay linearly with increasing h until $h = a$, at which point it is 0. Then the autocorrelation coefficient at any h is simply equal to the proportion of the filter that overlaps when the filter is translated by h. The variogram is obtained simply from relation (4.14) by

$$\begin{aligned}\gamma(h) &= C(0)-C(h)\\ &= \sigma_Y^2\frac{a-h}{a^2}=\frac{\sigma^2}{a}\left(\frac{h}{a}\right)=c\left(\frac{h}{a}\right),\end{aligned} \tag{5.18}$$

since $c = \sigma_Y^2/a = C(0)$.

Circular model

The formula for the circular variogram is

$$\gamma(h)=\begin{cases} c\left\{1-\frac{2}{\pi}\cos^{-1}\left(\frac{h}{a}\right)+\frac{2h}{\pi a}\sqrt{1-\frac{h^2}{a^2}}\right\} & \text{for } h\le a,\\ c & \text{for } h>a.\end{cases} \tag{5.19}$$

The parameters c and a are again the sill and range. The function curves tightly as it approaches the range (see Figure 5.4(b)) and its gradient at the origin is $4c/\pi a$. It is CNSD in $\mathbb{R}^1$ and $\mathbb{R}^2$, but not in $\mathbb{R}^3$.

This model can be derived in a way analogous to that of the bounded linear model from the area of intersection, A, of two discs of diameter a, the centres of which are separated by distance h. Matérn (1960) did this by considering the densities with which points are distributed at random by a Poisson process in two overlapping circles. This area is

$$A=\begin{cases}\frac{1}{2}a^2\cos^{-1}\left(\frac{h}{a}\right)-\frac{h}{2\pi}\sqrt{a^2-h^2} & \text{for } h\le a,\\ 0 & \text{for } h>a.\end{cases} \tag{5.20}$$

If we express this as a fraction of the area, $\pi a^2/4$, of one of the circles, in the same way as we expressed the fraction of the linear filter that overlapped along the line above, then we obtain the autocorrelation for the separation:

$$\rho(h)=\frac{2}{\pi}\left\{\cos^{-1}\left(\frac{h}{a}\right)-\frac{h}{a}\sqrt{1-\frac{h^2}{a^2}}\right\} \quad \text{for } h\le a. \tag{5.21}$$

Then from relation (4.14) the variogram, equation (5.19) above, follows.

Spherical model

By a similar line of reasoning we can derive the three-dimensional analogue of the circular model to obtain the spherical correlation function and variogram. The volume of intersection of two spheres of diameter a with their centres h apart is

$$V=\begin{cases}\frac{\pi}{4}c\left(\frac{2}{3}a^3-a^2h+\frac{1}{3}h^3\right) & \text{for } h\le a,\\ 0 & \text{otherwise.}\end{cases} \tag{5.22}$$

The volume of a sphere is $\frac{1}{6}\pi a^3$, and so dividing by it gives the autocorrelation

$$\rho(h)=\begin{cases}1-\frac{3h}{2a}+\frac{1}{2}\left(\frac{h}{a}\right)^3 & \text{for } h\le a,\\ 0 & \text{for } h>a,\end{cases} \tag{5.23}$$

and the variogram is

$$\gamma(h) = \begin{cases} c\left\{\dfrac{3h}{2a} - \dfrac{1}{2}\left(\dfrac{h}{a}\right)^3\right\} & \text{for } h \le a, \\ c & \text{for } h > a. \end{cases} \tag{5.24}$$

The spherical model seems the obvious one to describe variation in three-dimensional bodies of rock, and it has proved well suited to them. It would seem less obviously suited for describing the variation in one and two dimensions, which is usually what is needed in soil and land resource survey. Yet it nearly always fits experimental results from soil sampling better than the one- and two-dimensional analogues. The function curves more gradually than they do Figure 5.4(c), and the reason is probably that there are additional sources of variation at other scales that it can represent. Its gradient at the origin is $3c/2a$. It is CNSD in $\mathbb{R}^2$ and $\mathbb{R}^1$ as well as in $\mathbb{R}^3$.

The spherical function is one of the most frequently used models in geostatistics, in one, two and three dimensions. It represents transition features that have a common extent and which appear as patches, some with large values and others with small ones. The average diameter of the patches is represented by the range of the model. One can see this interpretation by simulating a large field of values using the function as the generator. Figures 5.5 and 5.6(a) are examples in which values have been simulated on a 256×256 square grid with unit interval. The model had a sill variance, $c = 1.0$, and ranges of $a = 15, 25$ and 50 units in Figures 5.5(a), 5.5(b) and 5.6(a), respectively. The maps show that the extents of the patches with large and small values increase as the range increases. The patches have a fairly regular form.

Pentaspherical model

Following Matérn (1960), McBratney and Webster (1986) extended the line of reasoning to obtain the five-dimensional analogue of the above, the pentaspherical function:

$$\gamma(h) = \begin{cases} c\left\{\dfrac{15}{8}\dfrac{h}{a} - \dfrac{5}{4}\left(\dfrac{h}{a}\right)^3 + \dfrac{3}{8}\left(\dfrac{h}{a}\right)^5\right\} & \text{for } h \le a, \\ c & \text{for } h > a\ . \end{cases} \tag{5.25}$$

It is useful in that its curve is somewhat more gradual than that of the spherical model Figure 5.4(d). Its gradient at the origin is $15c/8a$. Again it is CNSD in $\mathbb{R}^1$, $\mathbb{R}^2$ and $\mathbb{R}^3$.

Exponential model

A function that is also much used in geostatistics is the negative exponential:

$$\gamma(h) = c\left\{1 - \exp\left(-\frac{h}{r}\right)\right\}, \tag{5.26}$$

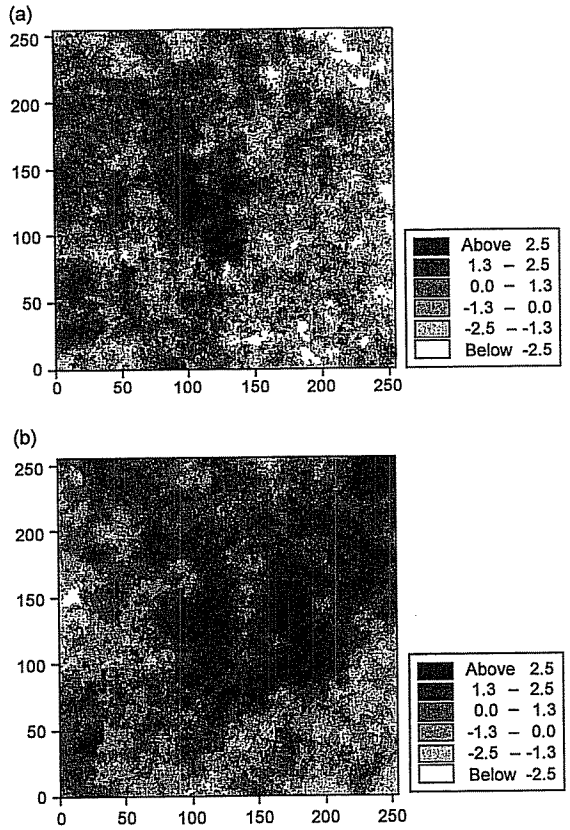


Figure 5.5 Simulated fields of values from spherical functions, equation (5.24), with distance parameters (a) $a = 15$, (b) $a = 25$.

with sill c, and a distance parameter, r, that defines the spatial extent of the model. The function approaches its sill asymptotically, and so it does not have a finite range. Nevertheless, for practical purposes it is convenient to assign it an effective range, and this is usually taken as the distance at which γ equals 95% of the sill variance, approximately $3r$. Its slope at the origin is c/r. Figure 5.7(a) shows it.

The function has an important place in statistical theory. It represents the essence of randomness in space. It is the variogram of first-order autoregressive and Markov processes. Its equivalent autocorrelation function has been the basis of several theoretical studies of the efficiency of sampling designs by, for example, Cochran (1946), Yates (1948), Quenouille (1949) and Matérn (1960). We should expect variograms of this form where differences in soil

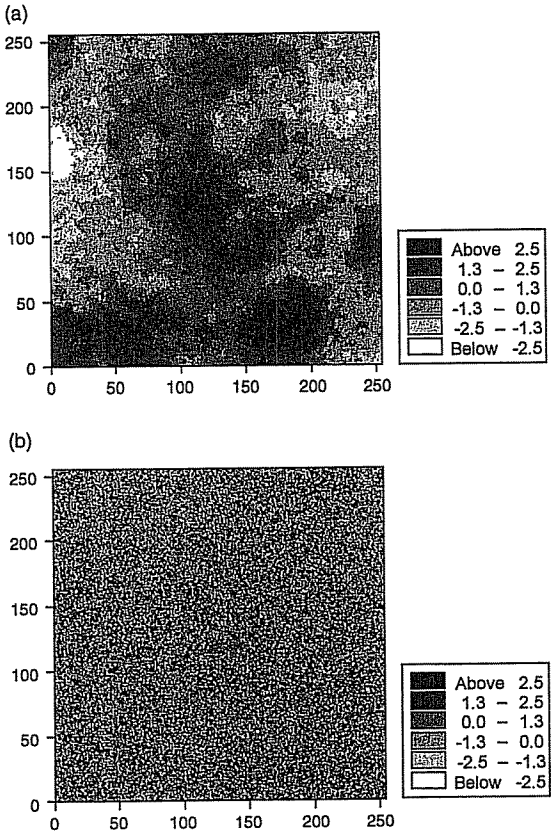


Figure 5.6 Simulated fields of values using: (a) a spherical function, equation (5.24), with distance parameter $a = 50$; (b) a pure nugget variogram, equation (5.33).

type are the main contributors to soil variation and where the boundaries between types occur at random as a Poisson process. Burgess and Webster (1984) found this to be the situation in many instances. If the intensity of the process is η then the mean distance between boundaries is $\bar{d} = 1/\eta$ and the variogram is

$$\begin{aligned} \gamma(h) &= c\{1 - \exp(-h/\bar{d})\} \\ &= c\{1 - \exp(-\eta h)\}. \end{aligned} \tag{5.27}$$

Put another way, this is the variogram of a transition process in which the structures have random extents.

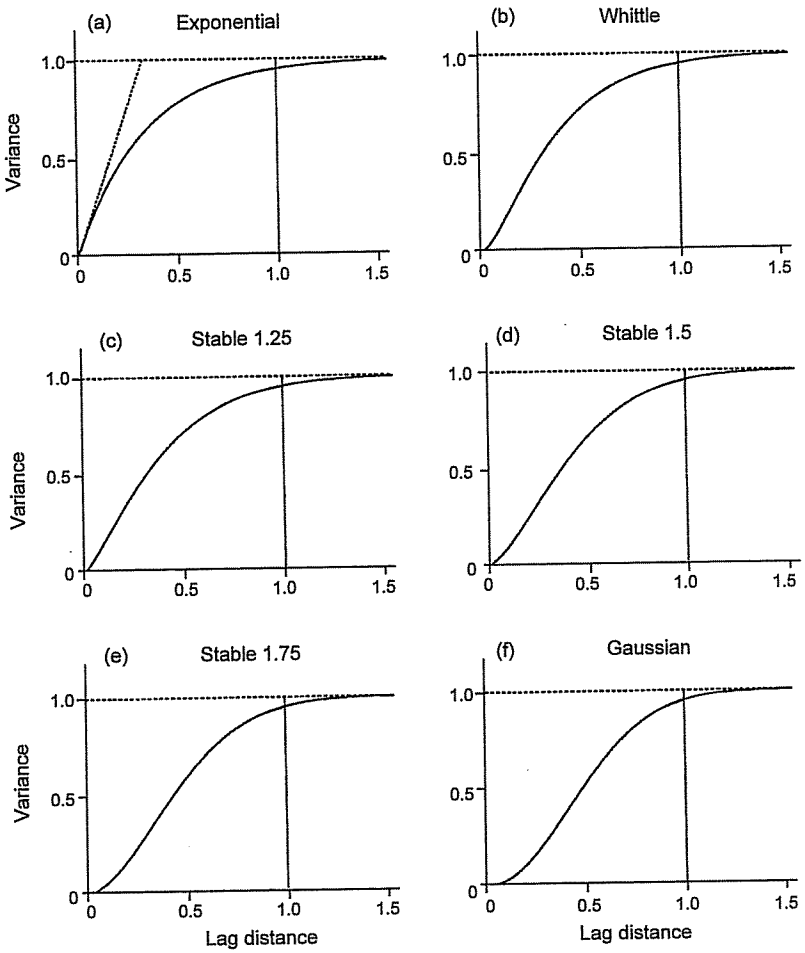


Figure 5.7 Models with asymptotic bounds. All are scaled so that the effective range where the function reaches 0.95 of its sill is approximately 1, marked by the vertical lines on the graphs. (a) $\alpha = 1$ (exponential), $r = 0.333$; (b) Whittle, $r = 0.25$; (c) $\alpha = 1.25$ (stable), $r = 0.416$; (d) $\alpha = 1.5$ (stable), $r = 0.478$; (e) $\alpha = 1.75$ (stable), $r = 0.533$; (f) $\alpha = 2$ (Gaussian), $r = 1/\sqrt{3}$.

Simulated fields obtained from an exponential function with an asymptote approaching 1.0 and distance parameters, r, of 5 and 16 are shown in Figure 5.8(a) and 5.8(b), respectively. The patches of large and small values in the two fields are similarly irregular, but the average sizes of the patches show the different spatial scales of the generator.

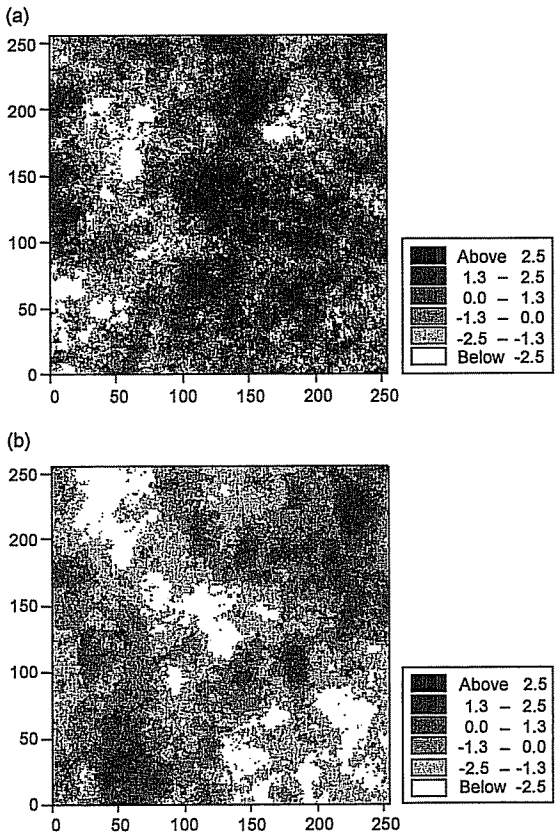


Figure 5.8 Simulated fields of values from exponential functions (equation (5.26)), with distance parameters (a) $r = 5$, (b) $r = 16$.

Whittle's elementary correlation

Whittle (1954) showed that a simple stochastic diffusion process also has an exponential variogram in one and three dimensions. In $\mathbb{R}^2$, however, the process leads to Whittle's *elementary correlation*, given by

$$\gamma(h) = c\left\{1 - \frac{h}{r}K_1\left(\frac{h}{r}\right)\right\}. \qquad (5.28)$$

The parameter c is the sill, as before, the *a priori* variance of the process, r is a distance parameter, and K_1 is the modified Bessel function of the second kind. Like the exponential function, Whittle's function approaches its sill

asymptotically and so has no definite range. Its effective range may be chosen as for the exponential function where the semivariance reaches 95% of the sill, and this is at approximately $4r$. The function approaches the origin with a decreasing gradient, however, and appears slightly sigmoid when plotted, Figure 5.7(b).

Gaussian model

Another function with reverse curvature near the origin recurs again and again in geostatistical texts and software packages. It is the so-called Gaussian model, Figure 5.7(f), with equation

$$\gamma(h) = c\left\{1 - \exp\left(-\frac{h^2}{r^2}\right)\right\}. \qquad (5.29)$$

Once more, c is the sill and r is a distance parameter. The function approaches its sill asymptotically, and it can be regarded as having an effective range of approximately $\sqrt{3}r$ where it reaches 95% of its sill variance.

A serious disadvantage of the model is that it approaches the origin with zero gradient, which we saw above as the limit for random variation and at which the underlying variation becomes continuous and twice differentiable. This can lead to unstable kriging equations, which we present in Chapter 8, and bizarre effects when used for estimation—see Wackernagel (2003) for examples.

In general we deprecate this model. If a variogram appears somewhat sigmoid then we recommend the theoretically attractive Whittle function. Alternatively, if the reverse curvature is stronger you may replace the exponent 2 in equation (5.29) by an additional parameter, say α, with a value less than 2:

$$\gamma(h) = c\left\{1 - \exp\left(-\frac{h^\alpha}{r^\alpha}\right)\right\}. \qquad (5.30)$$

Wackernagel (2003) calls these 'stable models'. Some examples of them are shown in Figure 5.7(c)–(e) with various values of α, and we have used the model with $\alpha = 1.965$ to describe topographic variation (Webster and Oliver, 2006).

Cubic model

Another bounded model with reverse curvature near the origin is the cubic function. Its formula is

$$\gamma(h) = \begin{cases} c\left\{7\left(\frac{h}{a}\right)^2 - 8.75\left(\frac{h}{a}\right)^3 + 3.5\left(\frac{h}{a}\right)^5 - 0.75\left(\frac{h}{a}\right)^7\right\} & \text{for } h \le a, \\ c & \text{for } h > a. \end{cases} \qquad (5.31)$$

The parameter a is a finite range which is approached much more gradually than in the spherical and pentaspherical models.

There are other simple models used in particular disciplines because of their theoretical attractions. Examples include the prismato-gravimetric and prismato-magnetic functions developed in geophysics to model gravimetric and magnetic anomalies (see Armstrong, 1998). If you work in such a special field then you should ask whether there are preferred functions for the particular applications.

Matérn function

The Matérn function is a generalization of several of the functions mentioned above and so appears attractive for this reason. Its formula is

$$\gamma(h) = c\left\{1 - \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{h}{r}\right)^{\nu} \mathrm{K}_{\nu}\left(\frac{h}{r}\right)\right\}. \qquad (5.32)$$

As in the exponential, Whittle and Gaussian models the function has a distance parameter r, and c is the sill. It also has a smoothness parameter, ν, analogous to α in the stable models, equation (5.30), though whereas α is limited to between 0 and 2, ν can vary in the range 0 (very rough) to infinity (very smooth). It includes the special cases of exponential when $\nu = 0.5$ and Whittle's function when $\nu = 1$. Figure 5.9 shows variograms for several values of ν.

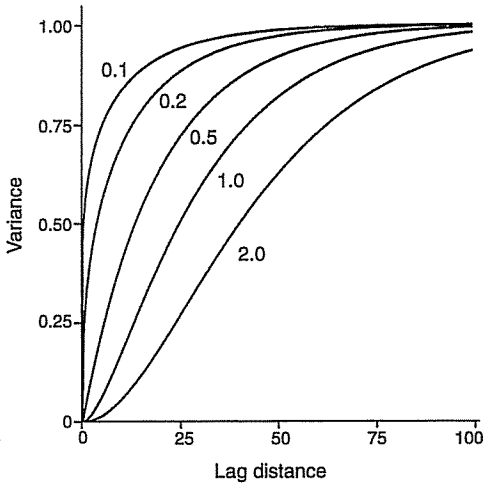


Figure 5.9 The Matérn function (5.32) with *a priori* variance $c = 1$ and distance parameter $r = 20$ and five values of the smoothness parameter ν, giving the five curves. The curve with $\nu = 0.5$ is the exponential and that with $\nu = 1$ is Whittle's function. After Minasny and McBratney (2005).

Unfortunately, when Minasny and McBratney (2005) examined its potential for describing soil properties they had difficulty fitting it to experimental variograms. They found that ν was poorly estimated by the usual method of weighted least squares (see below).

Pure nugget

Although the limiting value 0 of the exponent of equation (5.10) for the power function was excluded because it would give a constant variance, we do need some way of expressing such a constant because that is what appears in practice. We do so by defining a 'pure nugget' variogram as follows:

$$\gamma(h) = c_0\{1 - \delta(h)\}, \qquad (5.33)$$

where c_0 is the variance of the process, and $\delta(h)$ is the Kronecker δ which takes the value 1 when $h = 0$ and is zero otherwise. If the variable is continuous, as almost all properties of the soil and natural environment are, then a variogram that appears as pure nugget has almost certainly failed to detect the spatially correlated variation because the sampling interval was greater than the scale of spatial variation.

Since the nugget variance is constant for all $\mathbf{h}$, $|\mathbf{h}| > 0$, it is usually denoted simply by the variance c_0. Figure 5.6(b) shows the simulated field from a pure nugget variogram. There is no detectable pattern in the variation as there is in Figures 5.5, 5.6(a) and 5.8.

5.3 COMBINING MODELS

As is apparent in Figures 5.3, 5.4 and 5.7, all the above functions have simple shapes. In many instances, however, especially where we have many data, variograms appear more complex, and we may therefore seek more complex functions to describe them. The best way to do this is to combine two or more simple models. Any combination of CNSD functions is itself CNSD. Do not look for complex mathematical solutions the properties of which are unknown.

The most common requirement is for a model that has a nugget component in addition to an increasing, or structured, portion. So, for example, the equation for an exponential variogram with a nugget may be written as

$$\gamma(h) = c_0 + c\left\{1 - \exp\left(-\frac{h}{r}\right)\right\}, \qquad (5.34)$$

and an example is shown in Figure 5.10(a). Figure 5.11 shows the simulated fields for an exponential variogram with parameters $c_0 = 0.333$, $c = 0.667$ and distance parameters, r, of 5 and 16 as before. The speckled appearance within the patches is the result of the nugget variance.

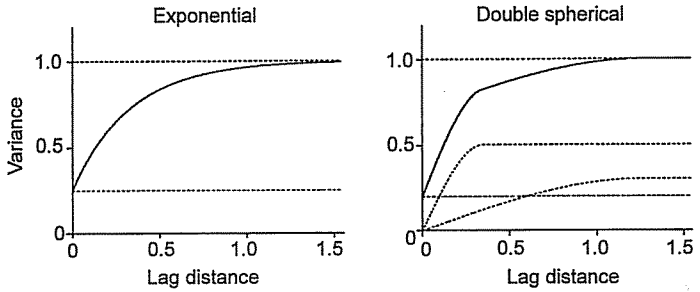


Figure 5.10. Combined (nested) models: (a) single exponential with sill 0.75 plus a nugget variance of 0.25; (b) double spherical with ranges 0.35 and 1.25 and corresponding sills 0.3 and 0.5 plus a nugget variance of 0.2 with the components shown separately.

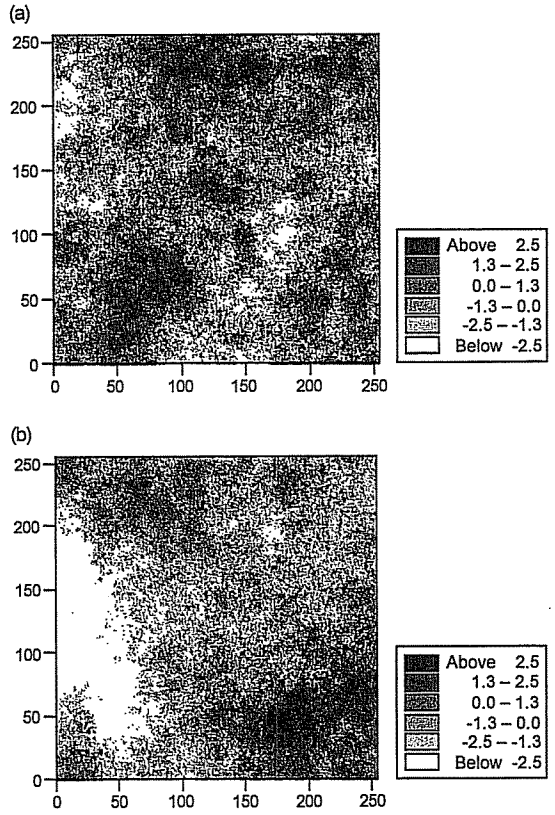


Figure 5.11 Simulated fields of values from exponential functions with nugget variance one-third of the total variance, equation (5.34): with distance parameters (a) $r = 5$; (b) $r = 16$.

Spatial dependence may occur at two distinct scales, and these may be represented in the variogram as two spatial components. The nested spherical, or double spherical, function is the one that has been used most often in these circumstances. Its equation is

$$\gamma(h) = \begin{cases} c_1\left\{\dfrac{3h}{2a_1} - \dfrac{1}{2}\left(\dfrac{h}{a_1}\right)^3\right\} + c_2\left\{\dfrac{3h}{2a_2} - \dfrac{1}{2}\left(\dfrac{h}{a_2}\right)^3\right\} & \text{for } 0 < h \le a_1, \\ c_1 + c_2\left\{\dfrac{3h}{2a_2} - \dfrac{1}{2}\left(\dfrac{h}{a_2}\right)^3\right\} & \text{for } a_1 < h \le a_2, \\ c_1 + c_2 & \text{for } h > a_2, \end{cases} \tag{5.35}$$

where c_1 and a_1 are the sill and range of the short-range component of the variation, and c_2 and a_2 are the sill and range of the long-range component. If it appears to need a nugget then that can be added as a third component, and Figure 5.10(b) shows this combination.

5.4 PERIODICITY

A variogram may seem to fluctuate more or less periodically, rather than increase monotonically, and we might try to describe it with a periodic function. The simplest such function is a sine wave, as shown in Figure 5.12(a), with equation

$$\gamma(h) = W\left\{1 - \cos\left(\frac{2\pi h}{\omega}\right)\right\}, \tag{5.36}$$

where W and ω are the amplitude and length of the wave, respectively.

The gradient at the origin is 0, which, as mentioned above, is undesirable. Usually, however, we find that the periodicity is superimposed on some other source of variation and that the combined model increases from the origin more steeply. Figure 5.12(b) shows an example of it superimposed on an exponential function. An example from actual soil survey is illustrated in Chapter 7.

We might be tempted to move the curve along the abscissa to fit the experimental values so that it increases more nearly linearly from lag 0. We have drawn such a function as the dashed line in Figure 5.12(a). In other words, we have introduced a phase shift, ϕ. If we designate the angle $2\pi h/\omega$ as θ for simplicity then the equation becomes

$$\gamma(h) = W\{1 - \cos(\theta - \phi)\}. \tag{5.37}$$

Unfortunately, the resulting function is not guaranteed to be CNSD, and so the temptation should be resisted.

Equation (5.36) is valid for one dimension only; it is not CNSD in $\mathbb{R}^2$ and $\mathbb{R}^3$. In two and three dimensions the fluctuation must damp, i.e. become less